

# Proyecto de IA - Simulación Major League Baseball

Alejandro Álvarez Lamazares  
Marian S. Álvarez Suri  
Carlos A. Bresó Sotto

Facultad de Matemática y Computación. Universidad de La Habana

**Resumen** Este reporte describe el desarrollo y resultados del proyecto de inteligencia artificial y simulación aplicado a la Major League Baseball (MLB). El objetivo del proyecto es simular juegos de béisbol utilizando técnicas de inteligencia artificial para tomar decisiones estratégicas. En particular, se realizó una simulación de los partidos de la temporada 2022 de la MLB, utilizando elementos de IA como algoritmos de búsqueda, conocimiento y procesamiento de lenguaje natural. En la simulación, los agentes fueron tratados desde el enfoque de Beliefs - Desires - Intentions (BDI), lo que permitió modelar de manera más realista las decisiones estratégicas y comportamientos de los managers.

## 1. Introducción

El objetivo de este proyecto es simular los juegos de la temporada 2022 de la Major League Baseball (MLB) [1] utilizando técnicas avanzadas de inteligencia artificial y simulación. La simulación permite observar cómo diferentes estrategias pueden afectar el resultado de un juego, proporcionando una herramienta valiosa para entrenadores, analistas y aficionados al béisbol.

Nuestra propuesta de solución incluye un algoritmo genético para optimizar la selección del lineup, asegurando que los jugadores más adecuados estén en el campo de juego. Además, el manager virtual toma decisiones estratégicas basadas en un sistema experto, que evalúa las estadísticas y situaciones del juego en tiempo real. Finalmente, hemos implementado un comentarista virtual utilizando procesamiento de lenguaje natural con la IA Gemini [2].

## 2. Datos y Preprocesamiento

Los datos utilizados en el proyecto incluyen estadísticas de jugadores y equipos de la temporada 2022 de la MLB, así como información sobre los partidos y resultados de la temporada. Los datos se obtuvieron de la página Baseball-Savant [3] y mediante el uso de la librería statsapi [4]. El preprocesamiento de los datos incluyó su limpieza, transformación y normalización para que fueran compatibles con los algoritmos de inteligencia artificial utilizados en el proyecto.

### 3. Simulación del Juego

La fase de simulación del juego es una parte crucial del proyecto, donde se modelan las dinámicas y reglas del béisbol para replicar los partidos de la temporada 2022 de la MLB. En esta fase, se utiliza la estructura de Beliefs-Desires-Intentions (BDI) para crear los agentes, que en este caso son los managers de los equipos.

#### 3.1. Estructura BDI para los Agentes

La estructura BDI es un modelo de agentes que se basa en tres componentes principales [5]:

- **Creencias (Beliefs)**: Representan la información que el agente tiene sobre el funcionamiento del juego y el estado actual del mismo. Constituyen el conjunto de condiciones que tienen que cumplirse para aplicar cierta regla.
- **Deseos (Desires)**: Son los objetivos que el agente quiere alcanzar, como anotar más carreras, defender la ventaja o tomar una alternativa conservadora. A partir del estado del juego, se determina cuáles de los objetivos quieren cumplirse con mayor interés y, por tanto, qué conjunto de reglas se deben considerar.
- **Intenciones (Intentions)**: Son las acciones que el agente decide llevar a cabo para alcanzar sus deseos, basándose en sus creencias.

#### 3.2. Tipos de Managers

En la simulación, pueden existir diferentes tipos de managers, cada uno con su propio estilo y estrategia:

- **Manager Agresivo**: Toma decisiones arriesgadas para intentar obtener una ventaja significativa. Este tipo de manager puede decidir cambiar a un bateador por uno emergente o intentar robos de base en situaciones críticas.
- **Manager Defensivo**: Se concentra en mantener la ventaja del juego indicando acciones como vigilar al corredor en base para evitar robos.
- **Manager Conservador**: Prefiere tomar decisiones seguras y minimizar riesgos. Por ejemplo, puede optar por otorgar una base por bolas intensional para mantener la ventaja.
- **Manager Neutral**: Considera todos los posibles tipos de reglas que pueden efectuarse, siempre que estas tengan sentido.

#### 3.3. Decisiones del Manager

Las decisiones que puede tomar cada manager incluyen, pero no se limitan a:

- **Cambio de Pitcher**: Decidir cuándo cambiar al pitcher actual por uno del bullpen, basado en el rendimiento y la fatiga del pitcher.

- **Robo de Base:** Decidir cuándo intentar un robo de base, considerando la velocidad del corredor.
- **textbfBateador Emergente:** Cambiar al bateador de turno por uno emergente que no se encontraba en la alineación pero que podría representar una clara ventaja ofensiva en el actual turno al bate.
- **Jugada de Corrido y Bateo:** Indicar un intento de jugada de corrido y bateo basado en el contacto del bateador y en la velocidad del corredor.
- **Pickoff:** Indicarle al pitcher que realice lanzamientos a la base donde se encuentre el corredor si considera que existe amenaza de robo de base.

El conjunto de decisiones que pueden tomar nuestros agentes se basa en qué porcentaje de cada tipo de manager es. Dado un estado determinado del juego, se determina cuán agresivo, defensivo, conservador y neutral debería ser un manager para, a partir de los porcentajes determinados, incluir un conjunto de reglas a ser consideradas. De esta forma, podemos tener managers totalmente agresivos, parcialmente conservadores y defensivos, o cualquier otra combinación posible.

## 4. Inteligencia Artificial

Los elementos de inteligencia artificial integrados permiten un mejor acercamiento a la realidad. Empleamos tres enfoques fundamentales durante las simulaciones con algoritmos de búsqueda, conocimiento y procesamiento de lenguaje natural.

### 4.1. Algoritmo de Búsqueda

El béisbol es un deporte que involucra una gran cantidad de variables y decisiones estratégicas. La selección del mejor lineup para un equipo es un problema complejo que requiere considerar múltiples factores, como las estadísticas de rendimiento de los jugadores, las restricciones de posicionamiento y la importancia del orden. Un **algoritmo genético** es particularmente adecuado para este problema debido a la posibilidad de explorar un amplio espacio de soluciones y encontrar lineups que maximicen el rendimiento del equipo. Basándonos en el principio de conservar los jugadores que maximicen la función de evaluación, se logra mejorar en cada iteración el resultado final y acomodar las posiciones definitivas que ocupará cada jugador. El proceso de selección del lineup implica los siguientes pasos:

1. **Inicialización:** Se genera una población inicial de posibles lineups de manera aleatoria.
2. **Evaluación:** Cada lineup se evalúa utilizando una función de aptitud que considera factores como las estadísticas de bateo de los jugadores y la efectividad defensiva.
3. **Selección:** Se seleccionan los mejores lineups basados en su aptitud para formar una nueva generación.

4. **Cruzamiento:** Se combinan pares de lineups seleccionados para crear nuevos lineups, intercambiando características entre ellos.
5. **Mutación:** Se introducen pequeñas modificaciones aleatorias en algunos lineups para mantener la diversidad genética y evitar la convergencia prematura.
6. **Iteración:** Los pasos de evaluación, selección, cruzamiento y mutación se repiten durante varias generaciones hasta que se encuentra un lineup óptimo o se alcanza un criterio de parada.

#### 4.2. Algoritmo de Conocimiento

Los managers constituyen una parte fundamental en el desarrollo de un juego de béisbol. La toma de decisiones estratégicas, como cuándo cambiar al pitcher o cuándo intentar un robo de base, puede influir significativamente en el resultado del juego. Para modelar el comportamiento de los managers, utilizamos un **sistema experto** que evalúa las situaciones del juego y toma decisiones basadas en reglas predefinidas. El sistema experto se basa en un conjunto de reglas que representan el conocimiento y la experiencia de un manager real, y se utiliza para determinar las acciones más adecuadas en cada situación. Las reglas se definen en función de los deseos del manager, que son determinados a partir del estado del juego.

#### 4.3. Algoritmo de Procesamiento de Lenguaje Natural

El comentarista virtual es una parte importante de la simulación, ya que proporciona información y análisis sobre el desarrollo del juego. Para implementar el comentarista virtual, utilizamos el **modelo de lenguaje Gemini** de OpenAI, que es capaz de generar texto coherente y relevante en función de una entrada dada. El comentarista virtual recibe información sobre el estado del juego en cada turno al bate, como el marcador actual, los jugadores en base y el número de outs, y genera comentarios que describen lo que está sucediendo.

### 5. Implementación

El proyecto está implementado en Python y utiliza bibliotecas como *scikit-learn* para los algoritmos de aprendizaje automático y *pandas* para el manejo de datos. La simulación del juego se realiza mediante un modelo basado en eventos discretos que representa las reglas y dinámicas del béisbol.

### 6. Resultados

Los resultados de la simulación muestran cómo diferentes decisiones estratégicas pueden influir en el resultado del juego. Se presentan estadísticas detalladas de varios juegos simulados, incluyendo el rendimiento de los jugadores y las decisiones tomadas por la inteligencia artificial.

## 7. Conclusiones

El proyecto demuestra que es posible utilizar técnicas de inteligencia artificial para simular juegos de béisbol y tomar decisiones estratégicas. Esto puede ser útil para entrenadores y analistas que buscan optimizar sus estrategias de juego.

## Referencias

1. MLB, “Major league baseball,” 2023, accedido: 2023-10-01. [Online]. Available: <https://www.mlb.com>
2. OpenAI, “Gemini: Advanced natural language processing,” 2023, accedido: 2023-10-01. [Online]. Available: <https://www.openai.com/gemini>
3. BaseballSavant, “Statcast,” 2022, accedido: 2022-12-01. [Online]. Available: [https://baseballsavant.mlb.com/statcast\\_home](https://baseballsavant.mlb.com/statcast_home)
4. PyPI, “Mlb-statsapi,” 2024, accedido: 2024-10-01. [Online]. Available: <https://pypi.org/project/MLB-StatsAPI/>
5. A. S. Rao and M. P. Georgeff, *BDI Agents: From Theory to Practice*. Berlin, Heidelberg: Springer, 1995. [Online]. Available: [https://link.springer.com/chapter/10.1007/3-540-59498-1\\_21](https://link.springer.com/chapter/10.1007/3-540-59498-1_21)