

Análisis de estrategias de juegos simultáneos de suma no-cero

Daniel Abad Fundora C411
Anabel Benítez González C411
Enzo Rojas D'Toste C411

22 de abril de 2024

Índice

1. Introducción	3
2. ¿Qué es un juego de suma no cero?	4
2.1. Dilema del prisionero	4
2.1.1. Variantes del dilema del prisionero	4
2.2. El dilema de los cazadores	4
2.3. La guerra de los sexos	5
2.4. Dinero Gratis!!!!	5
3. Búsqueda de estrategias óptimas para juegos de un solo turno. Estrategias egoístas	6
3.1. Estrategias básicas	6
3.1.1. Maximizar ganancia mínima	6
3.1.2. Maximizar de ganancia media	6
3.1.3. ¿Qué problema tienen las estrategias anteriores?	6
3.2. Creando estrategias que asumen racionalidad	6
3.2.1. ¿Qué es un equilibrio de Nash?	6
3.2.2. Creando un jugador que intenta alcanzar un equilibrio de Nash	7
4. Juegos iterados	8
4.1. Estrategias reactivas. El ojo por ojo	8
4.1.1. Utilización de lógica difusa para crear un ojo por ojo para matrices no simétricas	8
4.2. Uso de metaheurísticas para la creación estrategias	8
4.2.1. Una metaheurística sencilla	9
4.2.2. Jugador genético	9
4.2.3. Otro jugador genético	10
5. Modelación basada en agentes	11
5.1. Modelando el ambiente	11
5.2. Modelando los agentes	11
5.2.1. Agentes puramente reactivos	11
5.2.2. Agentes reactivos basados en modelos	12
5.2.3. Agentes adaptativos utilizando una arquitectura por capas híbrida	12
6. Análisis de Resultados	14
6.1. Escenario 1: Prisionero iterado sin utilización de jugadores metaheurísticos	14
6.2. Escenario 2: Juegos mixtos con jugadores no metaheurísticos	14
6.3. Escenario 3: Analizando a los jugadores difusos (1ra métrica, envidia)	15
6.4. Escenario 4: Jugadores difusos matrices distintas	17
6.4.1. Añadiendo al DeterministicFuzzyEyeForEye al escenario anterior	17
6.5. Probando otras métricas para el FuzzyEyeForEye	18
6.5.1. Razón entre ganancia total del oponente en la fila elegida y su ganancia total en la fila de máxima ganancia conjunta	18
6.5.2. Razón entre máximas ganancias posibles del oponente	20
6.6. Añadiendo jugadores metaheurísticos	21
6.7. Variando el entrenamiento de las metaheurísticas	23

7. ¿Qué aprendieron a jugar nuestras metaheurísticas?	24
7.1. Probando el modelo con una estrategia conocida	24
7.1.1. Intento de inferir al ojo por ojo	24
7.1.2. Intentando inferir al random	25
7.1.3. Intentando inferir al BadGuy	26
7.2. Análisis del NashGuy por el modelo	26
7.3. Análisis de las estrategias metaheurísticas por el modelo	29
8. El cuentacuentos	33
8.1. Asignándole nombres a los jugadores	33
8.2. Creando matrices y cuentos usando GPT 4	33
8.3. Llevando las matrices y cuentos a código usando GPT 3.5	34
8.4. Usando API de Gemini para generar historias	34
8.4.1. Describiendo rondas	35
8.4.2. Describiendo juegos	37
9. Conclusiones	40
10.Propuestas para futuros trabajos	41
11.Bibliografía	42

1. Introducción

La teoría de juegos es un campo matemático que se ocupa de analizar las interacciones estratégicas entre diferentes actores (jugadores) cuyas decisiones afectan mutuamente sus resultados. En este contexto, se exploran situaciones en las que las elecciones de un jugador no solo dependen de sus propias estrategias, sino también de las estrategias adoptadas por otros jugadores. En este trabajo utilizamos Inteligencia Artificial y Simulación para crear diferentes estrategias "inteligentes" para juegos de suma no cero. Este proyecto surgió inspirado por [este juego](#).

2. ¿Qué es un juego de suma no cero?

En teoría de juegos, un juego de suma no cero es un tipo de juego en el que la cantidad total de recursos o beneficios disponibles para los jugadores no es fija, es decir, la suma de las ganancias y pérdidas de los jugadores no necesariamente se iguala a cero. En estos juegos, los participantes pueden ganar o perder diferentes cantidades de recursos, y no necesariamente se cancelan entre sí. Esto contrasta con los juegos de suma cero, donde las ganancias y pérdidas totales de los jugadores siempre suman cero. Los juegos de suma no cero suelen ser más complejos y permiten una mayor variedad de estrategias y resultados. A diferencia de los juegos de suma cero, en los que, si son de información completa, el algoritmo minimax otorga una estrategia óptima, en los juegos de suma no cero, no es trivial hallar una estrategia óptima, ya que se pueden dar escenarios en que los jugadores cooperen o se enfrenten entre sí para maximizar ganancias. En este trabajo desarrollamos estrategias sobre juegos de suma no cero simultáneos, es decir, ambos jugadores juegan a la vez y no por turnos. A continuación se exponen ejemplos clásicos de juegos de suma no cero simultáneos.

2.1. Dilema del prisionero

Posiblemente el más clásico ejemplo de juego de suma no cero lo constituye el Dilema del prisionero, el problema se plantea de la siguiente forma: Dos delincuentes cometieron un asesinato, la policía los atrapa pero solo poseen evidencias para acusarlos por tenencia de armas de fuego, por lo que se les condena a 3 años de cárcel, cuando por homicidio se les condenaría a 7 años de cárcel. Ambos criminales son arrestados en celdas separadas sin la posibilidad de comunicarse entre sí. El jefe de la policía le hace una propuesta a cada uno de los criminales, si delata a su compañero, su condena se reduce en 2 años

Las posibles decisiones y resultados se resumen en la siguiente matriz

$$\begin{bmatrix} (-5, -5) & (-7, -1) \\ (-1, -7) & (-3, -3) \end{bmatrix}$$

2.1.1. Variantes del dilema del prisionero

Existen variantes con otros trasfondos del dilema del prisionero, por ejemplo el dilema de la caja mágica, aquí el problema se plantea de la siguiente forma: Dos personas, (sin posibilidad de comunicarse entre sí se encuentran con un dispositivo mágico. Si uno de los dos individuos introduce una moneda en este dispositivo, el otro obtendrá 3 monedas. Cada uno esta frente al dilema de si introducir o no una moneda en el dispositivo.

La matriz de decisiones y ganancias quedaría de la siguiente forma:

$$\begin{bmatrix} (3, 3) & (2, -1) \\ (-1, 2) & (0, 0) \end{bmatrix}$$

Nótese que este escenario es equivalente al dilema del prisionero, el mejor escenario para un jugador es que él traicione y el otro colabore, luego que ambos colaboren, luego que ambos traicionen y luego colaborar y ser traicionado por el otro jugador.

2.2. El dilema de los cazadores

Este es un ejemplo clásico de juego de suma no cero relativamente distinto al dilema del prisionero:

En cierta aldea hay dos cazadores. Cada uno posee un equipaje para cazar ciervos y otro para cazar liebres. Un ciervo posee 6 veces la carne de una liebre, pero un solo cazador no es capaz de cazarlo. En el bosque hay dos liebres y un ciervo, si ambos cazadores llevan el equipaje para cazar liebres, cazan una cada uno, si solo uno lo lleva, caza ambas liebres. EL ciervo solo será cazado si ambos llevan el equipaje para ciervos.

La matriz de decisiones y ganancias quedaría de la siguiente forma:

$$\begin{bmatrix} (3, 3) & (0, 2) \\ (2, 0) & (1, 1) \end{bmatrix}$$

2.3. La guerra de los sexos

Otro clásico ejemplo es el de la guerra de los sexos o dilema de la pareja. Se trata de una pareja que decide salir esta noche, sin embargo no llegaron a ponerse totalmente de acuerdo entre si ir a ver el ballet o el boxeo. El hombre prefiere ver el boxeo y la mujer el ballet, pero ambos quieren estar juntos, de hecho, si la otra persona no asiste al mismo lugar, cada uno volverá triste a su casa. No poseen forma de comunicarse entre sí.

Las diferentes decisiones y ganancias quedan representadas en la siguiente matriz

$$\begin{bmatrix} (2, 1) & (0, 0) \\ (0, 0) & (1, 2) \end{bmatrix}$$

2.4. Dinero Gratis!!!!

Es muy sencillo el problema, ambos jugadores encuentran un botón, si ambos votan sí, reciben 10 dólares cada uno, si no, ninguno recibe nada :(

En este caso la matriz de ganancias sería:

$$\begin{bmatrix} (10, 10) & (0, 0) \\ (0, 0) & (0, 0) \end{bmatrix}$$

3. Búsqueda de estrategias óptimas para juegos de un solo turno. Estrategias egoístas

Empecemos por el problema más simple que podemos plantearnos, hallar estrategias para juegos de suma no cero de un solo turno.

3.1. Estrategias básicas

3.1.1. Maximizar ganancia mínima

La estrategia egoísta más básica (aunque racional) que se nos puede ocurrir es la maximización de ganancia mínima, para esto en cada turno el jugador elegirá la fila/columna cuya ganancia mínima sea mayor.

3.1.2. Maximizar de ganancia media

Otra estrategia egoísta básica sería la de maximización de ganancia promedio, para esto se selecciona en cada turno la fila/columna cuya suma de ganancias sea máxima.

3.1.3. ¿Qué problema tienen las estrategias anteriores?

Estas estrategias son "óptimas" quizás contra jugadores aleatorios, pero en la práctica siempre podemos desechar las posibles estrategias del oponente que no tengan ningún sentido que las juegue.

Por ejemplo, en el dilema del prisionero, si un jugador siempre coopera independientemente de lo que haga su oponente, puede ser fácilmente explotado por un oponente que siempre traicione. Por lo tanto, en la práctica, es importante considerar las posibles estrategias del oponente y ajustar las propias estrategias en consecuencia.

3.2. Creando estrategias que asumen racionalidad

3.2.1. ¿Qué es un equilibrio de Nash?

El concepto de equilibrio de Nash es crucial en la teoría de juegos, ya que representa un estado en el que ningún jugador puede mejorar su posición cambiando unilateralmente de estrategia. Para comprenderlo mejor, examinemos los juegos discutidos en el capítulo 2:

El dilema del prisionero:

$$\begin{bmatrix} (-5, -5) & (-7, -1) \\ (-1, -7) & (-3, -3) \end{bmatrix}$$

En este caso, posee un único equilibrio de Nash, ya que hay una columna que es estrictamente menor término a término que la otra.

El dilema de los cazadores:

$$\begin{bmatrix} (3, 3) & (0, 2) \\ (2, 0) & (1, 1) \end{bmatrix}$$

Para este juego, hay varias estrategias que pueden constituir un equilibrio de Nash. Si ambos jugadores eligen la misma acción (ya sea liebre o ciervo), están en un equilibrio de Nash porque no tienen incentivo para cambiar. Además, si ambos jugadores eligen ciervo o liebre con una probabilidad de 0.5, también estarían en un equilibrio de Nash,

ya que cualquier cambio en la estrategia no mejoraría la posición de ninguno de los jugadores.

La guerra de los sexos:

$$\begin{bmatrix} (2, 1) & (0, 0) \\ (0, 0) & (1, 2) \end{bmatrix}$$

Esta matriz también posee 3 equilibrios de Nash, uno si ambos jugadores escogen ir al ballet, otro si ambos escogen ir al boxeo y otro si la mujer escoge ir al ballet con probabilidad $\frac{2}{3}$ y al boxeo con probabilidad $\frac{1}{3}$ y viceversa.

Uno de los teoremas más importantes de la teoría de juegos es el Teorema de Nash, que garantiza que en cualquier juego finito, siempre existe al menos un equilibrio de Nash. Esto proporciona una base sólida para el análisis de juegos y estrategias en una variedad de situaciones.

En general, encontrar estos equilibrios puede ser un problema computacionalmente complejo y costoso.

3.2.2. Creando un jugador que intenta alcanzar un equilibrio de Nash

Hemos creado un jugador que se encarga de buscar un equilibrio de Nash analizando la estrategia del oponente. Las principales características de su estrategia son:

- Valor esperado de las acciones: La estrategia de este jugador se basa en evaluar el valor esperado de cada acción disponible en el juego. El valor esperado representa la suma ponderada de los resultados posibles de una acción, donde las ponderaciones son las probabilidades de que esos resultados ocurran. Es decir, el jugador considera tanto los beneficios potenciales de una acción como la probabilidad de que ocurran esos beneficios.
- Utilización de información pasada: Para calcular el valor esperado de cada acción, el jugador utiliza información recopilada sobre las acciones y resultados pasados del oponente. Esto implica tener en cuenta las acciones que ha tomado el oponente en el pasado y los resultados que han ocurrido como consecuencia.
- Selección de acción óptima: Después de evaluar el valor esperado de cada acción disponible, el jugador selecciona la acción que maximiza este valor. En otras palabras, elige la acción que tiene la mayor expectativa de obtener un resultado favorable, dado el comportamiento pasado del oponente.
- Adaptación y aprendizaje: Este enfoque permite al jugador adaptarse y aprender durante el transcurso del juego, ajustando sus probabilidades y estrategias a medida que recopila más información sobre el comportamiento del oponente y los resultados del juego.

En secciones posteriores veremos que este jugador es el que mejor rendimiento logra alcanzar en juegos reiterados.

4. Juegos iterados

En esta sección complejizamos el problema a uno mucho más interesante, en lugar de jugar una sola ronda, los jugadores juegan un conjunto de rondas consecutivas. Nótese que las estrategias explicadas anteriormente pueden ser aplicadas para estos juegos ¿Pero serían óptimas?

4.1. Estrategias reactivas. El ojo por ojo

La más conocida de las estrategias en juegos iterados es el ojo por ojo, sorprendentemente eficaz para su sencillez, se basa en la siguiente idea: Ser tan bueno/malo con el otro jugador como él lo fue contigo.

En matrices simétricas, este jugador simplemente elige la misma fila/columna que eligió el otro jugador en el turno anterior pero... ¿Como saber qué jugar en un juego no simétrico?

4.1.1. Utilización de lógica difusa para crear un ojo por ojo para matrices no simétricas

Para esto debemos definir el conjunto difuso de jugadas colaborativas para medir qué tan egoísta fue la jugada anterior del otro jugador, para ello creamos diferentes funciones de membresía para evaluar la pertenencia o no a este conjunto y nuestro jugador ojo por ojo jugará en la próxima ronda la jugada que más se ajuste a dicha métrica:

A continuación algunas posibles funciones de membresía que hemos concebido. Digamos en todos los casos que el jugador B es el que analiza la jugada y A es su oponente.

1. Evaluar la relación entre la máxima ganancia obtenida por el jugador B en la columna seleccionada por A y su ganancia en el escenario de máxima ganancia conjunta en la última jugada.
2. Analizar la relación entre las ganancias de ambos jugadores en la casilla con la menor diferencia entre las ganancias en la columna elegida por A en la última ronda.
3. Determinar si la ganancia del jugador B es mayor o igual que su ganancia en el escenario de máxima ganancia conjunta.
4. Calcular la razón entre la suma de las ganancias de A y B en la columna seleccionada por A.
5. Aplicar cualquiera de las anteriores métricas, pero en lugar de a la jugada anterior al promedio de las últimas k jugadas o de todas las jugadas anteriores para obtener una evaluación más completa a lo largo del tiempo.

En todos los casos, el resultado se ajusta al intervalo $[0, 1]$ (se halla el mínimo entre el valor y 1), donde valores cercanos a 0 indican una estrategia egoísta de A que beneficia poco o nada a B.

4.2. Uso de metaheurísticas para la creación estrategias

La idea central detrás de las metaheurísticas es el balance entre dos procesos fundamentales: la exploración y la explotación. La exploración se refiere a la capacidad del algoritmo para investigar amplias áreas del espacio de búsqueda, identificando regiones prometedoras

que podrían contener soluciones óptimas. Por otro lado, la explotación se centra en profundizar dentro de estas regiones prometedoras para afinar y mejorar las soluciones encontradas; el desafío principal es mantener un equilibrio adecuado entre estos dos aspectos.

4.2.1. Una metaheurística sencilla

Una metaheurística que trabaja este concepto de manera sencilla es una que en cada jugada elija entre explorar y explotar con cierta probabilidad, y en ambas aprende del resultado obtenido, por ejemplo de elegir explorar simplemente selecciona una jugada aleatoria, y de elegir explotar juega la jugada que mejores resultados le ha dado. Para saber que jugada ha dado mejores resultados se usa un algoritmo sencillo:

- Se tiene una estructura de datos que dada la matriz y la jugada te dice cuán bien le ha ido en promedio.
- Se tiene otra estructura de datos que guarda la jugada con mejor promedio para cada matriz.

4.2.2. Jugador genético

1. **Población Inicial:** Se comienza con una población inicial de soluciones, cada una representada por un diccionario. En este contexto, las llaves del diccionario corresponden a un fragmento reciente de las jugadas del oponente y un vector que describe la matriz de juego en ese turno. El valor asociado a cada llave es un entero que indica el índice de la fila o columna que el jugador elegirá en respuesta.
2. **Evolución de la Población:**
 - **Generaciones:** El proceso evolutivo se lleva a cabo a través de múltiples iteraciones, denominadas generaciones. En cada generación, la población de soluciones se modifica y evoluciona a partir de la generación anterior.
 - **Función de aptitud (Fitness):** Cada solución se evalúa mediante una función de fitness específica para el problema. En este caso implica jugar n rondas donde, en cada ronda, se genera una nueva matriz y un historial aleatorio de jugadas del oponente. La solución determina la fila a jugar, se simula una respuesta del oponente, y se acumulan puntos según el resultado. El fitness de cada solución se calcula como el total de puntos obtenidos dividido por el número de rondas.
3. **Selección y Reproducción:**
 - **Selección:** Después de evaluar todas las soluciones se selecciona un subconjunto de ellas para la reproducción. Esto puede hacerse de forma aleatoria ponderada por el fitness o seleccionando directamente las soluciones con mejor desempeño.
 - **Cruce:** De las soluciones seleccionadas se elige aleatoriamente pares de "padres" para crear "hijos". La creación de un hijo implica combinar partes de los diccionarios de ambos padres, usualmente partiendo sus estructuras en una posición aleatoria y uniendo estos segmentos.
4. **Mutación:** Una vez generado un nuevo hijo, este puede ser sujeto a mutaciones aleatorias para introducir variabilidad en la población.
5. **Repetición y Conclusión:**

- **Iteración:** El proceso de selección, reproducción y mutación se repite a lo largo de varias generaciones, manteniendo constante el tamaño de la población.
- **Solución Óptima:** Al final de las generaciones se elige la solución con el mayor valor de fitness de la última generación como la mejor solución encontrada.

4.2.3. Otro jugador genético

Hemos desarrollado un nuevo jugador genético que sigue una estrategia similar a la previamente descrita, pero con mejoras sustanciales. Al comienzo del entrenamiento (antes de jugar), obtenemos una población global y una estrategia global a partir de la fase inicial. Durante el juego, después de cada ronda, el jugador analiza las jugadas del oponente y entrena una versión específica para ese torneo. Esta versión se basa en la población global, pero ahora considera todas las estrategias en lugar de solo la mejor. De esta manera, la población evoluciona y se adapta de forma dinámica para cada competición.

5. Modelación basada en agentes

Para realizar simulaciones de los juegos hacemos uso de agentes inteligentes, nuestro ambiente consistirá en el estado actual del juego, que está compuesto por sus jugadores, la estrategia que siguen, sus puntuaciones actuales, así como la lista de matrices (previas y posteriores a la actualidad), la matriz actual y el historial de decisiones de los jugadores. Nuestros agentes serán los jugadores y su función objetivo a maximizar será su puntuación.

5.1. Modelando el ambiente

Nuestro ambiente consta de las siguientes características:

1. Accesibilidad: Nuestro ambiente es poco accesible, cada jugador posee conocimiento acerca de la matriz del turno actual y de su historial, pero desconoce completamente quién es el otro jugador, las matrices futuras, así como el historial de juegos en el que él no está involucrado.
2. Determinista o no determinista: Nuestro ambiente es un ambiente no determinista, ya que, al no conocer la estrategia del oponente, existe incertidumbre acerca del estado del ambiente luego de realizar una acción. Aunque el ambiente en principio es no determinista, se comporta de forma determinista con respecto a las elecciones conjuntas de los jugadores.
3. Episódico o secuencial: Este ambiente es secuencial, ya que las decisiones del agente pueden influir de forma positiva o negativa en el futuro, pues los otros agentes tienen acceso a las jugadas realizadas previamente por sus oponentes y pueden jugar de forma acorde.
4. Estático o dinámico: Es un ambiente estático, ya que permanece inalterable mientras no se realice una acción sobre él, de hecho, ambos jugadores deben realizar su acción para pasar cada ronda.
5. Discreto o continuo: Estamos en presencia de un ambiente discreto. En cada turno, cada jugador puede realizar un número constante de acciones determinadas por la matriz de cada ronda.

Por lo tanto, estamos en presencia de un ambiente relativamente complejo.

5.2. Modelando los agentes

Como ya mencionamos previamente, nuestros agentes tendrán como función objetivo a maximizar su puntuación en el torneo.

Los jugadores pueden clasificarse como distintos tipos de agentes según su estrategia

5.2.1. Agentes puramente reactivos

Los siguientes jugadores fueron modelados como agentes puramente reactivos, ya que sus acciones dependen únicamente de la matriz del juego, no tienen en cuenta jugadas anteriores y juegan de forma determinista sin importar el oponente:

- BadGuy

- GoodGuy
- RandomGuy

5.2.2. Agentes reactivos basados en modelos

Los siguientes jugadores fueron modelados como agentes reactivos basados en modelos. Estos analizan jugadas previas del oponente para decidir qué jugar en el turno actual, ya sea para recompensar/castigar al oponente o para buscar una estrategia óptima según el juego del oponente:

- EyeForEye
- FuzzyEyeForEye
- DeterministicFuzzyEyeForEye

El primero de ellos, simplemente juega lo mismo que el jugador oponente, mientras que los otros, definen métricas a partir de jugadas anteriores y juegan acorde a dichas métricas. Para lograr este objetivo, almacenamos un historial de las jugadas realizadas por cada oponente, de manera que el jugador "recuerde" todo lo que le han jugado. La arquitectura utilizada para la concepción de este agente está basada en la arquitectura de categorización con ciertas diferencias importantes como que en este caso no se disparan comportamientos simultáneos.

5.2.3. Agentes adaptativos utilizando una arquitectura por capas híbrida

Los agentes adaptativos incorporan mecanismos de aprendizaje de forma tal que puedan ajustar dinámicamente sus estrategias y respuestas a lo largo del tiempo. Incorporamos los siguientes agentes con dichas características

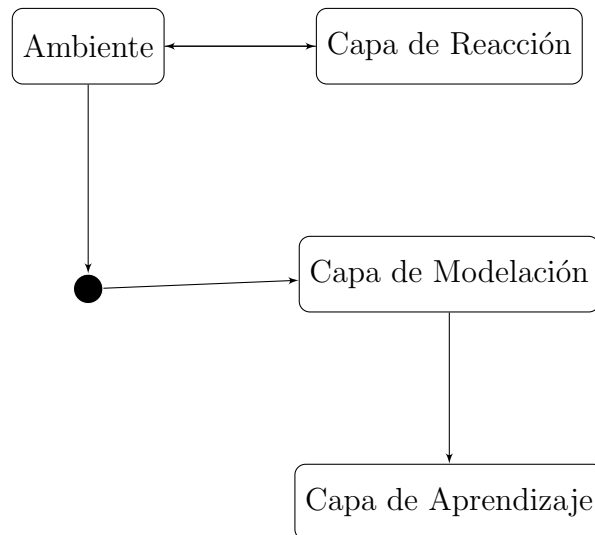
- NashGuy
- SimpleMetaheuristicGuy
- GeneticGuy

El primero de los dos jugadores intenta inferir la estrategia de su oponente dada cierta matriz e intenta jugar en función de la misma. Se obtienen buenos resultados con este agente en juegos donde se repitan matrices.

El segundo y tercer jugador utilizan metaheurísticas para entrenarse antes del juego, jugando según lo aprendido durante el entrenamiento.

Para la implementación de estos agentes utilizamos una arquitectura por capas híbrida. Fueron concebidas las siguientes capas:

- Capa reactiva
- Capa de modelación
- Capa de aprendizaje



El carácter híbrido de esta arquitectura se debe a que combina características de las Touring Machines con características de los InterRap. Ya que la capa reactiva se relaciona horizontalmente con respecto a las otras dos capas, las cuales se relacionan verticalmente entre sí. El ambiente en primer lugar envía la información del estado del juego, a lo cual el agente reacciona inmediatamente a través de la capa reactiva modelada en el método `play()` y luego, analiza la información del cambio en el ambiente luego de su acción, guardando la jugada del adversario en el historial del juego (capa de modelación) y aprendiendo según los resultados obtenidos (capa de aprendizaje).

6. Análisis de Resultados

En esta sección analizamos la eficiencia de las distintas estrategias en diferentes torneos.

6.1. Escenario 1: Prisionero iterado sin utilización de jugadores metaheurísticos

En primer lugar analizamos el siguiente torneo, consistirá en juegos de 32 rondas del dilema del prisionero (presentado en el capítulo 1).

Se introdujeron los siguientes jugadores:

- Random x3
- EyeForEye x4
- BadGuy x2
- GoodGuy x2
- NashGuy x1

Jugador	Estrategia	Puntaje
Gonzalo	Random	954
Nadia	Random	912
Samuel	Random	921
Ignacio	EyeForEye	1090
Hannah	EyeForEye	1090
Pablo	EyeForEye	1077
Julia	EyeForEye	1070
Daniela	BadGuy	1046
Kenia	BadGuy	1043
Carmen	GoodGuy	828
Laura	GoodGuy	804
Helena	NashGuy	1028

Cuadro 1: Resultados del juego "Iterated Prisoner's Dilemma".

Quedando como ganador en EyeForEye y como perdedor el GoodGuy.

6.2. Escenario 2: Juegos mixtos con jugadores no metaheurísticos

En este caso jugamos un torneo con juegos de 24 rondas donde se alternan los escenarios "El dilema de los cazadores", "Dilema del prisionero" y "Concurso de cocina" iterativamente.

Los jugadores fueron:

- Random x3
- EyeForEye x3
- BadGuy x4

- GoodGuy x2
- NashGuy x2

Se obtuvo el siguiente resultado:

Jugador	Estrategia	Puntaje
Amelia	Random	747
Omar	Random	694
Quintín	Random	720
Julia	EyeForEye	796
Gustavo	EyeForEye	793
Mónica	EyeForEye	793
Emilio	BadGuy	793
Nadia	BadGuy	779
Gabriela	BadGuy	775
Diego	BadGuy	761
Eva	GoodGuy	617
María	GoodGuy	624
Beatriz	GoodGuy	623
Uriel	NashGuy	817
Verónica	NashGuy	829

Cuadro 2: Resultados del juego "Hunter-Prisoner-Cooker x 8".

Resultando ganador el NashGuy y como perdedor el GoodGuy .

6.3. Escenario 3: Analizando a los jugadores difusos (1ra métrica, envidia)

A la lista de jugadores anteriores añadimos 3 jugadores FuzzyEyeForEye utilizando la métrica envy y lo analizamos en un torneo cuyos juegos consisten en 8 iteraciones de "El dilema de los cazadores", "Dilema del prisionero", "Concurso de cocina" y "Decisiones en el mercado de acciones" por lo que serían un total de 32 rondas por juego.

Se obtuvieron los siguientes resultados:

Jugador	Estrategia	Puntaje
Pilar	Random	1413
Iván	Random	1451
Rosa	Random	1420
Natalia	EyeForEye	1479
Sofía	EyeForEye	1440
Samuel	EyeForEye	1469
Zoe	BadGuy	1495
Penélope	BadGuy	1525
Xenia	BadGuy	1509
Verónica	BadGuy	1501
Flavia	GoodGuy	1222
Javier	GoodGuy	1232
Quirino	GoodGuy	1194
Rocío	NashGuy	1622
Zacarías	NashGuy	1597
Zara	FuzzyEyeForEye	1379
Carmen	FuzzyEyeForEye	1368
Nuria	FuzzyEyeForEye	1331

Cuadro 3: Resultados del torneo "Hunter-Prisoner-Cooker-ActionMarket".

Como se puede notar, el mejor jugador continúa siendo el NashGuy, nótese también que los FuzzyEyeForEye obtienen resultados ligeramente inferiores al EyeForEye, lo cual tiene sentido ya que estamos analizando juegos simétricos.

6.4. Escenario 4: Jugadores difusos matrices distintas

En esta sección jugamos con los mismos jugadores del torneo anterior, pero utilizando 25 matrices diferentes.

Se obtuvo el siguiente resultado:

Jugador	Estrategia	Puntaje
Ana	Random	2354
Rocío	Random	2386
Néstor	Random	2339
Yolanda	EyeForEye	3441
Zoe	EyeForEye	3459
Ignacio	EyeForEye	3487
Julia	BadGuy	3651
Carlos	BadGuy	3708
Olga	BadGuy	3716
Kai	BadGuy	3680
Felipe	GoodGuy	3470
Eva	GoodGuy	3459
Helena	GoodGuy	3484
Oriol	NashGuy	1926
Juan	NashGuy	1947
Alejandro	FuzzyEyeForEye	3098
Carmen	FuzzyEyeForEye	3140
Alberto	FuzzyEyeForEye	3154

Cuadro 4: Resultados del torneo "Every story once".

Nótese que en este caso el NashGuy, que era el mejor jugador hasta el momento quedó en último lugar, esto se debe a que, al ser distintas todas las matrices, no tiene forma de aprender la estrategia idónea para cada matriz, quedando esta vez el Badguy como ganador del torneo (recordar que al no haber visto previamente ninguna de estas matrices, el EyeForEye juega de forma similar al GoodGuy)

6.4.1. Añadiendo al DeterministicFuzzyEyeForEye al escenario anterior

Probamos añadir al escenario anterior el DeterministicFuzzyEyeForEye, tiene un funcionamiento diferente al FuzzyEyeForEye, analiza la jugada anterior bajo la misma métrica, si esta es menor que 0.7 juega como BadGuy, si es mayor juega como GoodGuy.

Aunque se obtienen mejores resultados, el BadGuy sigue siendo el ganador del torneo

Jugador	Estrategia	Puntaje
Uriel	Random	3595
Yolanda	Random	3620
Walter	Random	3512
Sara	EyeForEye	5261
Quintín	EyeForEye	5195
Jacinta	EyeForEye	5231
Patricia	BadGuy	5626
William	BadGuy	5582
Karen	BadGuy	5569
Nadia	BadGuy	5600
Daniela	BadGuy	5530
Bianca	BadGuy	5546
Olga	BadGuy	5624
Leandro	GoodGuy	5258
Néstor	GoodGuy	5262
Rosa	GoodGuy	5286
Kevin	NashGuy	2738
Oriol	NashGuy	2759
Elsa	FuzzyEyeForEye	4772
Isabel	FuzzyEyeForEye	4828
Óscar	FuzzyEyeForEye	4808
Eduardo	FuzzyDeterministicEyeForEye	5295
Xavier	FuzzyDeterministicEyeForEye	5323
Héctor	FuzzyDeterministicEyeForEye	5315
Javier	FuzzyDeterministicEyeForEye	5304
Miguel	FuzzyDeterministicEyeForEye	5311

Cuadro 5: Resultados del torneo "Adding DeterministicFuzzyGuy"

6.5. Probando otras métricas para el FuzzyEyeForEye

En esta sección simulamos el mismo escenario, pero cambiamos la función de membresía al conjunto de jugadas colaborativas.

6.5.1. Razón entre ganancia total del oponente en la fila elegida y su ganancia total en la fila de máxima ganancia conjunta

Cambiando esta métrica y simulando un torneo con las características del anterior se obtuvieron los siguientes resultados:

Jugador	Estrategia	Puntaje
Bianca	Random	3547
Alejandro	Random	3528
Sergio	EyeForEye	5330
Florencia	EyeForEye	5370
Uriel	EyeForEye	5363
Zoe	BadGuy	5709
Yasmin	BadGuy	5678
Raúl	BadGuy	5711
Natalia	BadGuy	5694
Luna	BadGuy	5730
Óliver	BadGuy	5721
Hannah	GoodGuy	5372
Nadia	GoodGuy	5328
Zacarías	GoodGuy	5390
Mónica	NashGuy	2799
César	NashGuy	2800
Yolanda	FuzzyEyeForEye	5153
Laura	FuzzyEyeForEye	5200
Carmen	FuzzyEyeForEye	5181
Elena	FuzzyEyeForEye	5185
David	FuzzyDeterministicEyeForEye	5422
Mariano	FuzzyDeterministicEyeForEye	5416
Eva	FuzzyDeterministicEyeForEye	5452
Flavia	FuzzyDeterministicEyeForEye	5403
Waldo	FuzzyDeterministicEyeForEye	5393
Xavier	FuzzyDeterministicEyeForEye	5446

Cuadro 6: Resultados del torneo "Analizing Second Metric".

Como se podrá observar, no se obtienen grandes mejoras con respecto a la métrica anterior

6.5.2. Razón entre máximas ganancias posibles del oponente

Procedemos de igual manera al epígrafe anterior, cambiando solamente la métrica del ojo por ojo a una tercera.

Jugador	Estrategia	Puntaje
Nora	Random	3668
Tatiana	Random	3794
Yasmin	EyeForEye	5340
Lorena	EyeForEye	5408
Dante	EyeForEye	5414
Zoe	BadGuy	5676
Adrián	BadGuy	5742
Tobías	BadGuy	5697
Ignacio	BadGuy	5689
Yolanda	BadGuy	5733
Hannah	BadGuy	5719
Alejandro	GoodGuy	5384
Kevin	GoodGuy	5377
Hilda	GoodGuy	5439
Violeta	NashGuy	2850
Emilio	NashGuy	2863
Pablo	FuzzyEyeForEye	5405
Paula	FuzzyEyeForEye	5379
Oriol	FuzzyEyeForEye	5431
Carlos	FuzzyEyeForEye	5436
Kenia	FuzzyDeterministicEyeForEye	5434
María	FuzzyDeterministicEyeForEye	5456
Gloria	FuzzyDeterministicEyeForEye	5480
Mónica	FuzzyDeterministicEyeForEye	5465
Penélope	FuzzyDeterministicEyeForEye	5395
William	FuzzyDeterministicEyeForEye	5445

Cuadro 7: Resultados del torneo "Analizing Third Metric".

Podemos observar que las tres métricas se comportan de manera similar en estos casos.

6.6. Añadiendo jugadores metaheurísticos

Añadimos en esta sección al SimpleMetaheuristic Guy y al genetic Guy Volviendo al caso más sencillo, del Dilema del Prisionero Iterado (con 1000 rondas)

Jugador	Estrategia	Puntaje
Camila	GoodGuy	20592
Eva	Random	25397
Ana	Random	25395
Sergio	EyeForEye	31601
Jacinta	EyeForEye	31577
Leandro	EyeForEye	31636
Inés	BadGuy	31279
Bianca	BadGuy	31321
Sara	GeneticGuy	20488
Ulises	GeneticGuy	31385
Violeta	SimpleMetaHeuristicGuy	30698
Lorena	SimpleMetaHeuristicGuy	30403

Cuadro 8: Resultados del torneo "Prisoner's with metaheuristics".

En este torneo podemos observar que tanto los SimpleMetaHeuristicGuy como uno de los GeneticGuy obtuvieron resultados competitivos quedando cerca del EyeForEye (estrategia más óptima en este escenario), sin embargo el otro GeneticGuy dejó mucho que desear, quedando en último lugar, incluso por debajo del GoodGuy.

Analicemos los resultados en el escenario visto anteriormente de "El dilema de los cazadores", "Dilema del prisionero" y "Concurso de cocina" x 8.

Jugador	Estrategia	Puntaje
Javier	GoodGuy	638
Luna	Random	607
Uriel	Random	644
Beatriz	EyeForEye	719
Tatiana	EyeForEye	710
Luis	EyeForEye	700
Diego	NashGuy	673
Pablo	BadGuy	687
Natalia	BadGuy	659
Eduardo	GeneticGuy	586
Raúl	GeneticGuy	681
Kenia	SimpleMetaHeuristicGuy	694
Néstor	SimpleMetaHeuristicGuy	657

Cuadro 9: Resultados del torneo "Hunter-Prisoner-Cooker x 8".

Nuestra experimentación reveló que, en promedio, el SimpleMetaHeuristicGuy supera al GeneticGuy, cosa que se puede observar en los ejemplos previamente presentados.

Jugador	Estrategia	Puntaje
Alejandro	GoodGuy	2306
Camila	Random	1707
Valentina	Random	1604
Dante	EyeForEye	2228
Flavia	EyeForEye	2240
Lorena	EyeForEye	2274
Laura	NashGuy	1316
Gabriela	BadGuy	2428
Leandro	BadGuy	2472
Ulises	GeneticGuy	1589
Quintín	GeneticGuy	1523
Zoe	SimpleMetaHeuristicGuy	1582
Julia	SimpleMetaHeuristicGuy	2256

Cuadro 10: Resultados del torneo "Every story once with metaheuristics".

Vemos que en este último torneo, el cual consiste en juegos de 25 matrices distintas entre sí, nuestras metaheurísticas presentan resultados inferiores a los agentes puramente reactivos, lo cual denota que el funcionamiento de las estrategias aprendidas por las metaheurísticas se basan parcialmente en el análisis de sus jugadas anteriores con los otros jugadores.

6.7. Variando el entrenamiento de las metaheurísticas

Variamos nuestra metaheurística para que aprenda durante el torneo, el entrenamiento de la metaheurística anterior lo guardamos como población global y estrategia global, pero ahora además, luego de cada ronda, vemos qué nos jugó el oponente y entrenamos una versión que se desprende de la global pero es única de cada torneo. Evaluamos los resultados dado que hubiese jugado con cada estrategia de la población y así el jugador evoluciona durante el torneo.

Se obtuvieron los siguientes resultados, en este caso simulamos usando 10 dilemas del prisionero y luego 10 dilemas del cazador, obteniéndose los siguientes resultados

Jugador	Estrategia	Puntaje
Carlos	GoodGuy	427
Julia	RandomGuy	401
Nora	RandomGuy	400
William	EyeForEye	506
Sergio	EyeForEye	512
Paula	EyeForEye	502
Néstor	BadGuy	469
Olga	BadGuy	474
Ramón	AnotherGeneticGuy	484
Bianca	AnotherGeneticGuy	465
Ciro	GeneticGuy	478
María	GeneticGuy	416

Cuadro 11: Resultados del torneo "Prisoner and hunter with learning genetic".

Luego simulamos con 10 prisioneros y se obtuvieron los siguientes resultados:

Jugador	Estrategia	Puntaje
Ximena	GoodGuy	268
Oriol	RandomGuy	288
Patricia	RandomGuy	289
David	EyeForEye	331
Juan	EyeForEye	334
Diego	EyeForEye	330
Waldo	BadGuy	361
Zacarías	BadGuy	358
Omar	AnotherGeneticGuy	351
Pilar	AnotherGeneticGuy	342
Paula	GeneticGuy	252
Karen	GeneticGuy	268

Cuadro 12: Resultados del torneo "Every story once with metaheuristics".

En los casos probados este nuevo tipo de genético obtiene resultados, obteniendo en ambos casos resultados cercanos al ganador del torneo. Sin embargo, este nuevo jugador es muy costoso computacionalmente, la primera de las simulaciones mostradas demoró 12 minutos y la segunda 8 minutos

7. ¿Qué aprendieron a jugar nuestras metaheurísticas?

Con el objetivo de analizar las estrategias que aprendieron a jugar nuestras metaheurísticas utilizamos un modelo de lenguaje (Gemini), al cual le enviamos el historial del torneo para que explique en lenguaje natural la estrategia del jugador.

7.1. Probando el modelo con una estrategia conocida

Para ajustar las consultas y probar la eficacia de este enfoque en primer lugar intentamos que explique estrategias de jugadores de los cuales conozcamos previamente la estrategia.

7.1.1. Intento de inferir al ojo por ojo

Enviándole el historial de un torneo de 64 rondas del dilema del prisionero contra un jugador Random, se obtuvieron los siguientes resultados:

Estrategia del Jugador:

El jugador está utilizando una estrategia de "Tit for Tat" modificada.

Bajo esta estrategia:

- En la primera ronda, el jugador coopera (fila 1).
- En las rondas posteriores, el jugador coopera si su oponente cooperó en la ronda anterior y no coopera si su oponente no cooperó.

Explicación:

- Primera ronda: El jugador siempre coopera en la primera ronda.
- Rondas posteriores: El jugador sigue cooperando si su oponente ha cooperado en rondas pasadas, por ejemplo, en las rondas 2, 4, 5, 8, 10, 12, 14, 16, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 y 47.
- El jugador deja de cooperar si su oponente no coopera, por ejemplo, en las rondas 3, 6, 7, 9, 11, 13, 15, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 49 y 50.

Esta estrategia alienta la cooperación mutua al castigar la no cooperación y recompensar la cooperación.

Estrategia de "tit-for-tat"

Razón:

El jugador elige la misma acción que el oponente tomó en el turno anterior. Esto es evidente en el historial de decisiones:

- Turno 1: El oponente juega "izquierda" (2). El jugador también juega "izquierda" (2).
- Turno 2: El oponente juega "derecha" (5). El jugador también juega "derecha" (5).
- Turno 3: El oponente juega "izquierda" (0). El jugador también juega "izquierda" (0).
- Turno 4: El oponente juega "derecha" (4). El jugador también juega "derecha" (4).

Estrategia del jugador:

El jugador parece estar siguiendo una estrategia de "ojo por ojo".

Justificación:

El jugador elige la misma fila que el jugador anterior en cada iteración:

- Jugador 1: Fila 1 (Cooperar)
- Jugador 2: Fila 1 (Cooperar)
- Jugador 1: Fila 2 (Desertar)
- Jugador 2: Fila 2 (Desertar)

Esta estrategia implica que el jugador coopera si su oponente cooperó en la ronda anterior y deserta si su oponente desertó. El objetivo es penalizar a los oponentes que no cooperan y recompensar a los que sí lo hacen.

En la matriz proporcionada, la estrategia de "ojo por ojo" es una estrategia evolutivamente estable, ya que fomenta la cooperación mutua a largo plazo. Ambos jugadores reciben una recompensa de 4 por cooperar en cada iteración, lo que supera la recompensa de 5 por desertar una vez y luego recibir 0 en iteraciones futuras.

Aunque comete ciertos errores en sus explicaciones, se puede notar que es relativamente bueno reconociendo este tipo de estrategias. Naturalmente varias veces identificó que la estrategia de este jugador era la de su oponente.

7.1.2. Intentando inferir al random

Analicemos su capacidad de inferir al jugador random, una estrategia notablemente más impredecible que el ojo por ojo:

Estrategia del Jugador:

El jugador está utilizando una estrategia de mezcla aleatoria.

Justificación:

El jugador elige las filas al azar con igual probabilidad. Esto se puede ver en el historial de decisiones:

- En la primera decisión, el jugador eligió la fila 1 con una probabilidad del 50
- En la segunda decisión, el jugador eligió la fila 2 con una probabilidad del 50

Al mezclar aleatoriamente sus elecciones, el jugador reduce la capacidad del oponente para predecir su siguiente movimiento y, por lo tanto, maximiza su ganancia media a largo plazo.

Aunque el modelo en varias ocasiones logró inferir la estrategia del random, en muchos otros casos se equivocó y dio una interpretación errónea, pues no describía la secuencia real de jugadas, por ejemplo decía que siempre elegía la primera fila cuando en realidad no lo hacía.

7.1.3. Intentando inferir al BadGuy

Nuestro BadGuy sigue una estrategia de maximización de ganancia mínima, elige siempre la fila donde su mínima ganancia sea mayor

Análisis de Gemini:

Estrategia del Jugador:

El jugador elige consistentemente la fila superior (fila 0) en todas las matrices.

Razón:

El jugador parece estar siguiendo una estrategia maximin, que implica seleccionar la fila con el menor peor rendimiento posible. En este caso, la fila superior tiene el peor rendimiento de 2 en ambos casos, mientras que la fila inferior tiene un peor rendimiento de 5 en un caso y 4 en el otro.

Al elegir siempre la fila superior, el jugador minimiza su posible pérdida máxima y asegura un pago decente de 2 en cada ronda. Este enfoque es conservador y reduce el riesgo de enfrentar pérdidas significativas. Sin embargo, también limita las posibles ganancias del jugador, ya que la fila inferior ofrece el potencial de pagos más altos en algunos casos.

En este caso logró inferir satisfactoriamente la estrategia.

7.2. Análisis del NashGuy por el modelo

Hagamos uso del LLM para analizar la estrategia del NashGuy

Primero lo hicimos jugar al dilema del prisionero con el jugador random, el LLM llegó a las siguientes conclusiones

Respuesta 1:

Estrategia Aparente del Jugador:

El jugador parece estar siguiendo una estrategia de "seguir al líder", en la que elige la misma fila o columna que su oponente en el round anterior.

Razonamiento:

Esta estrategia es común en juegos donde los jugadores son reacios al riesgo y prefieren evitar pérdidas seguras. Al seguir al líder, el jugador reduce el riesgo de elegir una opción que resulte en una pérdida, ya que ya ha sido probada por su oponente.

En este caso específico, el jugador elige la misma fila que su oponente en el 98 % de los casos (155 de 158 rondas). Esto sugiere que el jugador es muy dependiente de su oponente para tomar decisiones y no tiene una estrategia propia bien definida.

Estrategia Aparente del Oponente:

El oponente del jugador parece estar utilizando una estrategia de "columna aleatoria", en la que elige una columna al azar en cada ronda.

Razonamiento:

Esta estrategia es común entre los jugadores que desean minimizar sus pérdidas y maximizar sus ganancias potenciales. Al elegir una columna al azar, el jugador evita caer en un patrón predecible que pueda ser explotado por su oponente.

En este caso específico, el oponente elige una columna al azar en el 81 % de los casos (128 de 158 rondas). Esto sugiere que el oponente está tratando de evitar ser predecible y maximizar sus posibilidades de obtener un resultado positivo.

Conclusión:

En general, el jugador está siguiendo una estrategia de bajo riesgo al seguir a su oponente, mientras que el oponente está utilizando una estrategia de mayor riesgo al elegir columnas al azar. La estrategia del jugador es efectiva para reducir el riesgo, pero también limita sus ganancias potenciales. La estrategia del oponente es más arriesgada, pero también tiene el potencial de mayores recompensas.

Respuesta 2:

Estrategia del Jugador:

El jugador parece estar utilizando una estrategia de "promedio": alternando entre elegir la fila superior e inferior de la matriz. Esto se desprende de que el jugador elige la fila superior en los rounds pares (2, 4, 6, etc.) y la fila inferior en los rounds impares (1, 3, 5, etc.).

Justificación de la Estrategia:

La estrategia de promedio se basa en la idea de que elegir una fila aleatoria en una matriz de dos filas minimiza la pérdida esperada. Esta estrategia se considera un comportamiento racional en juegos repetidos, ya que evita que el jugador sea explotado por un oponente que anticipa sus movimientos.

Al alternar entre las dos filas, el jugador garantiza que sus ganancias sean relativamente estables, ya que es poco probable que elija la misma fila que su oponente dos veces seguidas. Además, la estrategia ayuda al jugador a evitar grandes pérdidas, ya que las dos filas ofrecen pagos similares.

Posible Estrategia del Oponente:

El oponente del jugador también parece estar utilizando una estrategia de promedio, pero alternando entre las dos columnas de la matriz. Esto se observa en que el oponente elige la columna izquierda en los rounds pares (2, 4, 6, etc.) y la columna derecha en los rounds impares (1, 3, 5, etc.).

Al utilizar una estrategia similar, ambos jugadores evitan ser explotados por el otro y mantienen un nivel relativamente constante de ganancias.

Analizando un juego contra el ojo por ojo el LLM dió la siguiente respuesta:

Estrategia del jugador:

El jugador parece estar usando una "estrategia de alternar columnas", seleccionando consistentemente la columna opuesta a la elegida por su oponente en el round anterior.

Fundamento de la estrategia:

Esta estrategia es efectiva en este juego porque:

- Evita la explotación: Si el jugador selecciona la misma columna repetidamente, el oponente podría aprender este patrón y seleccionar la otra columna en cada turno.
- Crea incertidumbre: Al alternar columnas, el jugador hace que sea más difícil para el oponente predecir sus selecciones y, por lo tanto, reduce las posibilidades de que sea explotado.
- Maximiza la ganancia esperada: A largo plazo, alternar columnas conduce a una ganancia esperada más alta en comparación con seleccionar aleatoriamente o repetir la misma columna.

En este juego específico, donde ambas columnas tienen el mismo valor esperado, la estrategia de alternar columnas garantiza que el jugador no sufra una pérdida y, potencialmente, obtenga ganancias al explotar la estrategia del oponente si se desvía de su propio patrón.

Y analizándolo contra otro NashGuy da la siguiente respuesta:

Estrategia aparente del jugador:

El jugador parece estar empleando una estrategia de "seguir al líder" en la que elige la misma opción que su oponente en el round anterior.

Razones para esta estrategia:

- Minimizar el riesgo: Al elegir la misma opción que el oponente, el jugador evita el riesgo de elegir una opción que podría resultar en una menor recompensa.
- Seguir la tendencia: El jugador puede asumir que el oponente tiene alguna información o conocimiento que influye en su elección, y al seguir al líder, puede beneficiarse de esa información.
- Evitar conflictos: Al elegir la misma opción que el oponente, el jugador evita el potencial de un conflicto o represalias.
- Facilidad de decisión: Esta estrategia no requiere que el jugador analice la matriz de pagos o considere las acciones del oponente, lo que la hace fácil de implementar.

Estrategia del oponente:

El oponente del jugador parece estar alternando entre las columnas 0 y 1, independientemente de las acciones del jugador. Esta estrategia no sigue un patrón lógico o discernible.

En general no se obtuvieron buenos resultados a la hora de usar el LLM para analizar la estrategia del NashGuy.

7.3. Análisis de las estrategias metaheurísticas por el modelo

Veamos cómo describe Gemini la estrategia del GeneticGuy contra distintos oponentes:

Oponente: Random

Estrategia del jugador:

El jugador parece estar empleando una estrategia de "alternancia" o "predicción de respuesta", donde alterna entre las filas superior e inferior de la matriz en función de la elección anterior de su oponente.

Por qué el jugador toma esta decisión:

Esta estrategia tiene como objetivo explotar el patrón de elecciones de su oponente. Al observar que su oponente a menudo elige la columna 1 cuando el jugador elige la fila 1, y la columna 0 cuando el jugador elige la fila 2, el jugador predice que su oponente continuará este patrón. Por lo tanto, alterna entre las filas para maximizar sus ganancias y minimizar las pérdidas.

Evidencia de la estrategia de alternancia:

- El jugador siempre elige la fila 1 después de que su oponente haya elegido la columna 0.
- El jugador siempre elige la fila 2 después de que su oponente haya elegido la columna 1.

Beneficios de la estrategia de alternancia:

- El jugador obtiene una ganancia de 4 cuando su oponente elige la columna 1 después de que el jugador haya elegido la fila 1.
- El jugador evita una pérdida de 5 cuando su oponente elige la columna 0 después de que el jugador haya elegido la fila 2.

Limitaciones de la estrategia de alternancia:

- Si el oponente se da cuenta del patrón y cambia su estrategia, el jugador puede comenzar a perder.
- Esta estrategia no es óptima contra oponentes que eligen columnas al azar o siguen patrones impredecibles.

Oponente: BadGuy

Estrategia del Jugador:

El jugador eligió sistemáticamente la primera fila de la matriz, independientemente de la estrategia de su oponente.

Razonamiento:

El jugador parece estar siguiendo una estrategia de "siempre elegir la primera fila". Esta estrategia es una estrategia simple y predecible que no tiene en cuenta las acciones de su oponente.

La razón potencial de esta estrategia podría ser:

- Falta de comprensión: El jugador puede no entender el juego y puede creer que la primera fila es la mejor opción independientemente de las acciones del oponente.
- Estrategia de prueba: El jugador puede estar probando diferentes estrategias para ver qué funciona mejor.
- Explotación de la falta de estrategia del oponente: El jugador puede haber observado que su oponente no tiene una estrategia clara y decidió explotar eso eligiendo siempre la primera fila.
- Simplificación: El jugador puede estar tratando de simplificar el juego eligiendo siempre la misma opción, lo que reduce la cantidad de opciones que necesita considerar.

Oponente: GeneticGuy

Estrategia del jugador:

El jugador ha optado sistemáticamente por la primera fila, sin variación a lo largo de 510 decisiones.

Razonamiento de la decisión:

Existen varias posibles razones para la estrategia constante del jugador de elegir la primera fila:

- Estrategia dominante: La primera fila puede ser una estrategia dominante, lo que significa que siempre es la mejor opción para el jugador, independientemente de la decisión de su oponente. Esto podría deberse a que la primera fila siempre resulta en un mayor pago o una menor pérdida que las otras filas.
- Creencia de un oponente predecible: El jugador puede creer que su oponente elegirá consistentemente la primera columna, lo que haría que la primera fila fuera la mejor opción para el jugador. Al elegir siempre la primera fila, el jugador puede maximizar sus ganancias o minimizar sus pérdidas.
- Estrategia a prueba de pérdidas: La primera fila puede ser una opción a prueba de pérdidas, lo que significa que garantiza un resultado positivo o al menos no negativo, independientemente de la decisión del oponente. Esto podría ser beneficioso para el jugador si quiere evitar pérdidas o si no está seguro de la estrategia del oponente.
- Estrategia de aleatorización: Aunque el jugador no varía su elección, la decisión constante de elegir la primera fila podría ser parte de una estrategia más amplia de aleatorización. Al no revelar un patrón, el jugador puede hacer más difícil para su oponente predecir su comportamiento.

8. El cuentacuentos

En la sección 2 presentamos un conjunto de matrices, cada una con una historia como trasfondo. En esta sección, en primer lugar utilizamos un LLM para generar un conjunto de matrices cada una con un cuento relacionado. Luego ponemos a los jugadores a jugar entre sí y luego pasamos los resultados del juego a un LLM para que cree una historia conjunta dado el comportamiento de cada uno de los jugadores.

8.1. Asignándole nombres a los jugadores

Utilizando Copilot, generamos una lista en Python con 110 nombres para asignar a nuestros jugadores, así como una función que recibe una lista de jugadores y le asigna a cada uno un elemento distinto de la lista. (En caso de recibir una lista de más de 110 jugadores, algunos jugadores quedarán sin nombre).

8.2. Creando matrices y cuentos usando GPT 4

Para crear una base de datos con suficientes matrices con cuentos de trasfondo hicimos la siguiente consulta a GPT 4:

”Necesito un conjunto de al menos 30 matrices diferentes de juegos de suma no cero cada una con un cuento diferente asociado que tenga sentido con respecto a la matriz, por ejemplo para la matriz

3:3 2:0
0:2 1:1

Un posible cuento sería:

En cierta aldea hay dos cazadores. Cada uno posee un equipaje para cazar ciervos y otro para cazar liebres. Un ciervo posee 6 veces la carne de una liebre, pero un solo cazador no es capaz de cazarlo. En el bosque hay dos liebres y un ciervo, si ambos cazadores llevan el equipaje para cazar liebres, cazan una cada uno, si solo uno lo lleva, caza ambas liebres. EL ciervo solo será cazado si ambos llevan el equipaje para ciervos.

Y para la matriz

-5:-5 -7:-1
-1:-7 -3:-3

Un posible cuento sería:

Dos delincuentes cometieron un asesinato, la policía los atrapa pero solo poseen evidencias para acusarlos por tenencia de armas de fuego, por lo que se les condena a 3 años de cárcel, cuando por homicidio se les condenaría a 7 años de cárcel. Ambos criminales son arrestados en celdas separadas sin la posibilidad de comunicarse entre sí. El jefe de la policía le hace una propuesta a cada uno de los criminales, si delata a su compañero, su condena se reduce en 2 años

Necesito que los cuentos y matrices generados sean lo más distintos posible entre sí y que las matrices sean cuadradas, algunas de 2x2, otras de 3x3 y otras de 4x4. Las matrices no deben ser simétricas necesariamente. Evidentemente cada matriz se debe ajustar a las ganancias de cada jugador según las decisiones tomadas.

Cada fila de la matriz debe estar asociada a una posible decisión del jugador 1 y cada columna de la matriz a una posible decisión del jugador 2. Las decisiones son simultáneas.”

De esta forma, generamos los cuentos almacenados en la base de datos del cuentacuentos. Las respuestas están ubicadas en el directorio `src/llm`. Se puede notar que las respuestas eran erróneas en el sentido de que cuando se mencionaba una matriz de $n \times n$ en la historia, GPT-4 indicaba que había n jugadores, esto fue corregido mediante un procesamiento manual, reemplazando la referencia a n jugadores por dos jugadores, y agrupando todo en un solo archivo.

8.3. Llevando las matrices y cuentos a código usando GPT 3.5

Teniendo las matrices y los cuentos en markdown, decidimos llevarlas a JSON, para esto realizamos el siguiente prompt:

”Necesito llevar las secciones `### Matrix` a una lista en formato json, cada elemento tiene la forma
`{"title": ..., "matrix": ..., "actions": ..., "story": ...}`,
actions es una lista de strings con las acciones disponibles,
matrix es una lista de lista de lista de 2 elementos de tipo int que representa la matriz,
story es el cuento.
Además necesito un título que describa el cuento y guardarlo en title. Necesito que el titulo sea único para cada matriz.”

Fue necesario dividir las matrices/cuentos en grupos de 8, ya que si le pasábamos una mayor cantidad, el modelo comenzaba a delirar.

Al prompt para obtener las últimas matrices se le añadió:

”Debes cambiar D_1, \dots, D_n , por sus respectivos textos en actions.
Debes cambiar Estrategia A, Estrategia B, ... por sus respectivos textos en actions.”

De esta forma creamos nuestra base de datos de matrices/cuentos, la cual actualmente se encuentra en el archivo `src/data.json`

8.4. Usando API de Gemini para generar historias

Con los nombres de nuestros jugadores asignados y los cuentos asociados a nuestras matrices, estamos listos para poner en marcha nuestro cuentacuentos.

8.4.1. Describiendo rondas

Comenzamos describiendo una ronda entre dos jugadores. Se envía a Gemini un prompt de la forma:

”A partir del siguiente cuento (cuento asociado) redacta una (tipo de historia deseada) en el que describas fielmente las decisiones tomadas por cada personaje que se listan a continuación
(nombre jugador 1): (acción jugador 1)
(nombre jugador 2): (acción jugador 2).
Ten en cuenta que (nombre jugador 1) y (nombre jugador 2) son personas.
Además ambos toman la decisión sin conocer la decisión del otro.”

Utilizando la matriz/cuento del dilema del prisionero, donde uno traiciona y el otro coopera, con historia deseada del tipo ”historia épica” se obtuvieron los siguientes resultados:

En las lúgubres celdas de una prisión, Amelia y Gonzalo esperaban ansiosamente su destino. Encarcelados por un delito que ambos habían cometido, se enfrentaban a un dilema que decidiría su futuro: cooperar o traicionar.

Amelia, egoísta y astuta, había decidido traicionar a Gonzalo. Creía que sacrificándolo, ella podría quedar en libertad mientras él asumía la carga de su culpabilidad. Sin saber la decisión de Gonzalo, ella se aferró a su plan.

Gonzalo, por otro lado, era un hombre honrado que valoraba la lealtad por encima de su propia libertad. Había decidido cooperar con Amelia, con la esperanza de que ambos recibieran una sentencia reducida. Desconocía el oscuro secreto que ocultaba su compañera.

Cuando llegó el momento de tomar una decisión, Amelia se dirigió con confianza a las autoridades y confesó su crimen. Gonzalo, horrorizado por su traición, permaneció en silencio, creyendo que su honor era más importante que su libertad.

El destino cumplió su cruel destino. Amelia fue liberada, mientras que Gonzalo recibió una sentencia más dura por su cooperación. Ambos habían tomado decisiones que definieron sus vidas, Amelia eligiendo el camino de la traición y Gonzalo el camino del sacrificio.

Utilizando la matriz/cuento del dilema de los cazadores, donde ambos deciden cazar al ciervo y eligiendo como tipo de historia "tragicomedia" se obtuvo el siguiente resultado:

Escena 1

INT. CABAÑA DE CAZADORES - DÍA

RAMÓN y AMELIA, dos cazadores experimentados, están reunidos en la cabaña. Se están preparando para ir de caza.

RAMÓN Amelia, ¿qué equipaje planeas llevar?

AMELIA (Sonriendo) Equipo para ciervos, Ramón. Creo que tendremos más suerte cazándolos hoy.

RAMÓN ¡Excelente! Yo también llevaré equipaje para ciervos. Siempre es más beneficioso cuando ambos lo llevamos.

Escena 2

EXT. BOSQUE - TARDE

Ramón y Amelia caminan por el bosque, cada uno con su equipaje para ciervos. Después de varias horas de caza, se encuentran.

RAMÓN ¡Amelia! ¡Mira lo que tengo!

Ramón sostiene un gran ciervo. Amelia sonríe y muestra su propia presa.

AMELIA ¡Yo también lo hice!

RAMÓN (Riendo) ¡Qué buena caza! Ambos elegimos sabiamente.

AMELIA Sí, Ramón. Es un resultado perfecto.

Escena 3

INT. CABAÑA DE CAZADORES - NOCHE

Ramón y Amelia están reunidos de nuevo en la cabaña, cenando el ciervo que cazaron.

AMELIA ¿Recuerdas cuando decidimos llevar equipaje para ciervos?

RAMÓN Por supuesto. Fue una gran decisión.

AMELIA Sí, pero imagina si hubiéramos elegido diferente. Si uno de nosotros hubiera llevado equipaje para liebres...

Ambos se ríen.

RAMÓN Habríamos vuelto con una liebre cada uno, o incluso nada si hubiéramos llevado ambos equipaje para liebres.

AMELIA Sí, afortunadamente, ambos tomamos la misma decisión sabia.

Ramón y Amelia continúan cenando, saboreando el resultado de su cuidadosa planificación y la bondad de su amistad.

Utilizando la matriz cuento del dilema del cazador donde ambos eligieron llevar el equipaje para cazar liebres y escogen la modalidad historia corta, se obtuvo el siguiente resultado

En la apacible aldea, Pepito y Juanito, dos cazadores experimentados, se preparaban para una expedición de caza. La decisión de qué equipo llevar era crucial, ya que determinaría su éxito en la adquisición de presas.

Sin saber lo que había elegido el otro, Pepito y Juanito optaron ambos por "equipaje para liebres". Su razonamiento era simple: cazar liebres era más fácil y, si tenían suerte, cada uno podía regresar con una.

A medida que los cazadores se adentraban en el bosque, desconocían que sus decisiones habían coincidido. Con el ligero equipo para liebres auestas, se movieron silenciosamente a través de los árboles, buscando rastros de sus escurridizas presas.

Después de horas de búsqueda, Pepito y Juanito se encontraron en el claro. Sonrieron al verse mutuamente con liebres en la mano. Ambos habían tenido éxito en su misión de cazar una liebre, demostrando que incluso sin coordinación previa, sus decisiones individuales habían conducido al resultado deseado.

Y así, los dos cazadores regresaron a la aldea, satisfechos con sus capturas y orgullosos de sus propias decisiones.

8.4.2. Describiendo juegos

En esta sección utilizamos el LLM para describir un juego, que sería equivalente a una historia conjunta de sus rondas. Intentando distintos enfoques, primero generar historias individuales de cada ronda y al final del juego enviárselas todas a Gemini. Con este enfoque, además de ser más ineficiente pues hacía demasiadas peticiones, no obtuvimos resultados satisfactorios.

Luego intentamos realizar una única consulta con los cuentos y las decisiones tomadas en cada ronda, obteniendo mejores resultados

Las consultas actualmente poseen la siguiente forma:

"Redacta un (tipo de historia) en el que describas cada round.

Empieza describiendo quiénes son los personajes (nombre de personaje 1) y (nombre de personaje 2).

Ten en cuenta que (nombre de personaje 1) y (nombre de personaje 2) son personas.

Luego narra qué ocurrió entre nuestros personajes en cada round.

(A continuación le concatenamos los cuentos relacionados a las matrices, y las decisiones tomadas por cada jugador en cada ronda, además en todas las rondas menos la 1ra se le pide un enlace con la ronda anterior)"

Introduciendo como tipo de historia "historia épica" y la secuencia de cuentos "El dilema de los cazadores", "El dilema del prisionero" y "La guerra de los sexos" donde en el primer caso toman decisiones distintas, en el 2do deciden cooperar y el 3ro deciden ir al ballet, Gemini dio la siguiente respuesta:

Néstor y Xavier: Una Historia Épica

En un pueblo remoto, vivían dos hábiles cazadores: Néstor, conocido por su ingenio, y Xavier, por su valentía. Juntos, se embarcaron en una serie de encuentros que probarían sus lazos y forjarían su destino.

Round 1: La Caza

Néstor y Xavier se encontraron en una encrucijada: podían cazar ciervos o liebres. Néstor, astuto como un zorro, optó por el equipaje para liebres, creyendo que podría superar en astucia a las ágiles criaturas. Xavier, sin embargo, apostó por el equipaje para ciervos, confiando en su fuerza bruta.

El resultado fue un revés para Xavier. Néstor regresó con dos liebres, mientras que Xavier se quedó con las manos vacías. Sin embargo, su desacuerdo sembró una semilla de desconfianza entre ellos.

Round 2: El Interrogatorio

Tiempo después, Néstor y Xavier fueron acusados injustamente de un delito. Los encerraron en celdas separadas y les ofrecieron un trato siniestro: cooperar o traicionar. Néstor, recordando su encuentro anterior, optó por la cooperación. Xavier, sin embargo, decidió traicionar, creyendo que podría escapar del castigo.

Pero su plan se hizo añicos cuando ambos cooperaron. El juez, impresionado por su honestidad, les impuso sentencias más leves. El recuerdo de su fracaso anterior los llevó a la redención.

Round 3: La Elección

Liberados de la prisión, Néstor y Xavier se enfrentaron a una nueva encrucijada: elegir entre asistir al ballet o al boxeo. Néstor, un alma sensible, prefería la gracia y la elegancia del ballet. Xavier, un guerrero de corazón, anhelaba la emoción de la pelea de boxeo.

Recordando sus pasadas diferencias, decidieron ir juntos al ballet. Aunque Xavier inicialmente se sintió incómodo, el ambiente refinado y la belleza de la actuación lo cautivaron. Néstor, a su vez, quedó impresionado por el poder y la pasión que Xavier exhibía en sus comentarios.

En ese momento, se dieron cuenta de que sus diferencias no los dividían, sino que les complementaban. Juntos, habían superado desafíos, aprendido de sus errores y encontrado el valor de la cooperación y la amistad. Y así, Néstor y Xavier, los cazadores un día, se convirtieron en compañeros inseparables, unidos por los lazos de la adversidad y el entendimiento mutuo.

Le pedimos una historia breve narrando un Cazadores-Prisionero-Ballet, donde en la primera ronda elegían diferentes equipos de caza, en la segunda ambos cooperaban y en la tercera ambos elegían ballet:

Personajes:

Omar: Un cazador astuto y calculador.

Diego: Un cazador experimentado y pragmático.

Round 1: La caza

En la penumbra del amanecer, Omar y Diego se prepararon para cazar. Omar optó por el equipaje para liebres, convencido de que podría superar en astucia a las escurridizas criaturas. Diego, por su parte, eligió el equipaje para ciervos, confiando en su habilidad para rastrear a los majestuosos animales.

En el vasto bosque, sus caminos se cruzaron. Omar se adentró en los arbustos, sus sentidos alerta ante el menor atisbo de movimiento. Diego se acercó sigilosamente a un claro, su mirada fija en el horizonte.

El sol comenzó a ponerse cuando Omar regresó al claro, con dos liebres colgando de su cinturón. Diego, sin embargo, se encontró con las manos vacías. Su búsqueda de ciervos había sido infructuosa.

Round 2: La traición

Después de compartir su escasa cena, Omar y Diego se retiraron a sus tiendas. Mientras la noche envolvía el bosque, Diego se revolvía en su incómoda cama. Sabía que si cooperaba con Omar y guardaba silencio sobre sus intenciones de caza, ambos serían sentenciados a un castigo leve. Sin embargo, la tentación de traicionar a Omar y obtener la libertad era abrumadora.

Omar, atormentado por sus propios pensamientos, también luchó con la elección. Cooperar significaba arriesgar su libertad, pero traicionar a su compañero cazador violaba su código moral.

Al amanecer, ambos hombres se enfrentaron en la plaza del pueblo. Omar preguntó: "¿Cooperamos o traicionamos?". Diego respondió: "Cooperemos". Y así, mantuvieron el silencio, esperando un resultado mejor que la traición.

Round 3: La elección

A medida que el sol alcanzaba su cenit, Omar y Diego recibieron una invitación para asistir a dos eventos muy diferentes: un ballet y un partido de boxeo. Omar anhelaba la gracia y la delicadeza del ballet, mientras que Diego ansiaba la emoción y la intensidad del boxeo.

Sin saber las preferencias del otro, ambos hombres tomaron una decisión. Omar se dirigió al teatro, su corazón palpitando con anticipación. Diego se dirigió al cuadrilátero, sus puños apretados con determinación.

Al final del día, se encontraron en el bar del pueblo. Omar se maravilló con la belleza del ballet, mientras que Diego se deleitó con la visceralidad del boxeo. Ambos lamentaron que hubieran tomado caminos separados, sabiendo que la felicidad mutua había estado a su alcance.

Y así, los tres rounds revelaron el intrincado juego de elecciones, cooperación y conflicto que entrelaza el destino de los individuos.

Podemos observar que en la última ronda deliró narrando cómo uno fue al ballet y el otro al boxeo, cuando en realidad ambos habían ido al ballet.

9. Conclusiones

En este trabajo, aplicamos nuestros conocimientos en inteligencia artificial (IA) y simulación para analizar diversas estrategias en juegos de suma no cero. Modelamos a los jugadores como agentes inteligentes de diferentes tipos y realizamos simulaciones para evaluar su desempeño. Aunque no logramos identificar una estrategia óptima de manera general, encontramos que ciertas estrategias son altamente efectivas en escenarios específicos. Por ejemplo, el enfoque de NashGuy demuestra una gran eficacia en juegos repetidos.

Nuestras metaheurísticas exhiben un comportamiento variable, produciendo resultados sorprendentemente satisfactorios en algunos casos y pésimos en otros. El rendimiento de estas estrategias varía según el juego en cuestión; descubrimos que entrenarlas con estrategias específicas o para juegos particulares puede mejorar su desempeño, aunque esto pueda limitar su aplicabilidad a situaciones más generales.

Los jugadores difusos que desarrollamos dependen en gran medida de la métrica utilizada y del contexto del juego. Si bien muestran un rendimiento ligeramente inferior al enfoque de "ojo por ojo" en juegos simétricos y repetidos, mantienen su efectividad en juegos asimétricos y no repetidos, donde el enfoque de "ojo por ojo" no es efectivo.

Al utilizar modelos de lenguaje para describir jugadores basado en sus decisiones en un torneo, obtuvimos resultados prometedores al analizar estrategias simples, aunque enfrentamos dificultades al tratar con estrategias más complejas.

Sin embargo, logramos excelentes resultados en la generación de contenido utilizando estos modelos. Por ejemplo, Bing fue capaz de generar una amplia variedad de nombres para jugadores, así como narrativas asociadas a las matrices de juego. Aunque algunas de estas narrativas contenían pequeñas imprecisiones, fueron fácilmente corregibles.

Además, utilizamos Gemini para crear narrativas relacionadas con el juego, obteniendo resultados divertidos y satisfactorios. Es importante tener en cuenta que, a medida que aumenta la complejidad del juego, la coherencia de las narrativas puede disminuir, pero en juegos más simples, Gemini demostró ser capaz de producir historias satisfactorias en la mayoría de los casos.

10. Propuestas para futuros trabajos

1. Definir nuevas funciones de membresía para el ojo por ojo difuso.
2. Probar distintos tipos de entrenamientos para la metaheurística
3. Probar otras metaheurísticas
4. Ampliar la base de datos de matrices/cuentos, añadiendo principalmente una mayor cantidad de matrices no simétricas
5. Analizar distintos tipos de torneos

11. Bibliografía

Referencias

- [1] Noam Nisan, Tim Roughgarden, Eva Tardos, & Vijay V. Vazirani (Eds.). *Algorithmic Game Theory*. Cambridge University Press, 2007.
- [2] William Spaniel. *Game Theory 101: The Complete Textbook*. Game Theory 101, 2011.
Disponible en: <https://gametheory101.com>
- [3] Luciano García Garrido, Luis Martí Orosa, & Luis Pérez Sánchez. *Temas de Simulación*. Facultad de Matemática y Computación, Universidad de la Habana, 200X.