

# Entorno interactivo para el estudio de estrategias de I.A. en juegos

José Miguel Horcas Aguilera

23 de abril de 2012

# Índice

1 Introducción

2 Juegos

3 Estrategias y heurísticos

4 La aplicación

5 Conclusiones

# Índice

1 Introducción

2 Juegos

3 Estrategias y heurísticos

4 La aplicación

5 Conclusiones

# Introducción

## Objetivos

Desarrollar un entorno interactivo que permita jugar, estudiar y comparar el rendimiento de estrategias de IA en juegos.

## Motivación

- Facilidad de representar el estado de los juegos.
- Definición precisa de las reglas.
- Requieren la capacidad de tomar alguna decisión.

# Índice

1 Introducción

2 Juegos

3 Estrategias y heurísticos

4 La aplicación

5 Conclusiones

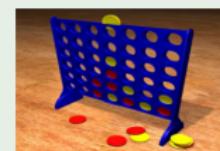
# Juegos

## Características de los juegos

Bipersonales, por turnos, de suma cero, de información perfecta y deterministas.

## Ejemplos: *classic board-games*

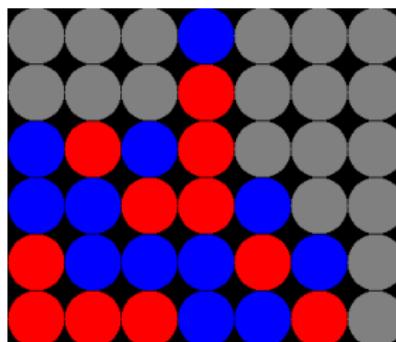
- 3 en raya
- **Conecta-4**
- Damas
- Othello (Reversi)
- Ajedrez
- **Go**



# Conecta-4

## Reglas

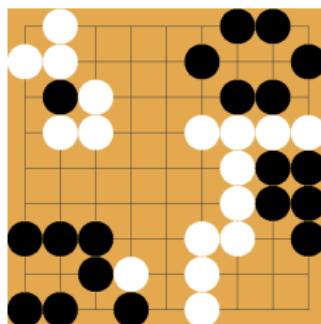
- **Objetivo:** Colocar 4 fichas en línea del mismo color.
- **Tablero:** 6x7.
- Las fichas se dejan caer por las columnas.



# Go

## Reglas

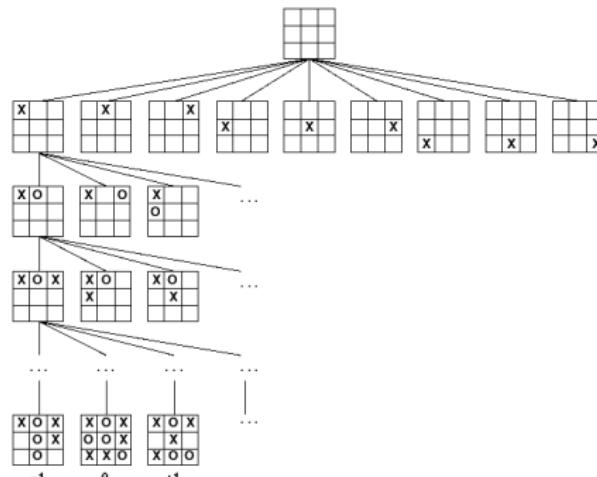
- **Objetivo:** controlar más tablero que el oponente.
- **Tablero:**  $19 \times 19$ ,  $17 \times 17$ ,  $13 \times 13$ ,  $9 \times 9$ .
- Las fichas se colocan sobre las intersecciones.
- Se capturan las fichas encerradas.
- Prohibido repetir situaciones y suicidarse.
- Gana el jugador con más puntos:
  - **Reglas japonesas:**  $1p/\text{inter.} - 1p/\text{ficha capturada}$
  - **Reglas chinas:**  $1p/\text{inter.} + 1p/\text{ficha tablero}$



# Problemas de búsqueda con adversarios

## Representación de los juegos

- **Estado inicial:** situación inicial de la partida y turno del jugador.
- **Función sucesor:** movimiento legal y el estado resultante.
- **Test terminal:** ¿cuándo se termina el juego?
- **Función de utilidad:** +1, -1 ó 0 si el resultado es un triunfo, una pérdida o un empate.



# Índice

1 Introducción

2 Juegos

3 Estrategias y heurísticos

4 La aplicación

5 Conclusiones

# Estrategias

## Estrategias desarrolladas

- Humana
- Aleatoria
- Monte-Carlo
- Monte-Carlo Tree Search
- Evaluador heurístico
- Minimax
- Poda Alfa-Beta
- Tablas de transposición

Jugadores	Prof. búsque- da	Nº simula- ciones	Límite tiempo	Heurísticos	Parámetros
Aleatorio					
Monte-Carlo		X	X		
Monte-Carlo Tree Search		X	X		
Evaluador heurístico				X	
Minimax	X		X	X	
Alfa-Beta	X		X	X	
Minimax (tabla transposición)	X		X	X	
Alfa-Beta (tabla transposición)	X		X	X	

# Estrategias básicas

## Jugador Humano

### Estrategia

Pide el movimiento a realizar por un dispositivo de entrada/salida.

Depende del juego.

## Jugador aleatorio

### Estrategia

Realiza movimientos aleatorios.

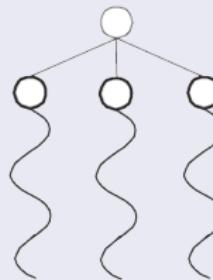
Es totalmente independiente del juego.

# Agentes basados en simulaciones

## Monte-Carlo

### Estrategia

- Realiza simulaciones a partir del estado actual.
- Una **simulación** es una partida completa con movimientos al azar.
- El movimiento con mayor valor de recompensa esperado es el elegido.



### Versión

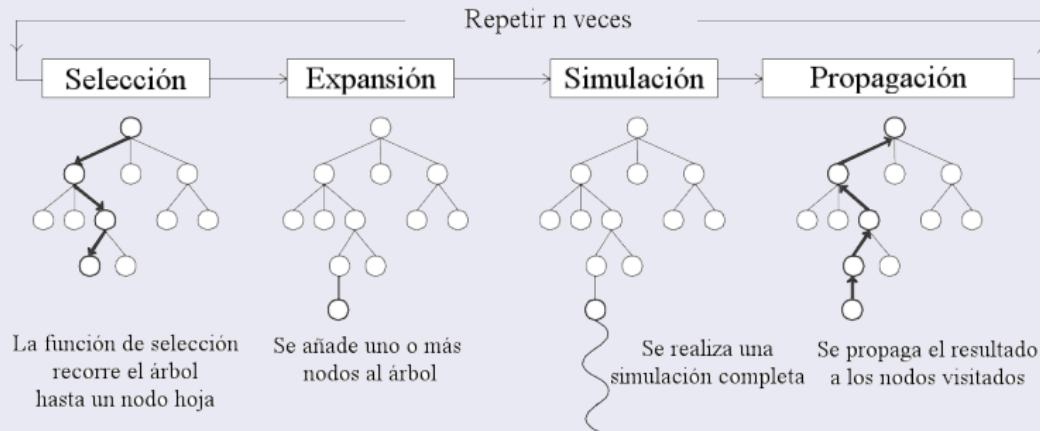
- Número de simulaciones.
- Tiempo limitado.



# Agentes basados en simulaciones

## Monte-Carlo Tree Search

### Estrategia



### Parámetros

- Constante de exploración.
- Posibilidad de reutilizar el árbol.

### Versión

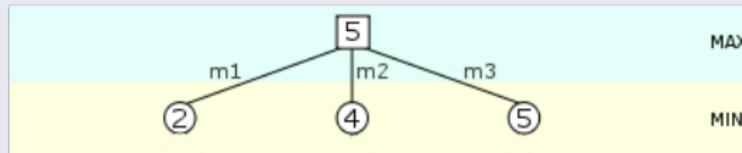
- Número de simulaciones.
- Tiempo limitado.

# Agentes con evaluador heurístico

## Evaluador heurístico

### Estrategia

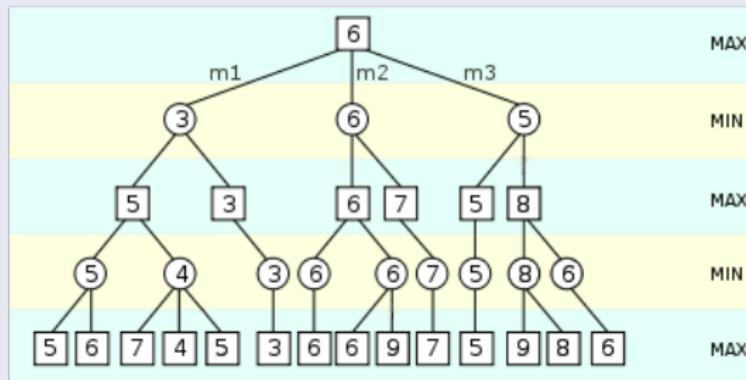
*Dado un estado, considera todos los movimientos inmediatos, los evalúa heurísticamente y escoge el mejor.*



# Agentes con evaluador heurístico

## Minimax

### Estrategia



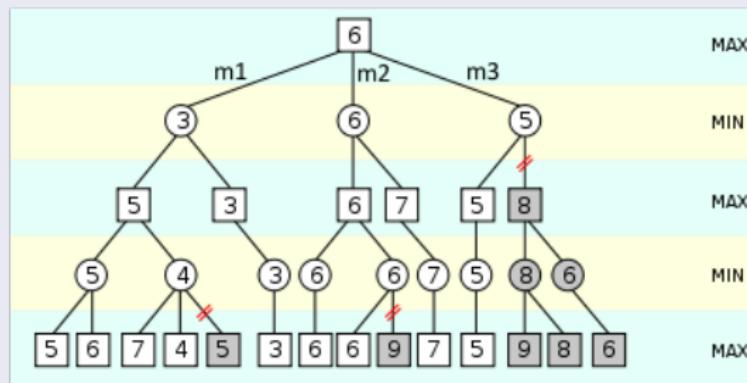
### Versiónes

- Profundidad máxima de búsqueda.
- Tiempo limitado.

# Agentes con evaluador heurístico

## Poda alfa-beta

### Estrategia



### Versiónes

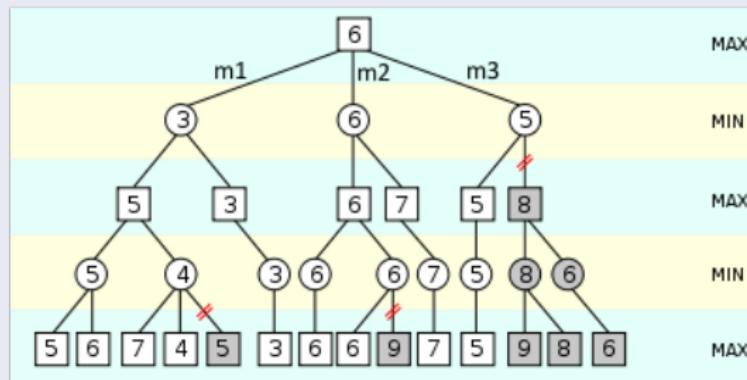
- Profundidad máxima de búsqueda.
- Tiempo limitado.

# Agentes con evaluador heurístico

## Tabla de transposición

### Estrategia

Una **transposición** es un estado que puede ser alcanzado por más de un camino.



### Versiones

- Minimax con tabla de transposición.
- Poda alfa-beta con tabla de transposición.



# Heurísticos

## Función de evaluación heurística

$$e(n) = \begin{cases} \infty & \text{si } n \text{ es un estado terminal y gana MAX} \\ > 0 & \text{si } n \text{ es favorable para MAX} \\ 0 & \text{si } n \text{ es indiferente ("empate")} \\ < 0 & \text{si } n \text{ es desfavorable para MAX} \\ -\infty & \text{si } n \text{ es un estado terminal y pierde MAX} \end{cases}$$

## Evaluadores heurísticos

- Tabla de valor
- Red neuronal
- Heurísticos del Conecta-4:
  - Matriz de posibilidades
- Heurísticos del Go:
  - Evaluador de territorios
  - Evaluador de puntos JP
  - Evaluador de puntos CH

# Heurísticos genéricos

## Tabla de valor

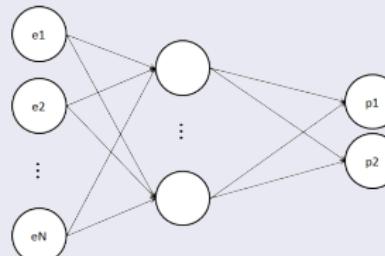
- $e(n) = \text{tabla}(n)$
- Aprendizaje mediante el método de las diferencias temporales:

$$\text{tabla}(n) \leftarrow \text{tabla}(n) + \alpha(\text{tabla}(s) - \text{tabla}(n))$$

- Evaluador independiente del juego.

## Red neuronal

- $e(n) = p_1 - p_2$
- Método de las diferencias temporales.
- **Depende del juego.**



# Heurísticos específicos de los juegos

## Conecta-K

- **Matriz de posibilidades:**

$$e(n) = \text{posibilidades}_{\text{MAX}}(n) - \text{posibilidades}_{\text{MIN}}(n)$$

3	4	5	7	5	4	3
4	6	8	10	8	6	4
5	8	11	13	11	8	5
5	8	11	13	11	8	5
4	6	8	10	8	6	4
3	4	5	7	5	4	3

## Go

- **Territorios:**  $e(n) = \text{territorios}_{\text{MAX}}(n) - \text{territorios}_{\text{MIN}}(n)$
- **Puntos JP:**  $e(n) = \text{puntos}_{\text{MAX}}^{\text{JP}}(n) - \text{puntos}_{\text{MIN}}^{\text{JP}}(n)$
- **Puntos CH:**  $e(n) = \text{puntos}_{\text{MAX}}^{\text{CH}}(n) - \text{puntos}_{\text{MIN}}^{\text{CH}}(n)$

# Índice

1 Introducción

2 Juegos

3 Estrategias y heurísticos

4 La aplicación

5 Conclusiones

# Índice

1 Introducción

2 Juegos

3 Estrategias y heurísticos

4 La aplicación

5 Conclusiones

# Evaluadores heurísticos en el Go

## Jugador 1



## Jugador 2

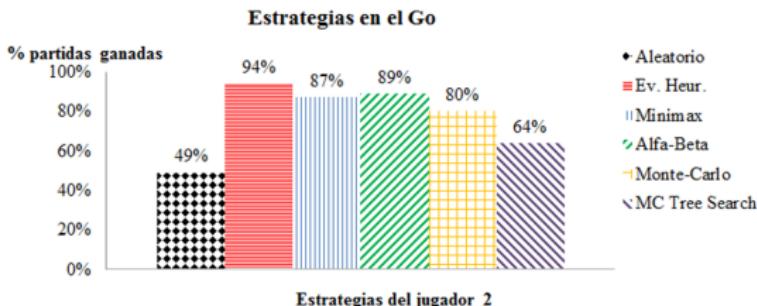


# Estrategias en el Go

## Jugador 1



## Jugador 2



# Conclusiones y trabajo futuro

## Conclusiones

- Herramienta didáctica para estudiar, analizar y comparar algoritmos.
- Marco de trabajo.

## Trabajo futuro

- Adaptación del módulo de razonamiento a otras clases de juegos.
- Nuevos algoritmos.

