Algoritmos para Juegos con Información Incompleta y No Determinismo

Rubmary Rojas

Universidad Simón Bolívar, Caracas, Venezuela

Enero 2020

Agenda

1. Teoría de juegos.

Agenda

1. Teoría de juegos.
2. Juegos en forma normal:
   - Modelo y estrategias.
   - Equilibrio de Nash.
   - Regret Matching: algoritmo, experimentos y conclusiones.

Agenda

1. Teoría de juegos.
2. Juegos en forma normal:
   - Modelo y estrategias.
   - Equilibrio de Nash.
   - Regret Matching: algoritmo, experimentos y conclusiones.
3. Juegos en forma extensiva:
   - Modelo y estrategias.
   - Kuhn Poker.
   - Counterfactual Regret Minimization (CFR): algoritmo, experimentos y conclusiones.

Agenda

1. Teoría de juegos.
2. Juegos en forma normal:
   - Modelo y estrategias.
   - Equilibrio de Nash.
   - Regret Matching: algoritmo, experimentos y conclusiones.
3. Juegos en forma extensiva:
   - Modelo y estrategias.
   - Kuhn Poker.
   - Counterfactual Regret Minimization (CFR): algoritmo, experimentos y conclusiones.
4. Conclusiones y Recomendaciones.
5. Demo.

Teoría de Juegos

Definición

- *Estudio de* **modelos matemáticos** *de* **conflicto** *y* **cooperación**.
- *Agentes que toman decisiones de forma* **racional** *e* **inteligente**.

Aplicaciones

Ciencias sociales

Economía

Matemática

Computación

Juegos no deterministas con información incompleta

No determinismo

Incertidumbre probabilística:
- Lanzar dados
- Repartir cartas

Información incompleta

Información parcial sobre algunas de las acciones que fueron tomadas previamente.

Interrogantes

- ¿Qué significa que un juego sea resuelto?
- ¿Cuándo un jugador juega de forma óptima?

Objetivo General

Comprender los conceptos en el área de juegos de dos personas que involucran información incompleta y no determinismo, así como implementar los algoritmos para resolverlos, realizando experimentos sobre distintos juegos que son capturados por el modelo.

Forma Normal o Estratégica

Juegos en Forma Normal o Estratégica

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	$0, 0$	$-1, 1$	$1, -1$
$\mathcal{P}$ (papel)	$1, -1$	$0, 0$	$-1, 1$
$\mathcal{S}$ (tijera)	$-1, 1$	$1,-1$	$0, 0$

Juegos en Forma Normal o Estratégica

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	$0, 0$	$-1, 1$	$1, -1$
$\mathcal{P}$ (papel)	$1, -1$	$0, 0$	$-1, 1$
$\mathcal{S}$ (tijera)	$-1, 1$	$1, -1$	$0, 0$

jugador 1

Juegos en Forma Normal o Estratégica

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)	jugador 2
$\mathcal{R}$ (piedra)	$0, 0$	$-1, 1$	$1, -1$	
$\mathcal{P}$ (papel)	$1, -1$	$0, 0$	$-1, 1$	
$\mathcal{S}$ (tijera)	$-1, 1$	$1, -1$	$0, 0$	

Juegos en Forma Normal o Estratégica

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	$0, 0$	$-1, 1$	$1, -1$
$\mathcal{P}$ (papel)	$1, -1$	$0, 0$	$-1, 1$
$\mathcal{S}$ (tijera)	$-1, 1$	$1, -1$	$0, 0$

primer jugador **gana** 1

Juegos en Forma Normal o Estratégica

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	$0, 0$	$-1, 1$	$1, -1$
$\mathcal{P}$ (papel)	$1, -1$	$0, 0$	$-1, 1$
$\mathcal{S}$ (tijera)	$-1, 1$	$1, -1$	$0, 0$

segundo jugador **pierde** 1

Juegos en Forma Normal o Estratégica

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	$0, 0$	$-1, 1$	$1, -1$
$\mathcal{P}$ (papel)	$1, -1$	$0, 0$	$-1, 1$
$\mathcal{S}$ (tijera)	$-1, 1$	$1,-1$	$0, 0$

Elementos

1. Jugadores.
2. Acciones o estrategias puras:
   $\mathcal{R}$, $\mathcal{P}$, $\mathcal{S}$.
3. Función de pago o utilidades.

Juegos en Forma Normal o Estratégica

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	$0, 0$	$-1, 1$	$1, -1$
$\mathcal{P}$ (papel)	$1, -1$	$0, 0$	$-1, 1$
$\mathcal{S}$ (tijera)	$-1, 1$	$1, -1$	$0, 0$

Juego de dos jugadores de suma cero

Elementos

1. Jugadores.
2. Acciones o estrategias puras: $\mathcal{R}$, $\mathcal{P}$, $\mathcal{S}$.
3. Función de pago o utilidades.

Juegos en Forma Normal o Estratégica

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	$0,0$	$-1,1$	$1,-1$
$\mathcal{P}$ (papel)	$1,-1$	$0,0$	$-1,1$
$\mathcal{S}$ (tijera)	$-1,1$	$1,-1$	$0,0$

Juego de dos jugadores de suma cero

Elementos

1. Jugadores.
2. Acciones o estrategias puras: $\mathcal{R}$, $\mathcal{P}$, $\mathcal{S}$.
3. Función de pago o utilidades.

Estrategias

1. Estrategias puras: siempre se elige la misma acción.
2. Estrategias mixtas: cada acción se elige con cierta probabilidad.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

José

		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

Juegos en Forma Normal o Estratégica

Batalla de los sexos

José

María		ballet	béisbol
	ballet	$2,1$	$0,0$
	béisbol	$0,0$	$1,2$

Ninguno obtiene ganancia.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

María obtiene una ganancia mayor que José.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

José obtiene una ganancia mayor que María.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

Conceptos

1. Ganancia Esperada

Valor promedio que un determinado jugador obtendría si jugara infinitas veces cuando cada jugador utiliza una estrategia dada.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

Conceptos

1. Ganancia Esperada
2. Mejor Respuesta

La mejor forma en que puede jugar un jugador dadas las estrategias seleccionadas de sus oponentes.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2,1$	$0,0$
	béisbol	$0,0$	$1,2$

Si María siempre elige ballet.

Conceptos

1. Ganancia Esperada
2. Mejor Respuesta

La mejor forma en que puede jugar un jugador dadas las estrategias seleccionadas de sus oponentes.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

Lo mejor para José es siempre elegir ballet.

Conceptos

1. Ganancia Esperada
2. Mejor Respuesta

La mejor forma en que puede jugar un jugador dadas las estrategias seleccionadas de sus oponentes.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

Conceptos

1. Ganancia Esperada
2. Mejor Respuesta
3. Equilibrio de Nash

Cada jugador utiliza una mejor respuesta frente a las estrategias de sus oponentes.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

Conceptos

1. Ganancia Esperada
2. Mejor Respuesta
3. Equilibrio de Nash

Cada jugador utiliza una mejor respuesta frente a las estrategias de sus oponentes.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

María no tiene motivos para cambiar su estrategia.

Conceptos

1. Ganancia Esperada
2. Mejor Respuesta
3. Equilibrio de Nash

Cada jugador utiliza una mejor respuesta frente a las estrategias de sus oponentes.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

José no tiene motivos para cambiar su estrategia.

Conceptos

1. Ganancia Esperada
2. Mejor Respuesta
3. Equilibrio de Nash

Cada jugador utiliza una mejor respuesta frente a las estrategias de sus oponentes.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

Conceptos

1. Ganancia Esperada
2. Mejor Respuesta
3. Equilibrio de Nash

Cada jugador utiliza una mejor respuesta frente a las estrategias de sus oponentes.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

María \ José	ballet	béisbol
ballet	$2, 1$	$0, 0$
béisbol	$0, 0$	$1, 2$

Conceptos

1. Ganancia Esperada
2. Mejor Respuesta
3. Equilibrio de Nash
4. Equilibrio Correlacionado

Puede haber cooperación entre los jugadores.

Juegos en Forma Normal o Estratégica

Batalla de los sexos

		José	
		ballet	béisbol
María	ballet	$2, 1$	$0, 0$
	béisbol	$0, 0$	$1, 2$

Lanzar una moneda

1. cara $\Longrightarrow$ ballet
2. sello $\Longrightarrow$ béisbol

Conceptos

1. Ganancia Esperada
2. Mejor Respuesta
3. Equilibrio de Nash
4. Equilibrio Correlacionado

Puede haber cooperación entre los jugadores.

Equilibrio de Nash

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	0, 0	-1, 1	1,-1
$\mathcal{P}$ (papel)	1,-1	0, 0	-1, 1
$\mathcal{S}$ (tijera)	-1, 1	1,-1	0, 0

Equilibrio de Nash

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	0, 0	-1, 1	1,-1
$\mathcal{P}$ (papel)	1,-1	0, 0	-1, 1
$\mathcal{S}$ (tijera)	-1, 1	1,-1	0, 0

Equilibrio de Nash

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	0, 0	-1, 1	1,-1
$\mathcal{P}$ (papel)	1,-1	0, 0	-1, 1
$\mathcal{S}$ (tijera)	-1, 1	1,-1	0, 0

Equilibrio de Nash

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	0, 0	-1, 1	1,-1
$\mathcal{P}$ (papel)	1,-1	0, 0	-1, 1
$\mathcal{S}$ (tijera)	-1, 1	1,-1	0, 0

Equilibrio de Nash

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	0, 0	-1, 1	1,-1
$\mathcal{P}$ (papel)	1,-1	0, 0	-1, 1
$\mathcal{S}$ (tijera)	-1, 1	1,-1	0, 0

Equilibrio de Nash

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	0, 0	-1, 1	1,-1
$\mathcal{P}$ (papel)	1,-1	0, 0	-1, 1
$\mathcal{S}$ (tijera)	-1, 1	1,-1	0, 0

Equilibrio de Nash

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	0, 0	-1, 1	1,-1
$\mathcal{P}$ (papel)	1,-1	0, 0	-1, 1
$\mathcal{S}$ (tijera)	-1, 1	1,-1	0, 0

Equilibrio de Nash

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	0, 0	-1, 1	1,-1
$\mathcal{P}$ (papel)	1,-1	0, 0	-1, 1
$\mathcal{S}$ (tijera)	-1, 1	1,-1	0, 0

No todos los juegos tienen un equilibrio de Nash en estrategias puras.

Equilibrio de Nash

Piedra, papel o tijera

	$\mathcal{R}$ (piedra)	$\mathcal{P}$ (papel)	$\mathcal{S}$ (tijera)
$\mathcal{R}$ (piedra)	0, 0	-1, 1	1,-1
$\mathcal{P}$ (papel)	1,-1	0, 0	-1, 1
$\mathcal{S}$ (tijera)	-1, 1	1,-1	0, 0

No todos los juegos tienen un equilibrio de Nash en estrategias puras.

Teorema de Nash

Todo juego finito tiene al menos un equilibrio de Nash (en estrategias mixtas).

Juegos de Dos Jugadores de Suma Cero

Juegos de Dos Jugadores de Suma Cero

- Equilibrio de Nash como principal concepto de solución.

Juegos de Dos Jugadores de Suma Cero

- Equilibrio de Nash como principal concepto de solución.
- Valor del juego (u): ganancia del primer jugador cuando ambos jugadores utilizan un equilibrio de Nash.

Juegos de Dos Jugadores de Suma Cero

- Equilibrio de Nash como principal concepto de solución.
- Valor del juego (u): ganancia del primer jugador cuando ambos jugadores utilizan un equilibrio de Nash.
- El juego "batalla de los sexos" no cumple estas condiciones.

Regret Matching: Cálculo del Equilibrio de Nash

Esquema General:

Regret Matching: Cálculo del Equilibrio de Nash

Esquema General:

1. Se juega de forma repetida a través del tiempo $t = 1, 2, 3, ...$

Regret Matching: Cálculo del Equilibrio de Nash

Esquema General:

1. Se juega de forma repetida a través del tiempo $t = 1, 2, 3, \ldots.$
2. A tiempo $t + 1$ cada jugador elige una acción siguiendo una estrategia mixta **determinada**.

Regret Matching: Cálculo del Equilibrio de Nash

Esquema General:

1. Se juega de forma repetida a través del tiempo $t = 1, 2, 3, ...$
2. A tiempo $t + 1$ cada jugador elige una acción siguiendo una estrategia mixta **determinada**.
3. La **estrategia empírica** converge a un equilibrio de Nash; en la práctica el algoritmo es detenido después de cierto número de iteraciones.

Regret Matching: Cálculo del Equilibrio de Nash

Esquema General:

1. Se juega de forma repetida a través del tiempo $t = 1, 2, 3, ...$
2. A tiempo $t + 1$ cada jugador elige una acción siguiendo una estrategia mixta **determinada**.
3. La **estrategia empírica** converge a un equilibrio de Nash; en la práctica el algoritmo es detenido después de cierto número de iteraciones.

Diferentes formas de calcular la estrategia mixta conducen a diferentes algoritmos:

1. Regret condicional.
2. Regret incondicional.
3. Vector invariante de probabilidad.

Regret Matching: Procedimientos

Regret

Métrica de **arrepentimiento** *de no haber elegido una acción en particular.*

Regret Matching: Procedimientos

Regret

Métrica de **arrepentimiento** *de no haber elegido una acción en particular.*

Tres procedimientos

$\mathcal{R},\mathcal{S}$	$\mathcal{R},\mathcal{P}$	$\mathcal{S},\mathcal{S}$	$\bar{u}$
1	-1	0	0

Regret Matching: Procedimientos

Regret

Métrica de **arrepentimiento** *de no haber elegido una acción en particular.*

Tres procedimientos

1. Regret condicional.

$\mathcal{R}, \mathcal{S}$	$\mathcal{R}, \mathcal{P}$	$\mathcal{S}, \mathcal{S}$	$\bar{u}$
1	-1	0	0

Regret Matching: Procedimientos

Regret

Métrica de **arrepentimiento** *de no haber elegido una acción en particular.*

Tres procedimientos

1. Regret condicional.

$\mathcal{R},\mathcal{S}$	$\mathcal{R},\mathcal{P}$	$\mathcal{S},\mathcal{S}$	$\bar{u}$
1	-1	0	0

$\mathcal{S},\mathcal{S}$	$\mathcal{S},\mathcal{P}$	$\mathcal{S},\mathcal{S}$	$\bar{u}$
0	1	0	$\frac{1}{3}$

$$R_1(\mathcal{R},\mathcal{S}) = \frac{1}{3} - 0 = \frac{1}{3}$$

Regret Matching: Procedimientos

Regret

Métrica de **arrepentimiento** *de no haber elegido una acción en particular.*

Tres procedimientos

1. Regret condicional.

$\mathcal{R}, \mathcal{S}$	$\mathcal{R}, \mathcal{P}$	$\mathcal{S}, \mathcal{S}$	$\bar{u}$
1	-1	0	0

$\mathcal{P}, \mathcal{S}$	$\mathcal{P}, \mathcal{P}$	$\mathcal{S}, \mathcal{S}$	$\bar{u}$
-1	0	0	-$\frac{1}{3}$

$$R_1(\mathcal{R}, \mathcal{P}) = -\frac{1}{3} - 0 = -\frac{1}{3}$$

Regret Matching: Procedimientos

Regret

Métrica de **arrepentimiento** *de no haber elegido una acción en particular.*

Tres procedimientos

1. Regret condicional.
2. Regret incondicional.

$\mathcal{R},\mathcal{S}$	$\mathcal{R},\mathcal{P}$	$\mathcal{S},\mathcal{S}$	$\bar{u}$
1	-1	0	0

Regret Matching: Procedimientos

Regret

Métrica de ***arrepentimiento*** *de no haber elegido una acción en particular.*

Tres procedimientos

1. Regret condicional.
2. Regret incondicional.

$\mathcal{R}, \mathcal{S}$	$\mathcal{R}, \mathcal{P}$	$\mathcal{S}, \mathcal{S}$	$\bar{u}$
1	-1	0	0

$\mathcal{S}, \mathcal{S}$	$\mathcal{S}, \mathcal{P}$	$\mathcal{S}, \mathcal{S}$	$\bar{u}$
0	1	0	$\frac{1}{3}$

$$R_1(\mathcal{S}) = \frac{1}{3} - 0 = \frac{1}{3}$$

Regret Matching: Procedimientos

Regret

Métrica de **arrepentimiento** *de no haber elegido una acción en particular.*

Tres procedimientos

1. Regret condicional.
2. Regret incondicional.

$\mathcal{R},\mathcal{S}$	$\mathcal{R},\mathcal{P}$	$\mathcal{S},\mathcal{S}$	$\bar{u}$
1	-1	0	0

$\mathcal{P},\mathcal{S}$	$\mathcal{P},\mathcal{P}$	$\mathcal{P},\mathcal{S}$	$\bar{u}$
-1	0	-1	-$\frac{2}{3}$

$$R_1(\mathcal{P}) = -\frac{2}{3} - 0 = -\frac{2}{3}$$

Regret Matching: Procedimientos

Regret

Métrica de **arrepentimiento** *de no haber elegido una acción en particular.*

Tres procedimientos

1. Regret condicional.
2. Regret incondicional.
3. Vector invariante de probabilidad.

$\mathcal{R}, \mathcal{S}$	$\mathcal{R}, \mathcal{P}$	$\mathcal{S}, \mathcal{S}$	$\bar{u}$
1	-1	0	0

Regret Matching: Observaciones

Regret Matching: Observaciones

1. Las probabilidades son elegidas proporcional a los regrets positivos.

Regret Matching: Observaciones

1. Las probabilidades son elegidas proporcional a los regrets positivos.
2. El regret positivo tiende a cero cuando el número de juegos tiende a infinito.

Regret Matching: Observaciones

1. Las probabilidades son elegidas proporcional a los regrets positivos.
2. El regret positivo tiende a cero cuando el número de juegos tiende a infinito.
3. Si el regret positivo es pequeño, la **estrategia empírica** es una aproximación a un equilibrio de Nash.

Regret Matching: Observaciones

1. Las probabilidades son elegidas proporcional a los regrets positivos.
2. El regret positivo tiende a cero cuando el número de juegos tiende a infinito.
3. Si el regret positivo es pequeño, la **estrategia empírica** es una aproximación a un equilibrio de Nash.

Estrategia Empírica

La probabilidad de que un determinado jugador elija una acción a es igual a:

$$p(a) = \frac{\text{número de veces que el jugador eligió } a}{\text{número total de juegos}}$$

Regret Matching: Experimentos

Regret Matching: Experimentos

1. 4 juegos para dos jugadores de suma cero.

Regret Matching: Experimentos

1. 4 juegos para dos jugadores de suma cero.
2. 10 corridas por cada uno de los juegos y cada uno de los procedimientos.

Regret Matching: Experimentos

1. 4 juegos para dos jugadores de suma cero.
2. 10 corridas por cada uno de los juegos y cada uno de los procedimientos.
3. Criterio de parada: regret incondicional menor que 0,005.

Regret Matching: Experimentos

1. 4 juegos para dos jugadores de suma cero.
2. 10 corridas por cada uno de los juegos y cada uno de los procedimientos.
3. Criterio de parada: regret incondicional menor que 0,005.
4. Gráficas del regret con respecto al número de iteraciones.

Regret Matching: Experimentos

1. 4 juegos para dos jugadores de suma cero.
2. 10 corridas por cada uno de los juegos y cada uno de los procedimientos.
3. Criterio de parada: regret incondicional menor que 0,005.
4. Gráficas del regret con respecto al número de iteraciones.
5. Verificación realizada con programación lineal.

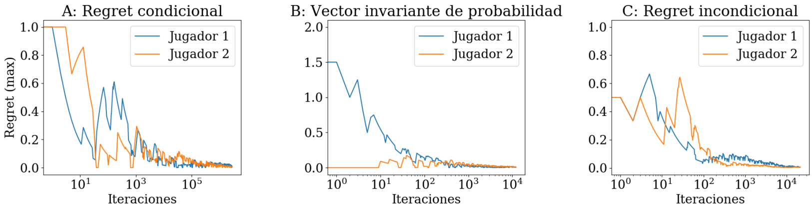
Regret Matching: Experimentos

1. 4 juegos para dos jugadores de suma cero.
2. 10 corridas por cada uno de los juegos y cada uno de los procedimientos.
3. Criterio de parada: regret incondicional menor que 0,005.
4. Gráficas del regret con respecto al número de iteraciones.
5. Verificación realizada con programación lineal.
6. Verificación: la explotabilidad mide la distancia entre la estrategia actual y un equilibrio de Nash.

Piedra, Papel o Tijera

Valor del juego (u): 0.

	A	B	C
Ganancia esperada $u(\sigma)$	$-0,000012$	$0,000004$	$0,000022$
Explotabilidad ε_σ	$0,006$	$0,010$	$0,009$
Tiempo T	$12,198$	$0,345$	$0,049$
Iteraciones I	$4.519.054,1$	$6.601,3$	$19.321,1$
T/I	$2,70\times10^{-6}$	$5,23\times10^{-5}$	$2,54\times10^{-6}$



Matching Pennies

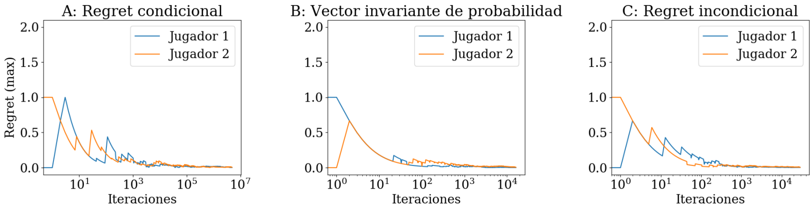
1. Cada jugador posee una moneda y elige cara o sello.
2. Misma elección: gana el primer jugador.
3. Elecciones diferentes: gana el segundo jugador.
4. Tabla de pagos.

	cara	sello
cara	1	-1
sello	-1	1

Matching Pennies

Valor del juego (u): 0.

	A	B	C
Ganancia esperada $u(\sigma)$	$0,000$	$0,000$	$0,000$
Explotabilidad ε_σ	$0,006$	$0,006$	$0,008$
Tiempo T	$10,276$	$0,777$	$0,042$
Iteraciones I	$3.892.550,4$	$25.616,6$	$16.260,5$
T/I	$2,64\times10^{-6}$	$3,03\times10^{-5}$	$2,58\times10^{-6}$



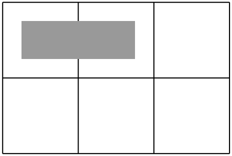
Ficha vs. Dominó

Jugador 1

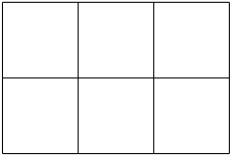
Jugador 2

Ficha vs. Dominó

Jugador 1

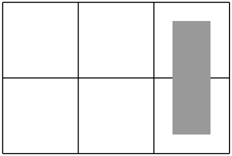


Jugador 2

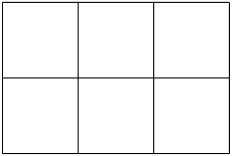


Ficha vs. Dominó

Jugador 1

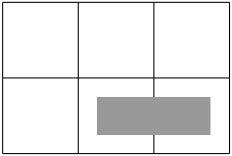


Jugador 2

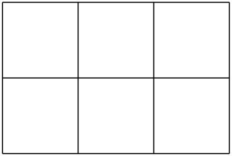


Ficha vs. Dominó

Jugador 1

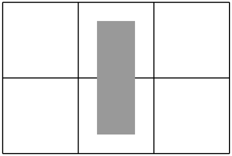


Jugador 2

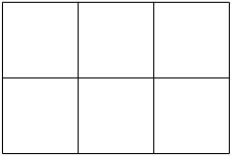


Ficha vs. Dominó

Jugador 1

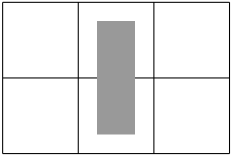


Jugador 2

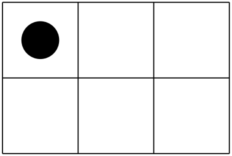


Ficha vs. Dominó

Jugador 1

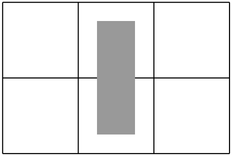


Jugador 2

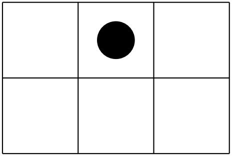


Ficha vs. Dominó

Jugador 1

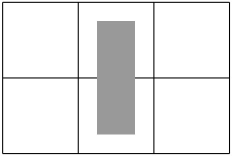


Jugador 2

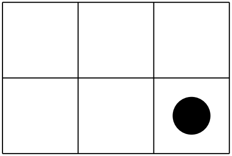


Ficha vs. Dominó

Jugador 1

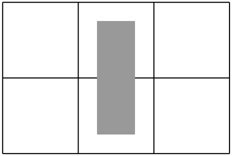


Jugador 2

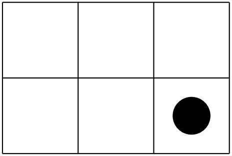


Ficha vs. Dominó

Jugador 1

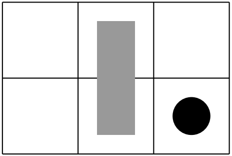


Jugador 2



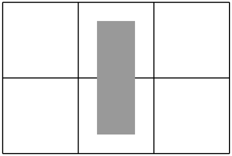
Resultado

1. La ficha y el dominó no se superponen: gana el jugador 1.

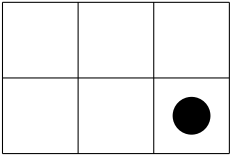


Ficha vs. Dominó

Jugador 1

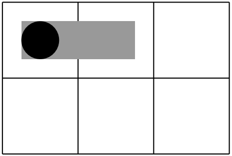


Jugador 2



Resultado

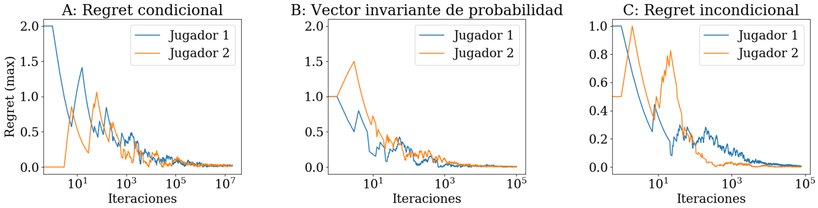
1. La ficha y el dominó no se superponen: gana el jugador 1.
2. La ficha y el dominó sí se superponen: gana el jugador 2.



Ficha vs. Dominó

Valor del juego (u): $\frac{1}{3}$.

	A	B	C
Ganancia esperada $u(\sigma)$	$0,333$	$0,334$	$0,334$
Explotabilidad ε_σ	$0,010$	$0,007$	$0,004$
Tiempo T	$319,179$	$11,275$	$0,237$
Iteraciones I	$108.319.272,4$	$75.250,2$	$84.318,5$
T/I	$2,95\times10^{-6}$	$1,50\times10^{-4}$	$2,81\times10^{-6}$

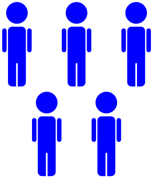


Coronel Blotto

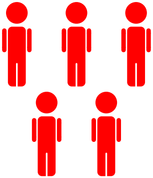
Coronel Blotto

- S soldados por jugador.

Jugador 1

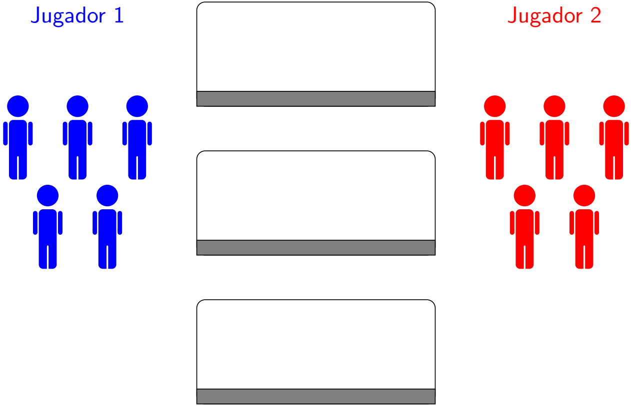


Jugador 2



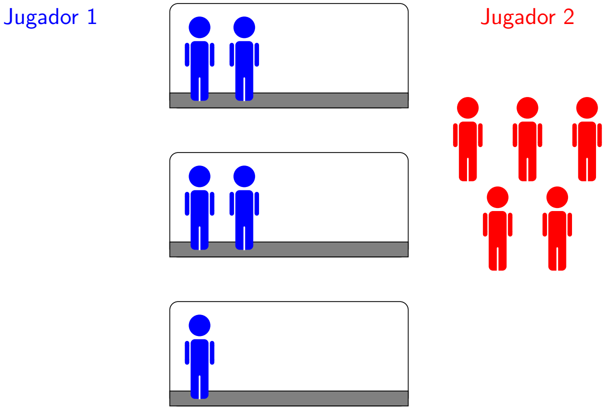
Coronel Blotto

- S soldados por jugador.
- N campos de batalla.



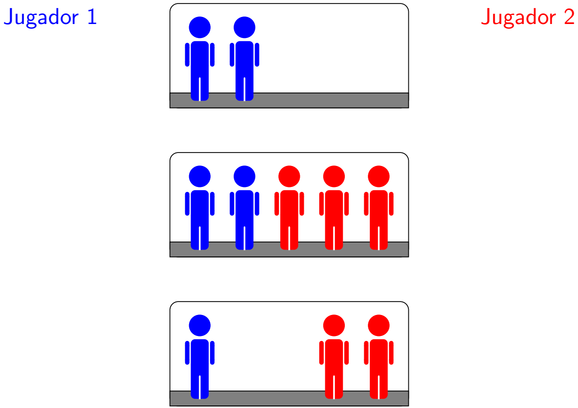
Coronel Blotto

- S soldados por jugador.
- N campos de batalla.



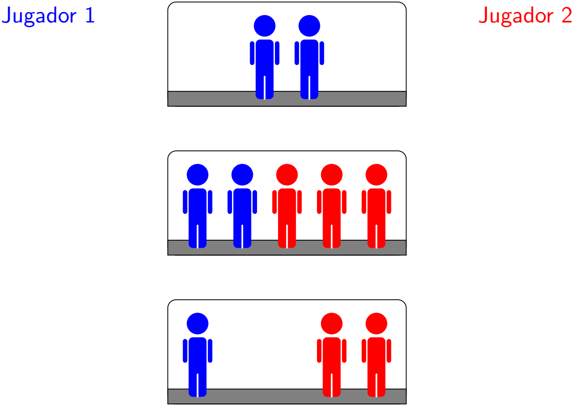
Coronel Blotto

- S soldados por jugador.
- N campos de batalla.



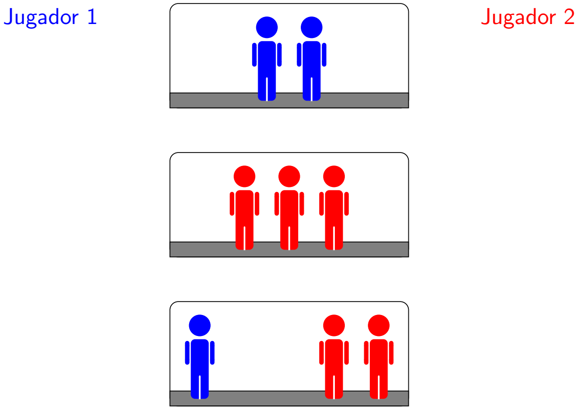
Coronel Blotto

- S soldados por jugador.
- N campos de batalla.



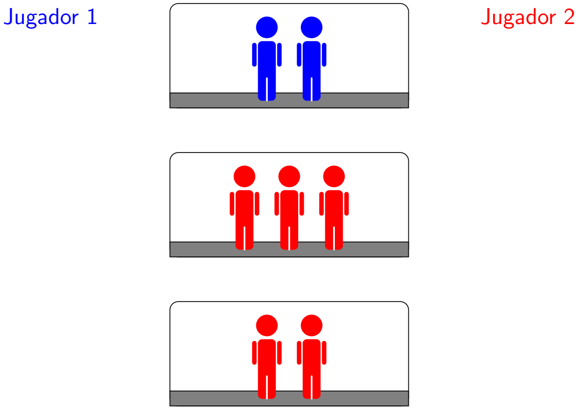
Coronel Blotto

- S soldados por jugador.
- N campos de batalla.



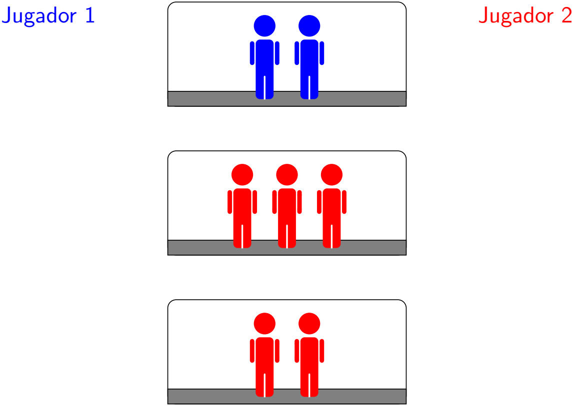
Coronel Blotto

- S soldados por jugador.
- N campos de batalla.



Coronel Blotto

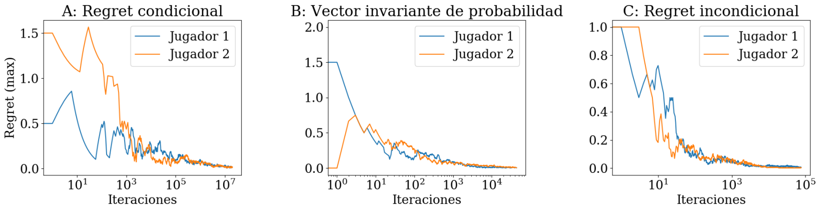
- S soldados por jugador.
- N campos de batalla.
- $u_1 = 1 - 2 = -1$
- $u_2 = 2 - 1 = 1.$



Coronel Blotto

Valor del juego (u): 0.

	A	B	C
Ganancia esperada $u(\sigma)$	$0,000219$	$0,000150$	$0,000024$
Explotabilidad ε_σ	$0,010$	$0,010$	$0,009$
Tiempo T	$875,533$	$70,453$	$0,166$
Iteraciones I	$190.222.305,3$	$58.794,4$	$48.613,5$
T/I	$4,60\times10^{-6}$	$1,20\times10^{-3}$	$3,41\times10^{-6}$



Forma Normal y Regret Matching: Conclusiones

Forma Normal y Regret Matching: Conclusiones

1. El equilibrio de Nash es un concepto de solución satisfactorio en juegos en forma normal de dos jugadores de suma cero.

Forma Normal y Regret Matching: Conclusiones

1. El equilibrio de Nash es un concepto de solución satisfactorio en juegos en forma normal de dos jugadores de suma cero.
2. Los algoritmos de Regret Matching permiten encontrar aproximaciones de un equilibrio de Nash en este tipo de juegos.

Forma Normal y Regret Matching: Conclusiones

1. El equilibrio de Nash es un concepto de solución satisfactorio en juegos en forma normal de dos jugadores de suma cero.
2. Los algoritmos de Regret Matching permiten encontrar aproximaciones de un equilibrio de Nash en este tipo de juegos.
3. El procedimiento "regret condicional" necesita mayor número de iteraciones, y por lo tanto mayor tiempo, para converger.

Forma Normal y Regret Matching: Conclusiones

1. El equilibrio de Nash es un concepto de solución satisfactorio en juegos en forma normal de dos jugadores de suma cero.
2. Los algoritmos de Regret Matching permiten encontrar aproximaciones de un equilibrio de Nash en este tipo de juegos.
3. El procedimiento “regret condicional” necesita mayor número de iteraciones, y por lo tanto mayor tiempo, para converger.
4. El procedimiento “vector invariante de probabilidad” es el procedimiento con las iteraciones más costosas.

Forma Normal y Regret Matching: Conclusiones

1. El equilibrio de Nash es un concepto de solución satisfactorio en juegos en forma normal de dos jugadores de suma cero.
2. Los algoritmos de Regret Matching permiten encontrar aproximaciones de un equilibrio de Nash en este tipo de juegos.
3. El procedimiento “regret condicional” necesita mayor número de iteraciones, y por lo tanto mayor tiempo, para converger.
4. El procedimiento “vector invariante de probabilidad” es el procedimiento con las iteraciones más costosas.
5. El procedimiento “regret incondicional” es el más eficiente.

Forma Normal y Regret Matching: Conclusiones

1. El equilibrio de Nash es un concepto de solución satisfactorio en juegos en forma normal de dos jugadores de suma cero.
2. Los algoritmos de Regret Matching permiten encontrar aproximaciones de un equilibrio de Nash en este tipo de juegos.
3. El procedimiento “regret condicional” necesita mayor número de iteraciones, y por lo tanto mayor tiempo, para converger.
4. El procedimiento “vector invariante de probabilidad” es el procedimiento con las iteraciones más costosas.
5. El procedimiento “regret incondicional” es el más eficiente.
6. Si el regret incondicional es menor a 0,005 por jugador, la estrategia empírica tiene una explotabilidad no mayor que 0,01.

Forma Normal y Regret Matching: Conclusiones

1. El equilibrio de Nash es un concepto de solución satisfactorio en juegos en forma normal de dos jugadores de suma cero.
2. Los algoritmos de Regret Matching permiten encontrar aproximaciones de un equilibrio de Nash en este tipo de juegos.
3. El procedimiento "regret condicional" necesita mayor número de iteraciones, y por lo tanto mayor tiempo, para converger.
4. El procedimiento "vector invariante de probabilidad" es el procedimiento con las iteraciones más costosas.
5. El procedimiento "regret incondicional" es el más eficiente.
6. Si el regret incondicional es menor a 0,005 por jugador, la estrategia empírica tiene una explotabilidad no mayor que 0,01.
7. Fue posible resolver diversos juegos en forma normal.

Forma Extensiva

Forma Extensiva: Modelo

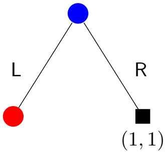
Juegos secuenciales

Forma Extensiva: Modelo

Juegos secuenciales

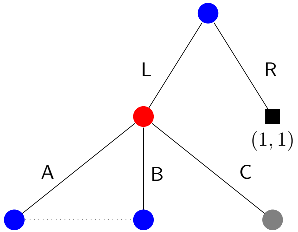
Forma Extensiva: Modelo

Juegos secuenciales



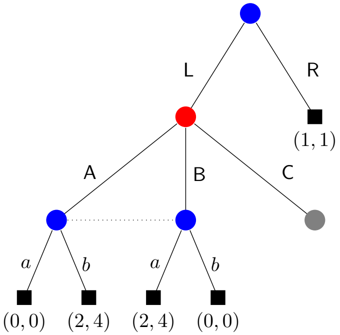
Forma Extensiva: Modelo

Juegos secuenciales



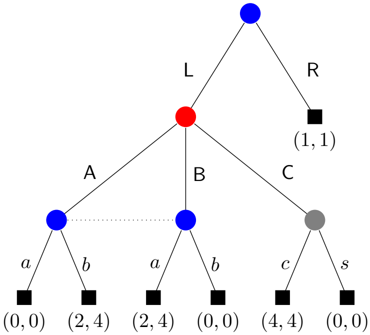
Forma Extensiva: Modelo

Juegos secuenciales



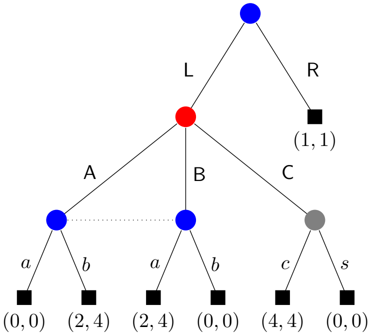
Forma Extensiva: Modelo

Juegos secuenciales



Forma Extensiva: Modelo

Juegos secuenciales

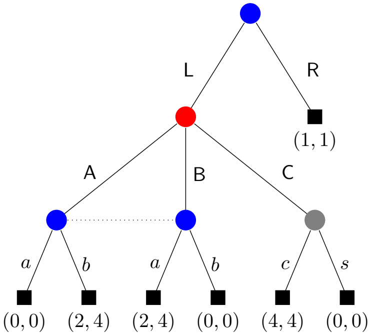


Elementos

1. Historias o nodos.
   Ej: $\emptyset$, LA, LBb, R.

Forma Extensiva: Modelo

Juegos secuenciales

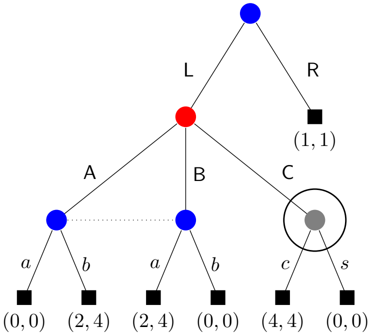


Elementos

1. Historias o nodos.
   Ej: $\emptyset$, LA, LBb, R.
2. Función que asigna a cada historia (nodo) no terminal un jugador.

Forma Extensiva: Modelo

Juegos secuenciales

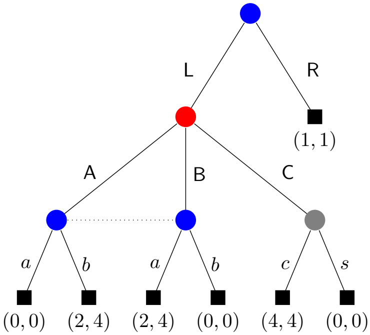


Elementos

1. Historias o nodos.
   Ej: $\emptyset$, LA, LBb, R.
2. Función que asigna a cada historia (nodo) no terminal un jugador.
   - Nodos de azar.

Forma Extensiva: Modelo

Juegos secuenciales

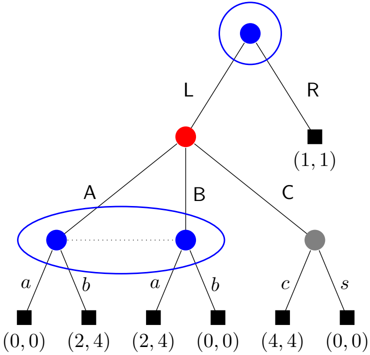


Elementos

1. Historias o nodos.
   Ej: $\emptyset$, LA, LBb, R.
2. Función que asigna a cada historia (nodo) no terminal un jugador.
   - Nodos de azar.
3. Conjuntos de información.

Forma Extensiva: Modelo

Juegos secuenciales

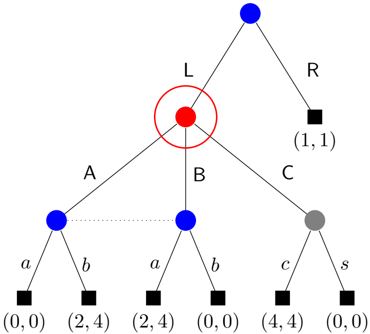


Elementos

1. Historias o nodos.
   Ej: $\emptyset$, LA, LBb, R.
2. Función que asigna a cada historia (nodo) no terminal un jugador.
   - Nodos de azar.
3. Conjuntos de información.

Forma Extensiva: Modelo

Juegos secuenciales

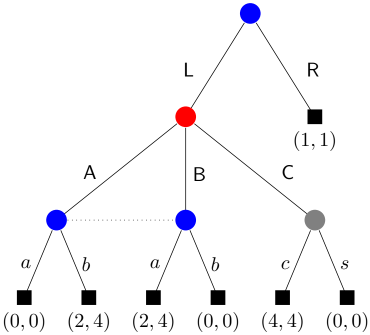


Elementos

1. Historias o nodos.
   Ej: $\emptyset$, LA, LBb, R.
2. Función que asigna a cada historia (nodo) no terminal un jugador.
   - Nodos de azar.
3. Conjuntos de información.

Forma Extensiva: Modelo

Juegos secuenciales

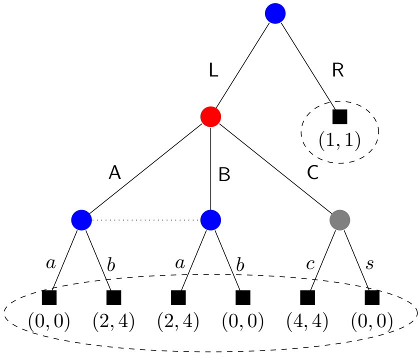


Elementos

1. Historias o nodos.
   Ej: $\emptyset$, LA, LBb, R.
2. Función que asigna a cada historia (nodo) no terminal un jugador.
   - Nodos de azar.
3. Conjuntos de información.
4. Función que asigna por cada historia (nodo) terminal y cada jugador una utilidad.

Forma Extensiva: Modelo

Juegos secuenciales

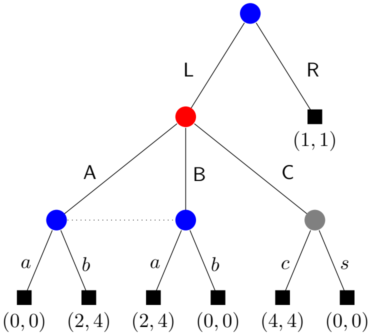


Elementos

1. Historias o nodos.
   Ej: $\emptyset$, LA, LBb, R.
2. Función que asigna a cada historia (nodo) no terminal un jugador.
   - Nodos de azar.
3. Conjuntos de información.
4. Función que asigna por cada historia (nodo) terminal y cada jugador una utilidad.

Forma Extensiva: Modelo

Juegos secuenciales

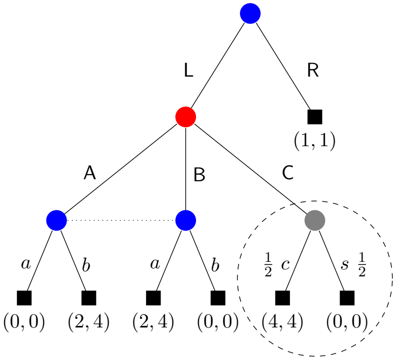


Elementos

1. Historias o nodos.
   Ej: $\emptyset$, LA, LBb, R.
2. Función que asigna a cada historia (nodo) no terminal un jugador.
   - Nodos de azar.
3. Conjuntos de información.
4. Función que asigna por cada historia (nodo) terminal y cada jugador una utilidad.
5. Distribución de probabilidad sobre el conjunto de acciones en cada nodo de azar.

Forma Extensiva: Modelo

Juegos secuenciales

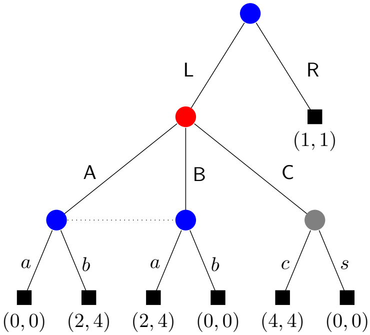


Elementos

1. Historias o nodos.
   Ej: $\emptyset$, LA, LBb, R.
2. Función que asigna a cada historia (nodo) no terminal un jugador.
   - Nodos de azar.
3. Conjuntos de información.
4. Función que asigna por cada historia (nodo) terminal y cada jugador una utilidad.
5. Distribución de probabilidad sobre el conjunto de acciones en cada nodo de azar.

Forma Extensiva: Modelo

Juegos secuenciales



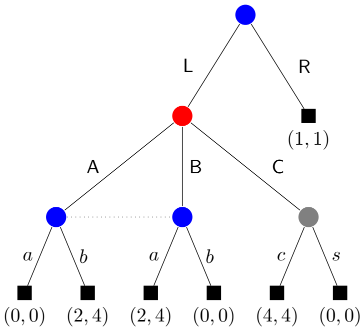
Perfect Recall

Se recuerda de forma perfecta todo lo que se ha visto.

Elementos

1. Historias o nodos. Ej: $\emptyset$, LA, LBb, R.
2. Función que asigna a cada historia (nodo) no terminal un jugador.
    - Nodos de azar.
3. Conjuntos de información.
4. Función que asigna por cada historia (nodo) terminal y cada jugador una utilidad.
5. Distribución de probabilidad sobre el conjunto de acciones en cada nodo de azar.

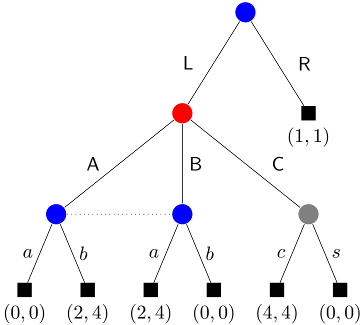
Forma Extensiva: Estrategias



Forma Normal

	A	B	C
(L, a)	$0, 0$	$2, 4$	$2, 2$
(L, b)	$2, 4$	$0, 0$	$2, 2$
(R, a)	$1, 1$	$1, 1$	$1, 1$
(R, b)	$1, 1$	$1, 1$	$1, 1$

Forma Extensiva: Estrategias



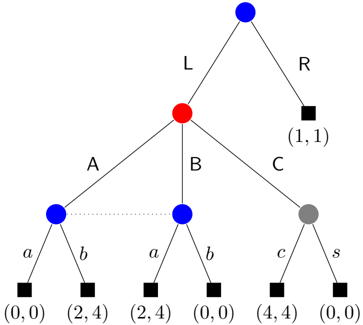
Forma Normal

	A	B	C
(L, a)	$0, 0$	$2, 4$	$2, 2$
(L, b)	$2, 4$	$0, 0$	$2, 2$
(R, a)	$1, 1$	$1, 1$	$1, 1$
(R, b)	$1, 1$	$1, 1$	$1, 1$

Estrategias

1. Estrategias Puras.

Forma Extensiva: Estrategias



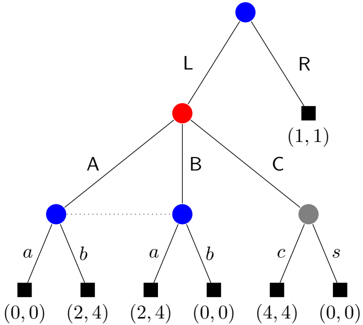
Estrategias

1. Estrategias Puras.

Forma Normal

	A	B	C
(L, a)	$0, 0$	$2, 4$	$2, 2$
(L, b)	$2, 4$	$0, 0$	$2, 2$
(R, a)	$1, 1$	$1, 1$	$1, 1$
(R, b)	$1, 1$	$1, 1$	$1, 1$

Forma Extensiva: Estrategias



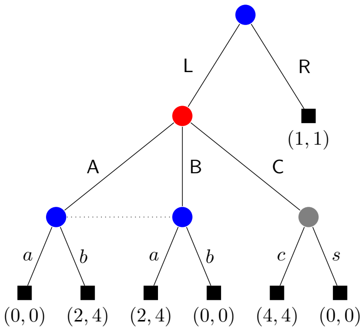
Forma Normal

	A	B	C
(L, a)	$0, 0$	$2, 4$	$2, 2$
(L, b)	$2, 4$	$0, 0$	$2, 2$
(R, a)	$1, 1$	$1, 1$	$1, 1$
(R, b)	$1, 1$	$1, 1$	$1, 1$

Estrategias

1. Estrategias Puras.

Forma Extensiva: Estrategias



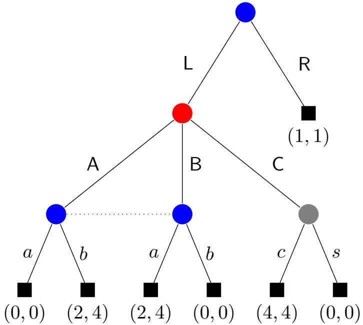
Estrategias

1. Estrategias Puras.
2. Estrategias Mixtas.

Forma Normal

	A	B	C
(L, a)	$0,0$	$2,4$	$2,2$
(L, b)	$2,4$	$0,0$	$2,2$
(R, a)	$1,1$	$1,1$	$1,1$
(R, b)	$1,1$	$1,1$	$1,1$

Forma Extensiva: Estrategias



Estrategias

1. Estrategias Puras.
2. Estrategias Mixtas.

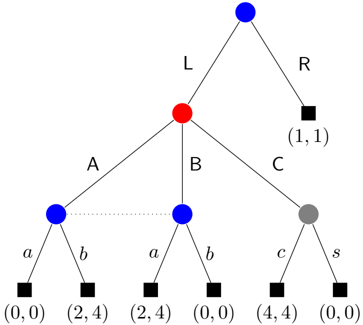
(L, a)	(L, b)	(R, a)	(R, b)
0.45	0.30	0.00	0.25

A	B	C
0.25	0.25	0.50

Forma Normal

	A	B	C
(L, a)	$0, 0$	$2, 4$	$2, 2$
(L, b)	$2, 4$	$0, 0$	$2, 2$
(R, a)	$1, 1$	$1, 1$	$1, 1$
(R, b)	$1, 1$	$1, 1$	$1, 1$

Forma Extensiva: Estrategias



Forma Normal

	A	B	C
(L, a)	$0, 0$	$2, 4$	$2, 2$
(L, b)	$2, 4$	$0, 0$	$2, 2$
(R, a)	$1, 1$	$1, 1$	$1, 1$
(R, b)	$1, 1$	$1, 1$	$1, 1$

Estrategias

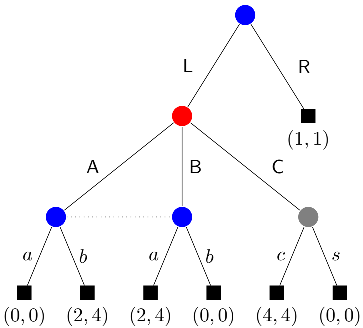
1. Estrategias Puras.
2. Estrategias Mixtas.

(L, a)	(L, b)	(R, a)	(R, b)
0.45	0.30	0.00	0.25

A	B	C
0.25	0.25	0.50

3. Estrategias de Comportamiento.

Forma Extensiva: Estrategias



Forma Normal

	A	B	C
(L, a)	$0, 0$	$2, 4$	$2, 2$
(L, b)	$2, 4$	$0, 0$	$2, 2$
(R, a)	$1, 1$	$1, 1$	$1, 1$
(R, b)	$1, 1$	$1, 1$	$1, 1$

Estrategias

1. Estrategias Puras.
2. Estrategias Mixtas.

(L, a)	(L, b)	(R, a)	(R, b)
0.45	0.30	0.00	0.25

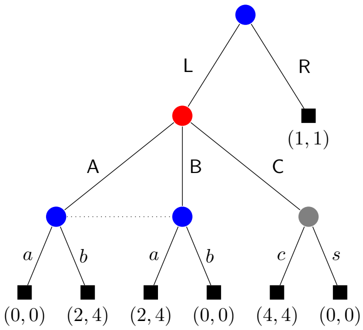
A	B	C
0.25	0.25	0.50

3. Estrategias de Comportamiento.

L	R
0.65	0.35

a	b
0.40	0.60

Forma Extensiva: Estrategias



Forma Normal

	A	B	C
(L, a)	$0, 0$	$2, 4$	$2, 2$
(L, b)	$2, 4$	$0, 0$	$2, 2$
(R, a)	$1, 1$	$1, 1$	$1, 1$
(R, b)	$1, 1$	$1, 1$	$1, 1$

Estrategias

1. Estrategias Puras.
2. Estrategias Mixtas.

(L, a)	(L, b)	(R, a)	(R, b)
0.45	0.30	0.00	0.25

A	B	C
0.25	0.25	0.50

3. Estrategias de Comportamiento.

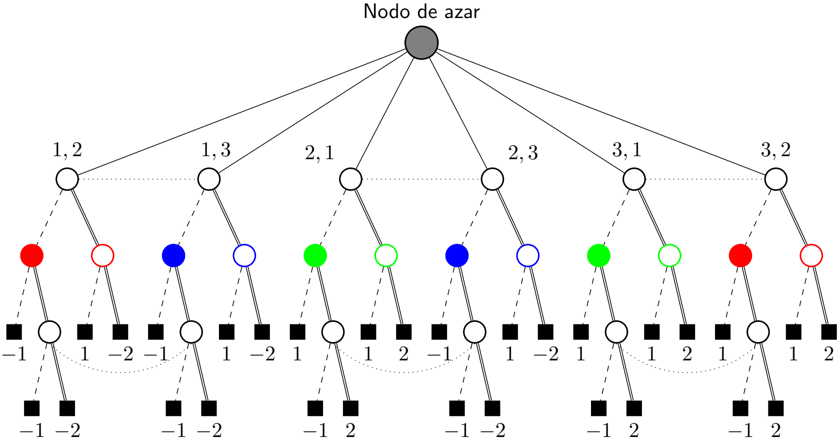
L	R
0.65	0.35

a	b
0.40	0.60

Equilibrio de Nash

Mejor respuesta frente a la estrategias de sus oponentes.

Kuhn Poker



Nodo de azar
1, 2
1, 3
2, 1
2, 3
3, 1
3, 2
−1 1 −2 −1 1 −2 1 1 2 −1 1 −2 1 1 2 1 1 2
−1 −2 −1 −2 −1 2 −1 −2 −1 2 −1 2

Counterfactual Regret Minimization (CFR)

Counterfactual Regret Minimization (CFR)

1. El árbol del juego se recorre repetidamente.

Counterfactual Regret Minimization (CFR)

1. El árbol del juego se recorre repetidamente.
2. Inicia con una estrategia uniforme en cada conjunto de información.

Counterfactual Regret Minimization (CFR)

1. El árbol del juego se recorre repetidamente.
2. Inicia con una estrategia uniforme en cada conjunto de información.
3. En cada iteración se elige una mejor estrategia revisando las decisiones pasadas, utilizando las métricas de regret.

Counterfactual Regret Minimization (CFR)

1. El árbol del juego se recorre repetidamente.
2. Inicia con una estrategia uniforme en cada conjunto de información.
3. En cada iteración se elige una mejor estrategia revisando las decisiones pasadas, utilizando las métricas de regret.
4. La estrategia promedio converge a un equilibrio de Nash; en la práctica el algoritmo es detenido después de cierto número de iteraciones.

Counterfactual Regret Minimization (CFR)

1. El árbol del juego se recorre repetidamente.
2. Inicia con una estrategia uniforme en cada conjunto de información.
3. En cada iteración se elige una mejor estrategia revisando las decisiones pasadas, utilizando las métricas de regret.
4. La estrategia promedio converge a un equilibrio de Nash; en la práctica el algoritmo es detenido después de cierto número de iteraciones.

¿Cómo se mejora la estrategia en cada iteración?

Counterfactual Regret Minimization (CFR)

1. El árbol del juego se recorre repetidamente.
2. Inicia con una estrategia uniforme en cada conjunto de información.
3. En cada iteración se elige una mejor estrategia revisando las decisiones pasadas, utilizando las métricas de regret.
4. La estrategia promedio converge a un equilibrio de Nash; en la práctica el algoritmo es detenido después de cierto número de iteraciones.

¿Cómo se mejora la estrategia en cada iteración?

- Sumar el regret contrafactual que se tiene en cada conjunto de información por cada acción.

Counterfactual Regret Minimization (CFR)

1. El árbol del juego se recorre repetidamente.
2. Inicia con una estrategia uniforme en cada conjunto de información.
3. En cada iteración se elige una mejor estrategia revisando las decisiones pasadas, utilizando las métricas de regret.
4. La estrategia promedio converge a un equilibrio de Nash; en la práctica el algoritmo es detenido después de cierto número de iteraciones.

¿Cómo se mejora la estrategia en cada iteración?

- Sumar el regret contrafactual que se tiene en cada conjunto de información por cada acción.
- **Regret contrafactual**: cuánto mejor lo habría hecho en todos los juegos hasta ahora si siempre hubiera jugado esta acción en este conjunto de información.

Counterfactual Regret Minimization (CFR)

1. El árbol del juego se recorre repetidamente.
2. Inicia con una estrategia uniforme en cada conjunto de información.
3. En cada iteración se elige una mejor estrategia revisando las decisiones pasadas, utilizando las métricas de regret.
4. La estrategia promedio converge a un equilibrio de Nash; en la práctica el algoritmo es detenido después de cierto número de iteraciones.

¿Cómo se mejora la estrategia en cada iteración?

- Sumar el regret contrafactual que se tiene en cada conjunto de información por cada acción.
- **Regret contrafactual**: cuánto mejor lo habría hecho en todos los juegos hasta ahora si siempre hubiera jugado esta acción en este conjunto de información.
- **Regret Matching**: en la nueva estrategia las acciones son elegidas con probabilidades proporcionales a los regrets positivos.

CFR: Experimentos

CFR: Experimentos

1. CFR con muestreo en los nodos de azar.

CFR: Experimentos

1. CFR con muestreo en los nodos de azar.
2. Implementación propia del algoritmo.

CFR: Experimentos

1. CFR con muestreo en los nodos de azar.
2. Implementación propia del algoritmo.
   - Input: definición del juego.
   - **Definición del juego**: implementación de las funciones que permiten recorrer el árbol de forma ímplicita.

CFR: Experimentos

1. CFR con muestreo en los nodos de azar.
2. Implementación propia del algoritmo.
   - Input: definición del juego.
   - **Definición del juego**: implementación de las funciones que permiten recorrer el árbol de forma ímplicita.
3. Tres clases de juegos, cada uno con diferentes parámetros.

CFR: Experimentos

1. CFR con muestreo en los nodos de azar.
2. Implementación propia del algoritmo.
   - Input: definición del juego.
   - **Definición del juego**: implementación de las funciones que permiten recorrer el árbol de forma ímplicita.
3. Tres clases de juegos, cada uno con diferentes parámetros.
4. 10 horas de entrenamiento.

CFR: Experimentos

1. CFR con muestreo en los nodos de azar.
2. Implementación propia del algoritmo.
   - Input: definición del juego.
   - **Definición del juego**: implementación de las funciones que permiten recorrer el árbol de forma ímplicita.
3. Tres clases de juegos, cada uno con diferentes parámetros.
4. 10 horas de entrenamiento.
5. Gráfica del regret con respecto al número de iteraciones.

CFR: Experimentos

1. CFR con muestreo en los nodos de azar.
2. Implementación propia del algoritmo.
   - Input: definición del juego.
   - **Definición del juego**: implementación de las funciones que permiten recorrer el árbol de forma ímplicita.
3. Tres clases de juegos, cada uno con diferentes parámetros.
4. 10 horas de entrenamiento.
5. Gráfica del regret con respecto al número de iteraciones.
6. Verificación: la explotabilidad mide la distancia entre la estrategia actual y un equilibrio de Nash.

CFR: Experimentos

1. CFR con muestreo en los nodos de azar.
2. Implementación propia del algoritmo.
   - Input: definición del juego.
   - **Definición del juego**: implementación de las funciones que permiten recorrer el árbol de forma ímplicita.
3. Tres clases de juegos, cada uno con diferentes parámetros.
4. 10 horas de entrenamiento.
5. Gráfica del regret con respecto al número de iteraciones.
6. Verificación: la explotabilidad mide la distancia entre la estrategia actual y un equilibrio de Nash.
7. Un juego se considera resuelto si la explotabilidad es menor que el 1% de la mínima ganancia positiva posible.

One Card Poker (OCP)

One Card Poker (OCP)

- Generalización del Juego Kuhn Poker.

One Card Poker (OCP)

- Generalización del Juego Kuhn Poker.
- N: número de cartas.

One Card Poker (OCP)

- Generalización del Juego Kuhn Poker.
- N: número de cartas.
- OCP(N).

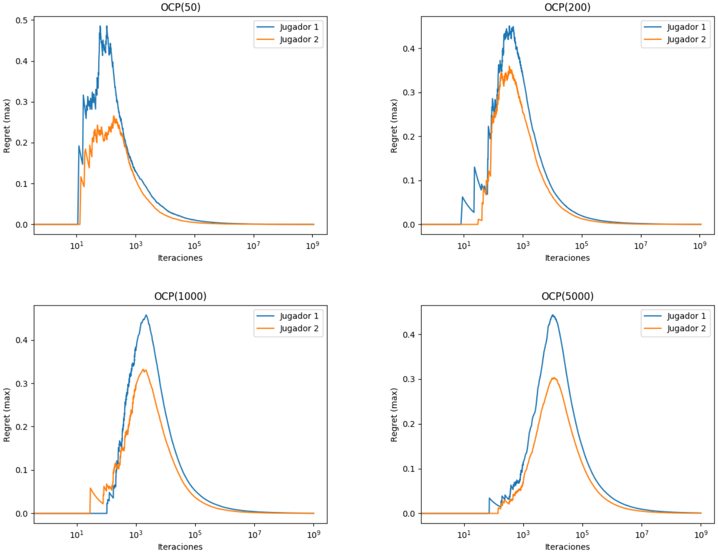
One Card Poker (OCP)

- Generalización del Juego Kuhn Poker.
- N: número de cartas.
- OCP(N).

Tabla de Resultados

Juego	N	I	Iteraciones	$u(\sigma)$	ε_σ (%)	Resuelto
OCP(3)	55	12	1.181.763.638	-0,056	0,0098	✓
OCP(12)	1.189	48	1.147.919.240	-0,062	0,0032	✓
OCP(50)	22.051	200	1.145.291.974	-0,058	0,0099	✓
OCP(200)	358.201	800	1.128.993.847	-0,056	0,0078	✓
OCP(1000)	8.991.001	4.000	1.087.573.694	-0,056	0,0098	✓
OCP(5000)	224.955.001	20.000	1.038.367.354	-0,056	0,0241	✓

One Card Poker (OCP)



Dudo

Dudo

- Juego de dados y apuestas.

Dudo

- Juego de dados y apuestas.
- K: número de caras de los dados.

Dudo

- Juego de dados y apuestas.
- K: número de caras de los dados.
- D_1, D_2: número de dados del primer y segundo jugador, respectivamente.

Dudo

- Juego de dados y apuestas.
- K: número de caras de los dados.
- D_1, D_2: número de dados del primer y segundo jugador, respectivamente.
- Dudo(K, D_1, D_2).

Dudo

- Juego de dados y apuestas.
- K: número de caras de los dados.
- D_1, D_2: número de dados del primer y segundo jugador, respectivamente.
- Dudo(K, D_1, D_2).

Tabla de Resultados

Juego	N	I	Iteraciones	$u(\sigma)$	ε_σ (%)	Resuelto
Dudo$(4,1,1)$	8.177	512	18.697.532	-0,125	0,0259	✓
Dudo$(4,1,2)$	327.641	14.366	1.215.600	-0,508	0,0971	✓
Dudo$(4,2,1)$	327.641	14.366	1.213.799	0,552	0,3701	✓
Dudo$(4,2,2)$	13.107.101	327.680	63.109	0,0069	2,1132	✗
Dudo$(5,1,1)$	51.176	2.560	4.521.208	-0,120	0,1186	✓
Dudo$(5,1,2)$	4.915.126	163.840	151.235	-0,565	0,6197	✓
Dudo$(5,2,1)$	4.915.126	163.840	143.698	0,581	0,0122	✓
Dudo$(5,2,2)$	471.858.976	7.864.320	3.826	0,836	15,1963	✗
Dudo$(6,1,1)$	294.877	12.288	1.067.782	-0,111	0,0975	✓
Dudo$(6,1,2)$	66.060.163	1.769.472	17.702	-0,593	4,5781	✗
Dudo$(6,2,1)$	66.060.163	1.769.472	17.221	0,592	3,9594	✗

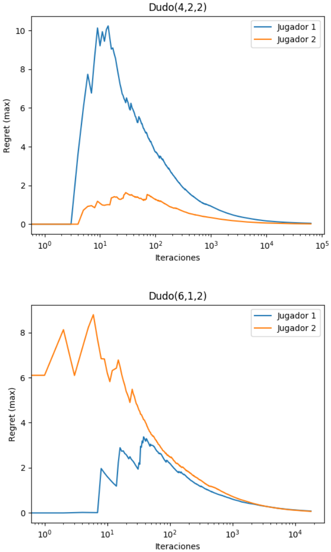
Dudo

- Juego de dados y apuestas.
- K: número de caras de los dados.
- D_1, D_2: número de dados del primer y segundo jugador, respectivamente.
- Dudo(K, D_1, D_2).

Tabla de Resultados

Juego	N	I	Iteraciones	$u(\sigma)$	ε_σ (%)	Resuelto
Dudo$(4,1,1)$	8.177	512	18.697.532	-0,125	0,0259	✓
Dudo$(4,1,2)$	327.641	14.366	1.215.600	-0,508	0,0971	✓
Dudo$(4,2,1)$	327.641	14.366	1.213.799	0,552	0,3701	✓
Dudo$(4,2,2)$	13.107.101	327.680	63.109	0,0069	2,1132	✗
Dudo$(5,1,1)$	51.176	2.560	4.521.208	-0,120	0,1186	✓
Dudo$(5,1,2)$	4.915.126	163.840	151.235	-0,565	0,6197	✓
Dudo$(5,2,1)$	4.915.126	163.840	143.698	0,581	0,0122	✓
Dudo$(5,2,2)$	471.858.976	7.864.320	3.826	0,836	15,1963	✗
Dudo$(6,1,1)$	294.877	12.288	1.067.782	-0,111	0,0975	✓
Dudo$(6,1,2)$	66.060.163	1.769.472	17.702	-0,593	4,5781	✗
Dudo$(6,2,1)$	66.060.163	1.769.472	17.221	0,592	3,9594	✗

Dudo



Dudo(4,2,2)
Dudo(5,2,2)
Dudo(6,1,2)
Dudo(6,2,1)
Jugador 1
Jugador 2
Regret (max)
Iteraciones

Dominó

Dominó

- Versión para dos jugadores.

Dominó

- Versión para dos jugadores.
- M: máximo número de puntos en una cara de una pieza.

Dominó

- Versión para dos jugadores.
- M: máximo número de puntos en una cara de una pieza.
- N: número de piezas de la mano inicial para cada jugador.

Dominó

- Versión para dos jugadores.
- M: máximo número de puntos en una cara de una pieza.
- N: número de piezas de la mano inicial para cada jugador.
- Domino(M, N).

Dominó

- Versión para dos jugadores.
- M: máximo número de puntos en una cara de una pieza.
- N: número de piezas de la mano inicial para cada jugador.
- Domino(M, N).

Tabla de Resultados

Juego	N	I	Iteraciones	$u(\sigma)$	ε_σ (%)	Resuelto
Domino$(2, 2)$	7.321	102	540.186.366	2,4000	0,0000	✓
Domino$(3, 2)$	46.534.657	88.947	400.047.334	2,8767	0,0315	✓
Domino$(3, 3)$	246.760.993	107.854	72.492.951	2,1539	0,3854	✓
Domino$(3, 4)$	1.547.645.185	104.050	11.213.463	3,2034	1,4871	✗

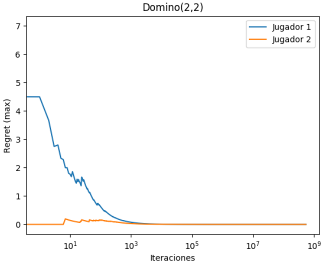
Dominó

- Versión para dos jugadores.
- M: máximo número de puntos en una cara de una pieza.
- N: número de piezas de la mano inicial para cada jugador.
- Domino(M, N).

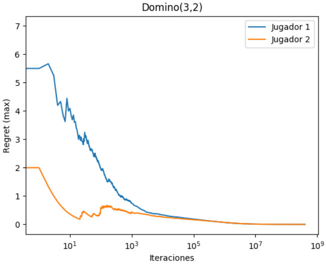
Tabla de Resultados

Juego	N	I	Iteraciones	$u(\sigma)$	ε_σ (%)	Resuelto
Domino$(2,2)$	7.321	102	540.186.366	2,4000	0,0000	✓
Domino$(3,2)$	46.534.657	88.947	400.047.334	2,8767	0,0315	✓
Domino$(3,3)$	246.760.993	107.854	72.492.951	2,1539	0,3854	✓
Domino$(3,4)$	1.547.645.185	104.050	11.213.463	3,2034	1,4871	✗

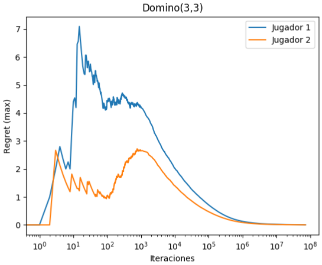
Dominó



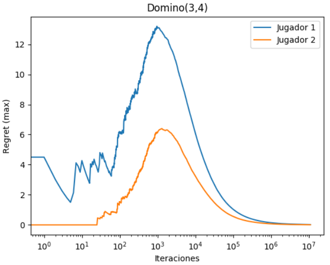
Domino(2,2)
Jugador 1
Jugador 2
Regret (max)
Iteraciones



Domino(3,2)
Jugador 1
Jugador 2
Regret (max)
Iteraciones



Domino(3,3)
Jugador 1
Jugador 2
Regret (max)
Iteraciones



Domino(3,4)
Jugador 1
Jugador 2
Regret (max)
Iteraciones

Experimentos Adicionales

Experimentos Adicionales

- Juegos no resueltos con 10 horas de entrenamiento.

Experimentos Adicionales

- Juegos no resueltos con 10 horas de entrenamiento.
- 200 horas de entrenamiento.

Experimentos Adicionales

- Juegos no resueltos con 10 horas de entrenamiento.
- 200 horas de entrenamiento.

Tabla de Resultados

Juego	N	I	Iteraciones	$u(\sigma)$	ε_σ (%)	Resuelto
Dudo$(4,2,2)$	13.107.101	327.680	2.276.259	0,00875	0,2382	✓
Dudo$(5,2,2)$	471.858.976	7.864.320	133.863	-0,00004	1,7695	✗
Dudo$(6,1,2)$	66.060.163	1.769.472	543.485	-0,597	0,5102	✓
Dudo$(6,2,1)$	66.060.163	1.769.472	513.786	0,597	0,6727	✓
Domino$(3,4)$	1.547.645.185	104.050	365.484.932	3.2027	0,1812	✓

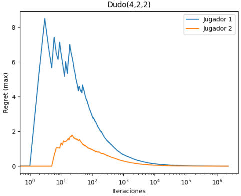
Experimentos Adicionales

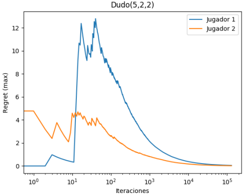
- Juegos no resueltos con 10 horas de entrenamiento.
- 200 horas de entrenamiento.

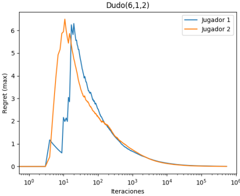
Tabla de Resultados

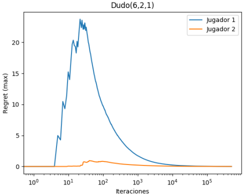
Juego	N	I	Iteraciones	$u(\sigma)$	ε_σ (%)	Resuelto
Dudo$(4,2,2)$	13.107.101	327.680	2.276.259	0,00875	0,2382	✓
Dudo$(5,2,2)$	471.858.976	7.864.320	133.863	-0,00004	1,7695	✗
Dudo$(6,1,2)$	66.060.163	1.769.472	543.485	-0,597	0,5102	✓
Dudo$(6,2,1)$	66.060.163	1.769.472	513.786	0,597	0,6727	✓
Domino$(3,4)$	1.547.645.185	104.050	365.484.932	3.2027	0,1812	✓

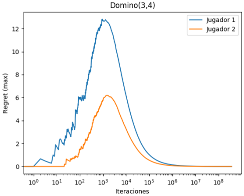
Experimentos Adicionales











CFR: Conclusiones

1. La forma extesiva es un modelo adecuado para representar juegos secuenciales con información incompleta y no determinismo.
2. El algoritmo CFR permite encontrar un equilibrio de Nash en juegos en forma extensiva de dos jugadores de suma cero.
3. Se resolvieron diversos juegos captados por el modelos:
   - OCP: el segundo tiene ventaja sobre el primer jugador: $u \in (-0.7, -0.5)$.
   - Dudo: el jugador con mayor número de dados tiene ventaja.
   - Domino: el primer jugador tiene ventaja sobre el segundo jugador: $u \in (2, 3.5)$.

Conclusiones y Recomendaciones

Conclusiones y Recomendaciones

Conclusiones

1. Los modelos utilizados son adecuados para los juegos planteados.

Conclusiones y Recomendaciones

Conclusiones

1. Los modelos utilizados son adecuados para los juegos planteados.
2. El equilibrio de Nash es un concepto de solución satisfactorio en juegos de dos jugadores de suma cero. No lo es cuando el juego no es de suma cero o tiene más de dos jugadores.

Conclusiones y Recomendaciones

Conclusiones

1. Los modelos utilizados son adecuados para los juegos planteados.
2. El equilibrio de Nash es un concepto de solución satisfactorio en juegos de dos jugadores de suma cero. No lo es cuando el juego no es de suma cero o tiene más de dos jugadores.
3. Se utilizó la explotabilidad como métrica para medir la distancia entre la estrategia obtenida y un equilibrio de Nash.

Conclusiones y Recomendaciones

Conclusiones

1. Los modelos utilizados son adecuados para los juegos planteados.
2. El equilibrio de Nash es un concepto de solución satisfactorio en juegos de dos jugadores de suma cero. No lo es cuando el juego no es de suma cero o tiene más de dos jugadores.
3. Se utilizó la explotabilidad como métrica para medir la distancia entre la estrategia obtenida y un equilibrio de Nash.
4. Resolución de juegos: se encontraron aproximaciones con una explotabilidad no mayor que el 1% de la mínima ganancia positiva posible.

Conclusiones y Recomendaciones

Conclusiones

1. Los modelos utilizados son adecuados para los juegos planteados.
2. El equilibrio de Nash es un concepto de solución satisfactorio en juegos de dos jugadores de suma cero. No lo es cuando el juego no es de suma cero o tiene más de dos jugadores.
3. Se utilizó la explotabilidad como métrica para medir la distancia entre la estrategia obtenida y un equilibrio de Nash.
4. Resolución de juegos: se encontraron aproximaciones con una explotabilidad no mayor que el 1% de la mínima ganancia positiva posible.

Recomendaciones

1. Resolver instancias mayores del juego de dominó para 2 personas considerando abstracciones.

Conclusiones y Recomendaciones

Conclusiones

1. Los modelos utilizados son adecuados para los juegos planteados.
2. El equilibrio de Nash es un concepto de solución satisfactorio en juegos de dos jugadores de suma cero. No lo es cuando el juego no es de suma cero o tiene más de dos jugadores.
3. Se utilizó la explotabilidad como métrica para medir la distancia entre la estrategia obtenida y un equilibrio de Nash.
4. Resolución de juegos: se encontraron aproximaciones con una explotabilidad no mayor que el 1% de la mínima ganancia positiva posible.

Recomendaciones

1. Resolver instancias mayores del juego de dominó para 2 personas considerando abstracciones.
2. Experimentos sobre el juego para 4 personas considerando cada pareja como un único jugador.

Demo

Gracias por su atención

¿Preguntas?