# Toma de Decisiones Multiagente

Artificial Intelligence: A Modern Approach, 4th US ed.

by Stuart Russell and Peter Norvig

Diego Esteban Quintero Rey

Modelos Estocásticos y Simulación en Computación y Comunicaciones
Jorge Eduardo Ortiz Triviño (Profesor Asociado)
Universidad Nacional de Colombia

# Toma de decisiones multiagente

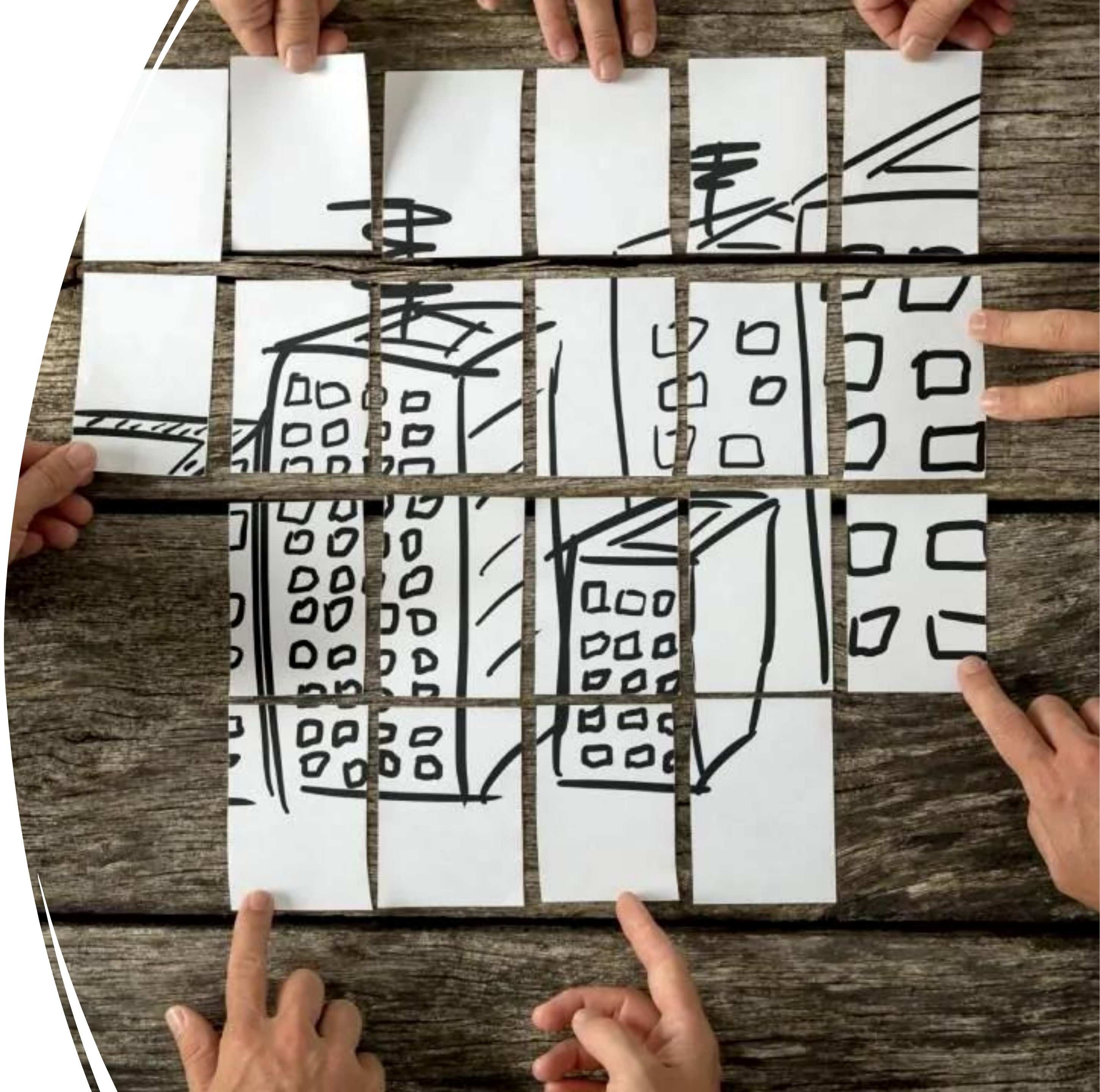
En este capítulo se examina qué se hace cuándo varios agentes habitan un ambiente.

# Propiedades de los ambientes multiagente

Un ambiente que contiene varios actores se denomina como: **sistema multiagente**.

Los agentes que se encuentran en este sistema se enfrentan al **Problema de Planeación Multiagente**.

Resolver este problema requiere de técnicas que son diferentes dependiendo de las relaciones que tengan los agentes entre sí.



# Un tomador de decisiones

El tomador de decisiones elabora planes y les dice a los demás agentes qué deben hacer.

Se asume que los agentes van a cumplir las órdenes.

Esto se llama: **Suposición del agente benevolente**.

# Sincronización

los agentes deben garantizar sincronización que puede ser de tres tipos:

- *Simultanea: como* cantar un dueto,
- *En diferentes tiempos:* para acciones mutuamente excluyentes, como por ejemplo cargar una batería cuando solo hay un enchufe.
- *Secuencial:* Cuando una tarea es precondición de otra, como, por ejemplo, lavar los platos y luego secarlos.

# Multiefectores

Los agentes tomadores de decisiones también podrían tener múltiples efectores que pueden operar concurrentemente. Los efectores son como las extremidades o las salidas de los agentes. Por ejemplo, un humano, puede hablar y caminar al mismo tiempo.

A esto se le llama **Planeación de Multiefectores**.

# Planeación de multicuerpos

Estos efectores podrían ser también unidades que se despegan del agente, como por ejemplo una flota de robots. A esto se le llama **planeación de multicuerpos**.

La planeación de multicuerpos puede ser vista como un problema de un solo agente, si toda la información que es recolectada por los cuerpos puede ser recolectada en una base central o en cada cuerpo, de manera que en todo momento se tenga un estimado del estado del mundo. Los múltiples cuerpos pueden entonces pensarse como un solo cuerpo. En ocasiones no se puede establecer comunicación entre las partes en todo momento.

A esto se le llama **planeación descentralizada**.

# Múltiples tomadores de decisiones

Cada agente tiene sus preferencias y elige ejecutar su propio plan. Los agentes en este tipo de ambientes se llaman contrapartes.

# Meta común

Por ejemplo, los empleados de una compañía. En este caso, el problema al que se enfrentan los agentes es el de la **coordinación**, es decir, deben asegurarse de que todos se están moviendo en la misma dirección.

# Propias preferencias

Los agentes tienen sus propias preferencias, que cada uno intenta perseguir lo mejor que puedan. Estas preferencias pueden ser diametralmente opuestas como ocurre en los **juegos de suma cero** (un ejemplo es el ajedrez). Pero la mayoría de los ambientes multiagente son más complicados.

# Juego suma cero

Un juego de suma cero es aquel en el que las recompensas de los jugadores son iguales y opuestas. Es decir, lo que uno gana, el otro lo pierde y viceversa.

# Teoría de juegos

La teoría de juegos es la teoría que estudia la toma de decisiones estratégica. Lo estratégico aquí, es que los jugadores toman en cuenta cómo podrían actuar los otros jugadores. Eso es precisamente lo que distingue la teoría de juegos de la teoría de la decisión.

# Aplicaciones de la teoría de juegos

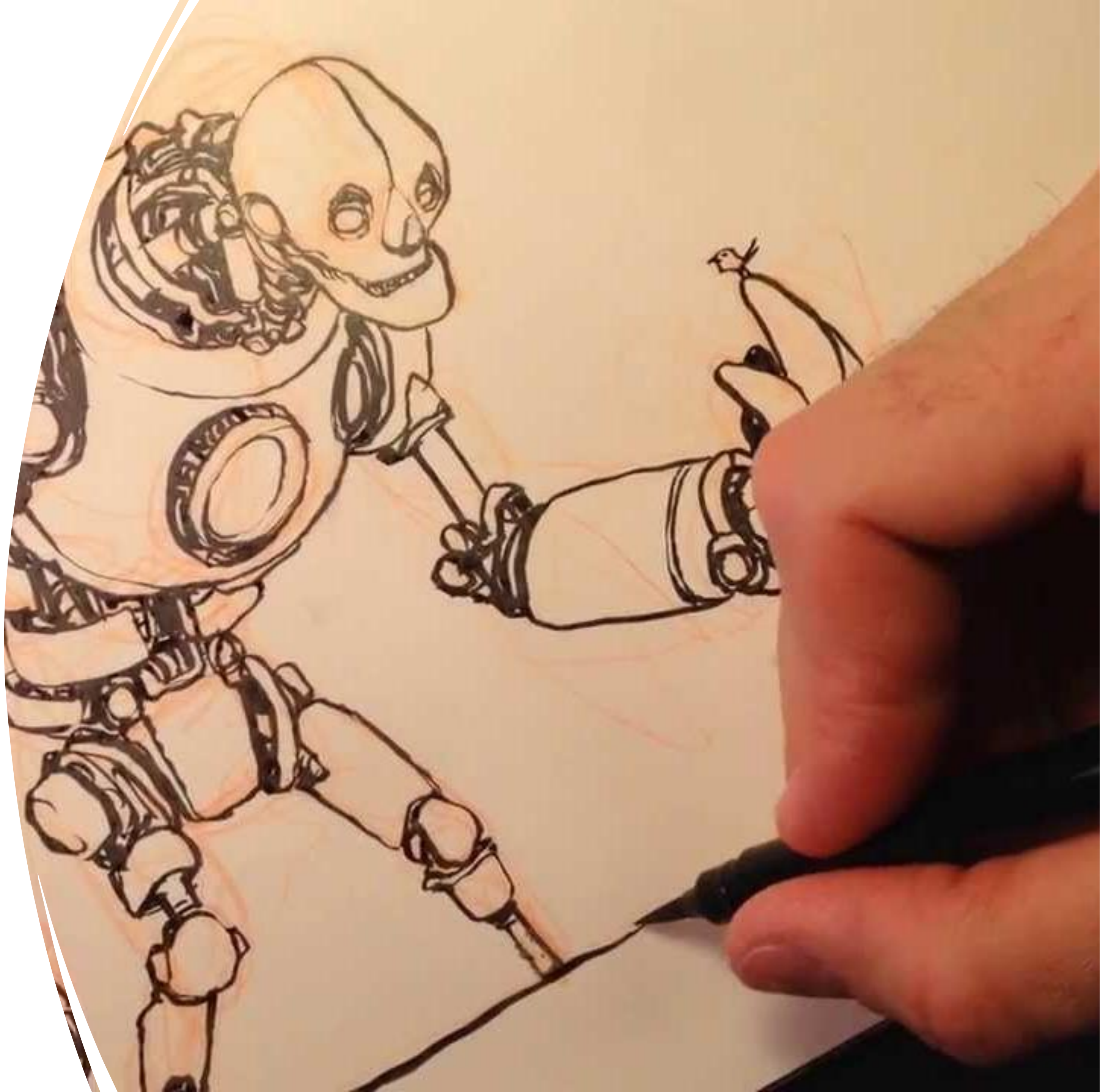
La subasta de derechos de perforación petrolera y derechos de espectro de frecuencias inalámbricas, procedimientos de quiebra, desarrollo de productos y decisiones sobre fijación de precios, y defensa nacional

# Ambientes mixtos

Algunos ambientes pueden ser mixtos, por ejemplo, en una compañía de envío de paquetes, la planeación de las rutas de sus camiones y aviones puede hacerse de manera centralizada, y luego los conductores y pilotos seguirían las instrucciones. Pero, luego los conductores podrían tomar sus propias decisiones dependiendo de las condiciones del tráfico o del clima. Las metas de los conductores y la compañía se alinean con el pago de incentivos (salarios y bonos).

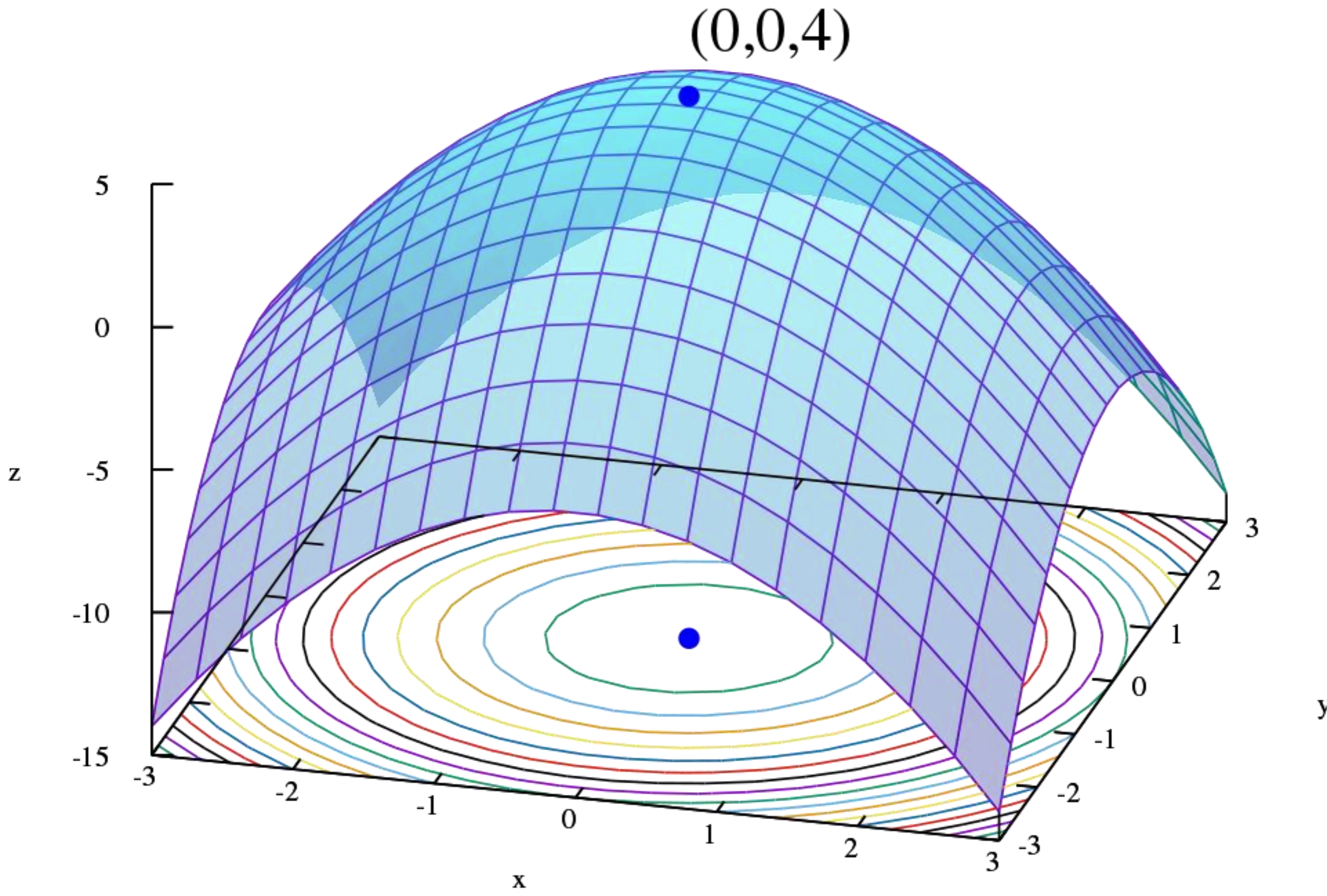
# ¿Cómo se usa la teoría de juegos en la IA?

- Diseño de agentes
- Diseño de mecanismos



# Diseño de agentes

En el diseño de agentes, la teoría de juegos puede usarse para que el agente analice sus posibles decisiones y calcule la utilidad esperada para cada una, asumiendo que los otros agentes actúan racionalmente. Actuar racionalmente en el contexto de la teoría de juegos es tomar las decisiones que maximizan la utilidad esperada.



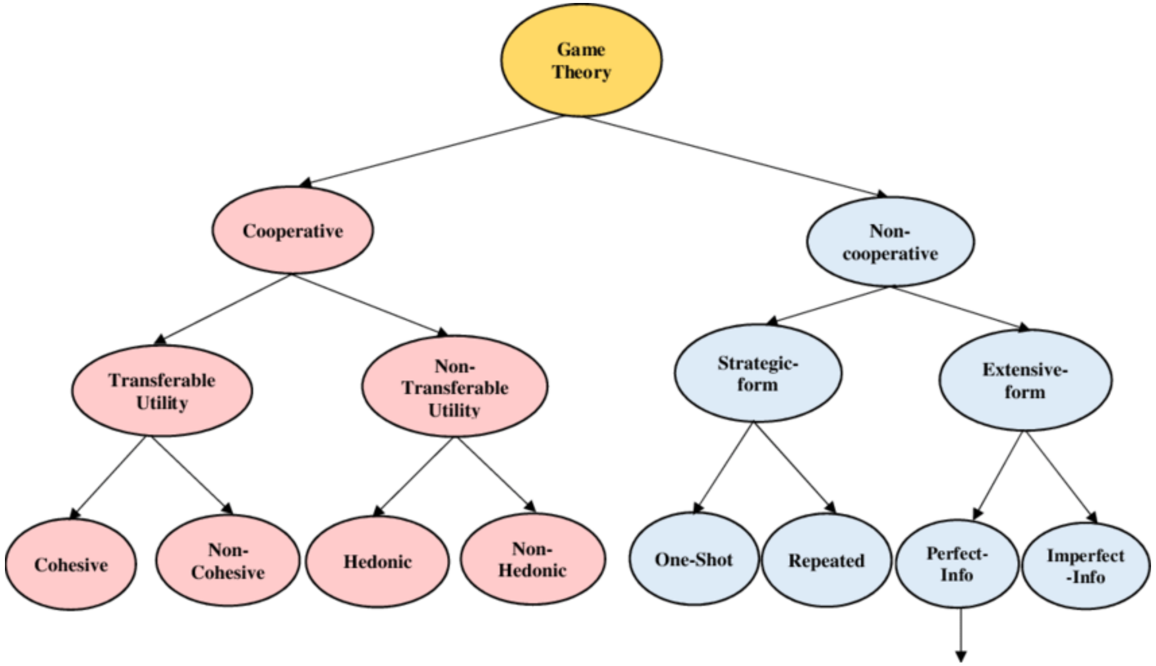
# Diseño de mecanismos

En el diseño de mecanismos, la teoría de juegos puede usarse para definir las reglas del ambiente de manera que se maximice el bienestar colectivo cuando estos actúan racionalmente y maximizan su propia utilidad.

Un ejemplo es el diseño de protocolos para una colección de enrutadores de internet, de manera que cada uno tenga un incentivo para actuar de manera que el rendimiento global es maximizado.

# Modelos de la teoría de juegos

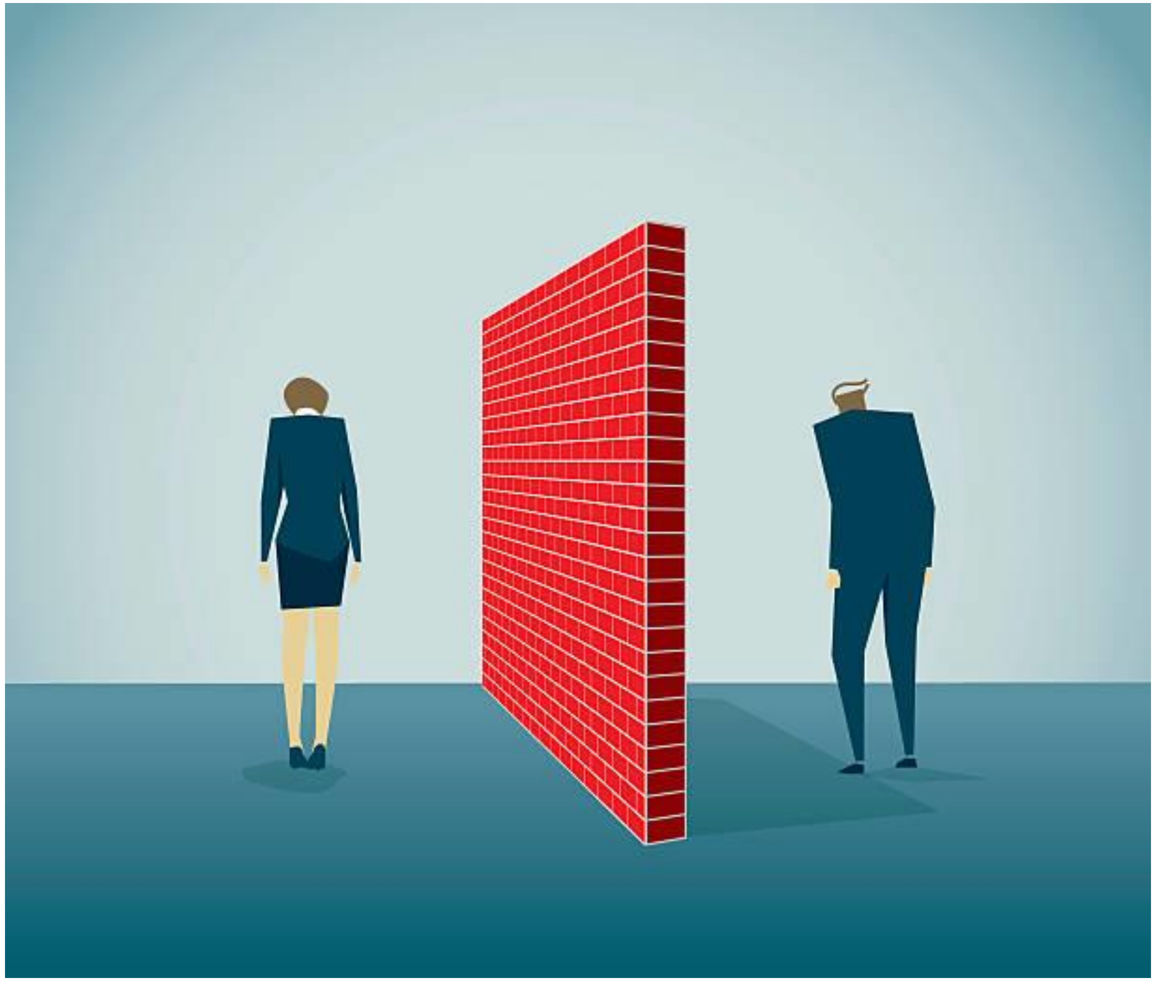
- No cooperativos
- Cooperativos




Tomado de: [Game-Theoretic Foundations for Forming Trusted Coalitions of Multi-Cloud Services in the Presence of Active and Passive Attacks - Spectrum: Concordia University Research Repository]

# Juegos no cooperativos

En los no cooperativos, los jugadores no pueden hacer acuerdos. Esto no necesariamente significa que el juego sea netamente competitivo. Lo único es que en estos juegos no hay un acuerdo central que vincule a los agentes y que garantice que cooperen. Pero puede que lleguen a cooperar entre sí, si esto les interesa.

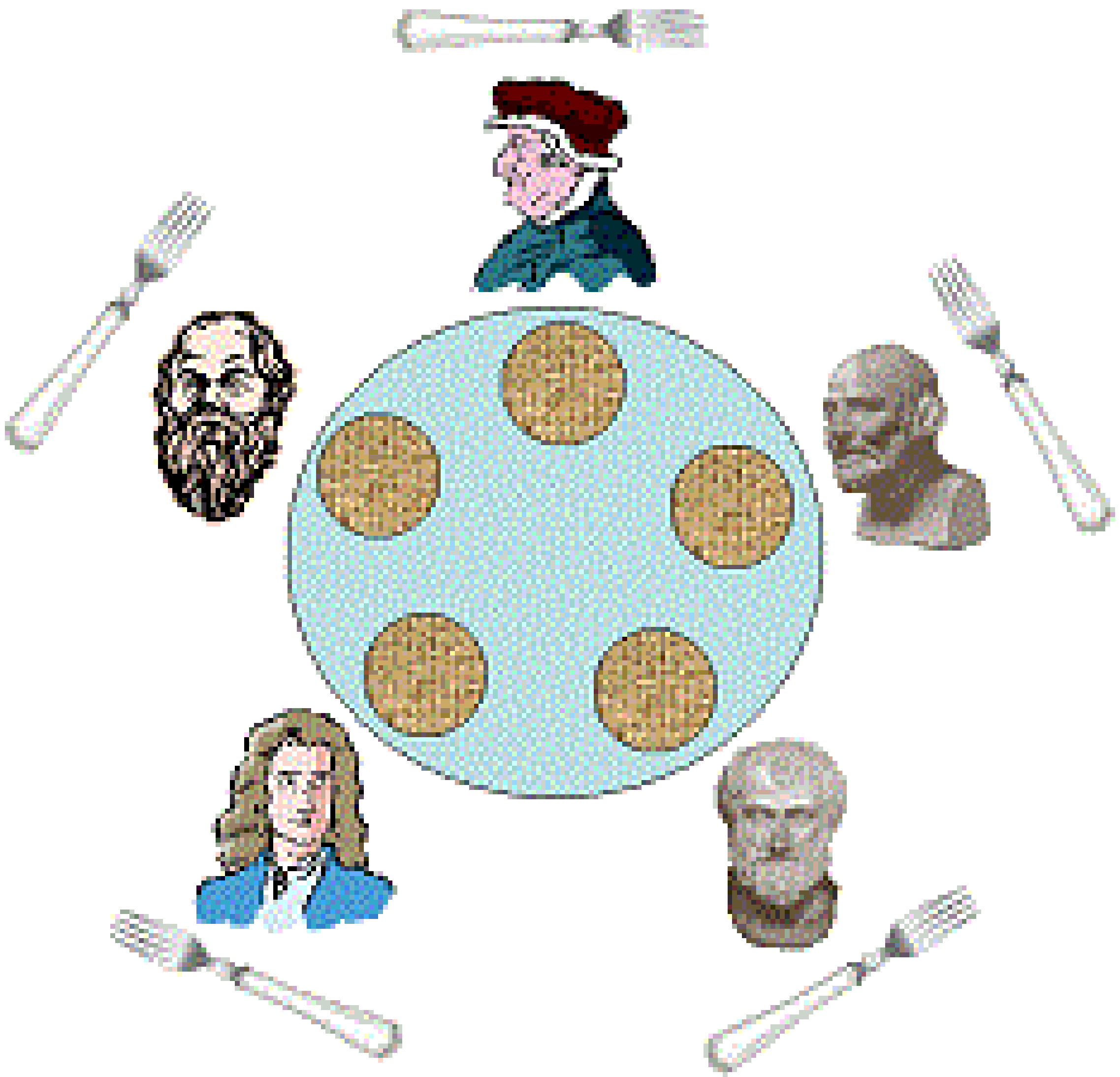


# Juegos cooperativos

En los juegos cooperativos, los jugadores pueden hacer acuerdos, permitiendo cooperación robusta. En el mundo real, son los contratos legales y las normas sociales lo que ayuda a establecer estos acuerdos.

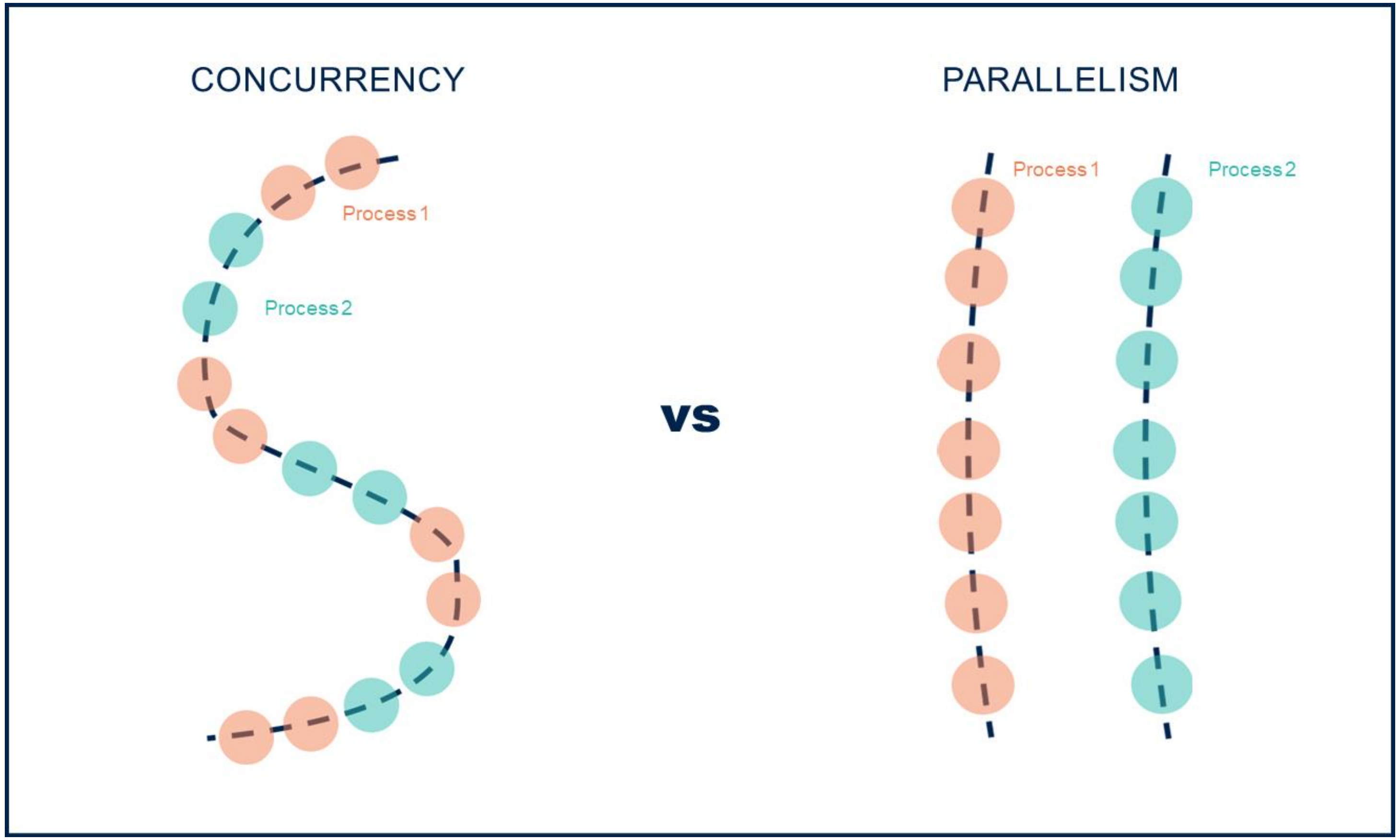
# Concurrencia

El objetivo de la concurrencia es que los planes de los agentes se ejecuten simultáneamente. Pero también, los agentes deben tener en cuenta la forma en la que sus propias acciones interactúan con las acciones de otros agentes. Los agentes deben tener en cuenta si las acciones de otros afectarán las precondiciones de sus propias acciones, por ejemplo, o si los recursos que necesita son compartidos, o pueden ser agotados por otros agentes. También debe tener en cuenta si las acciones son mutuamente excluyentes, o si sus acciones pueden facilitar las acciones de los otros.



# Modelos de concurrencia

- Intercalada
- Verdadera
- Perfecta

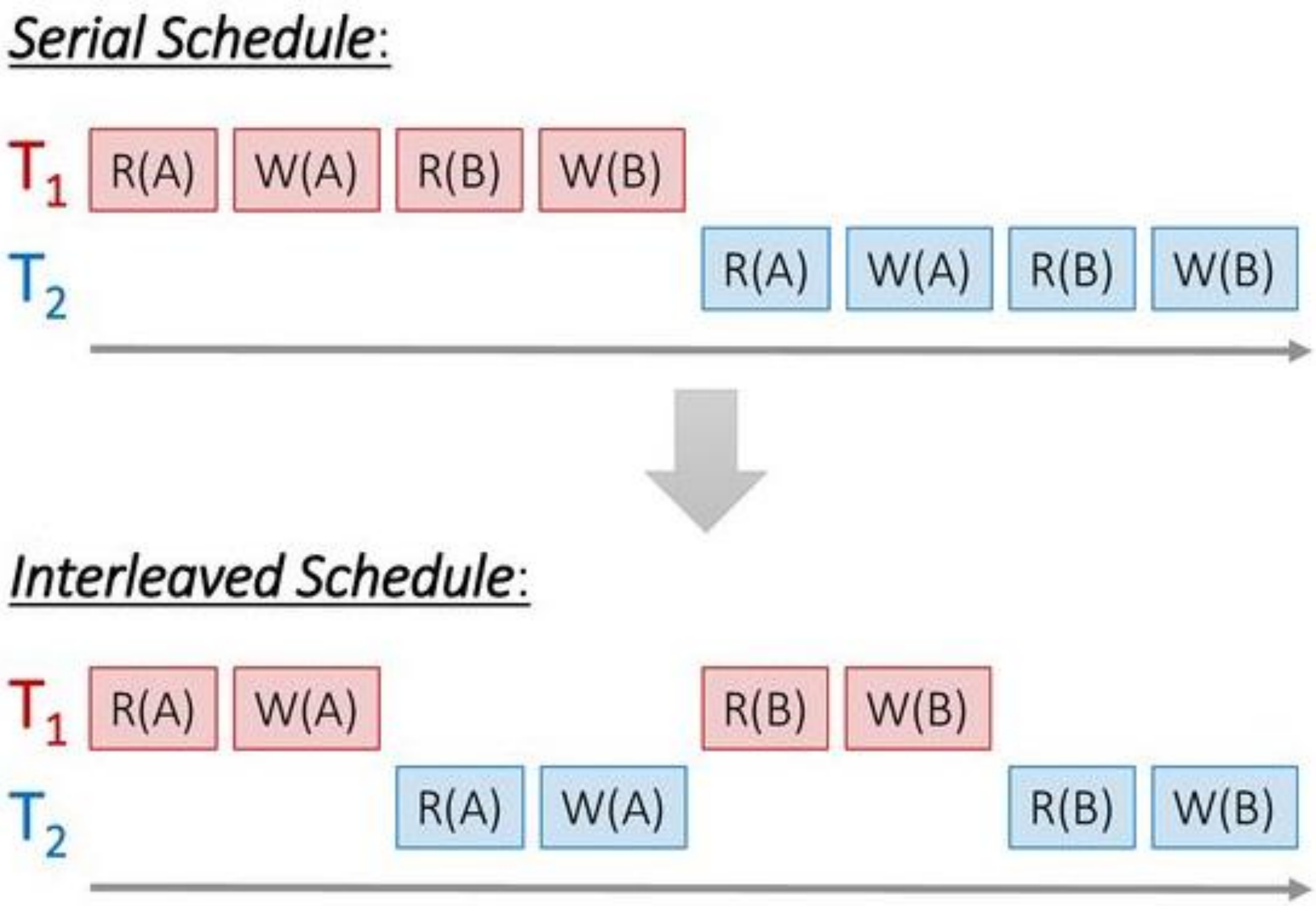



Tomado de: [Concurrency in Flutter. Let us understand concurrency in… | by DLT Labs | DLT Labs | Medium]

# Concurrencia intercalada

En la ejecución intercalada, digamos que tenemos dos agentes $A$, $B$ cada uno con un plan $[a_1, a_2]$ y $[b_1, b_2]$. Podemos organizar estos planes en planes de cuatro tareas en los que se garantice que el orden de las tareas de los planes individuales se conserva, como por ejemplo $[a_1, b_1, b_2, a_2]$.Un plan determinado es correcto si, no importa cuál sea la intercalación, se cumple la meta.

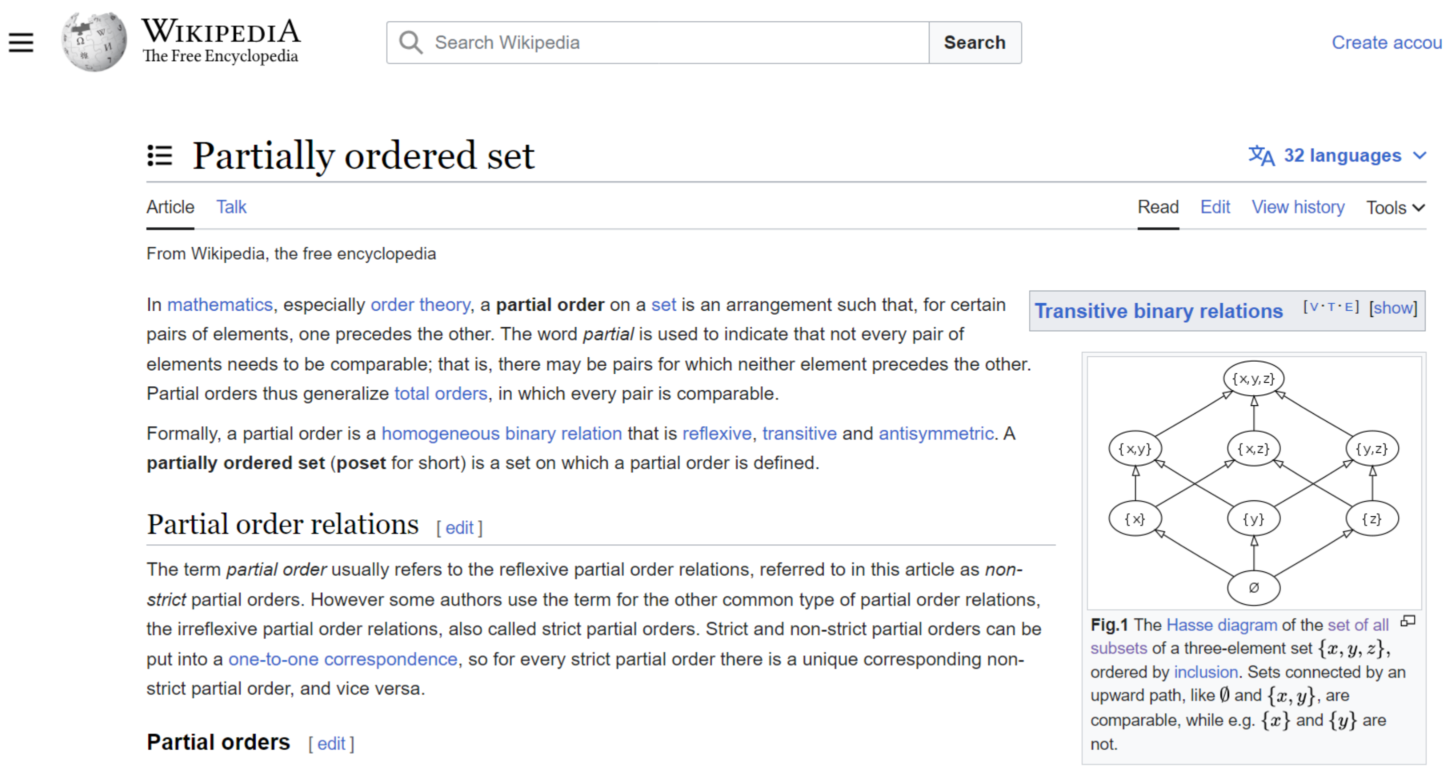
Este número de intercalaciones crece exponencialmente, y además no considera el caso en el que dos tareas deben realizarse simultáneamente.




Tomado de: [Lecture 21: Concurrency & Locking - ppt download (slideplayer.com)]

# Verdadera concurrencia

La verdadera concurrencia, es similar al modelo anterior, pero en este caso las intercalaciones son ordenamientos parciales, lo que quiere decir, que no nos importan definir si una tarea de $A$ ocurre antes de una tarea de $B$ y viceversa y, de hecho, podrían ocurrir simultáneamente.


WIKIPEDIA
The Free Encyclopedia

Search Wikipedia | Search

Create accou

## Partially ordered set

32 languages

Article Talk | Read Edit View history Tools

From Wikipedia, the free encyclopedia

In mathematics, especially order theory, a **partial order** on a set is an arrangement such that, for certain pairs of elements, one precedes the other. The word *partial* is used to indicate that not every pair of elements needs to be comparable; that is, there may be pairs for which neither element precedes the other. Partial orders thus generalize total orders, in which every pair is comparable.

Formally, a partial order is a homogeneous binary relation that is reflexive, transitive and antisymmetric. A **partially ordered set** (**poset** for short) is a set on which a partial order is defined.

Transitive binary relations [V · T · E] [show]



**Fig.1** The Hasse diagram of the set of all subsets of a three-element set $\{x, y, z\}$, ordered by inclusion. Sets connected by an upward path, like $\emptyset$ and $\{x, y\}$, are comparable, while e.g. $\{x\}$ and $\{y\}$ are not.

### Partial order relations [ edit ]

The term *partial order* usually refers to the reflexive partial order relations, referred to in this article as *non-strict* partial orders. However some authors use the term for the other common type of partial order relations, the irreflexive partial order relations, also called strict partial orders. Strict and non-strict partial orders can be put into a one-to-one correspondence, so for every strict partial order there is a unique corresponding non-strict partial order, and vice versa.

**Partial orders** [ edit ]

# Sincronización perfecta

Todos los agentes tienen acceso a reloj global, todas las tares tienen la misma duración y se ejecutan paso a paso de manera simultánea. Cuando un agente debe esperar que se complete otra tarea, ejecuta una tarea llamada **no-op**, que es prácticamente una espera.

# Tenis dobles, jugadores $A$, $B$

- Objetivo: golpear una pelota
- Acciones: golpear, desplazarse

# Tenis dobles, jugadores *A*, *B*

Objetivo: golpear una pelota

Acciones: golpear, desplazarse

$$
\begin{aligned}
&Action(Hit(actor, Ball),\\
&\quad \text{CONCURRENT:} \forall b \;\; b \neq actor \Rightarrow \neg Hit(b, Ball)\\
&\quad \text{PRECOND:} Approaching(Ball, loc) \wedge At(actor, loc)\\
&\quad \text{EFFECT:} Returned(Ball))\,.
\end{aligned}
$$

$$
\begin{aligned}
&Action(Go(actor, to),\\
&\quad \text{PRECOND:} At(actor, loc) \wedge to \neq loc,\\
&\quad \text{EFFECT:} At(actor, to) \wedge \neg At(actor, loc))
\end{aligned}
$$

$$
\begin{aligned}
\text{PLAN 1:} \qquad & A: \; [Go(A, RightBaseline), Hit(A, Ball)]\\
& B: \; [NoOp(B), NoOp(B)]\,.
\end{aligned}
$$

$$
\begin{aligned}
\text{PLAN 2:} \qquad & A: \; [Go(A, LeftNet), NoOp(A)]\\
& B: \; [Go(B, RightBaseline), Hit(B, Ball)]\,.
\end{aligned}
$$

Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Restricciones de concurrencia

Las restricciones de concurrencia garantizan que, por ejemplo, dos agentes no hagan una acción mutuamente excluyente al mismo tiempo, por ejemplo, golpear la pelota. También puede usarse para garantizar que se tiene el número de agentes adecuado para realizar una acción, por ejemplo, levantar un objeto pesado.

# Planes

$$\text{Plan 1:} \quad \begin{aligned} A &: [Go(A, RightBaseline), Hit(A, Ball)] \\ B &: [NoOp(B), NoOp(B)]\,. \end{aligned}$$

$$\text{Plan 2:} \quad \begin{aligned} A &: [Go(A, LeftNet), NoOp(A)] \\ B &: [Go(B, RightBaseline), Hit(B, Ball)]\,. \end{aligned}$$

Si ambos escogen 1 o ambos escogen 2, el objetivo de golpear la pelota se cumple, pero si $A$ escoge 2 y $B$ 1, la nadie golpea la pelota. Al contrario, los dos intentarán golpear la pelota, pero los dos van a fallar debido a la restricción de concurrencia.

# Convenciones

Se necesita de convenciones, los cuales pueden ser normas sociales o **comunicación** para que logren coordinarse y elegir los planes que realmente les ayude cumplir con el objetivo. Otra posibilidad es que un agente empiece a ejecutar un plan, y los otros reconozcan cuál es el plan en los primeros pasos. Esto solo es posible si son reconocibles sin ambigüedad en esos primeros pasos. A esto se le llama **reconocimiento de planes**.

# Teoría de juegos no cooperativa

# Juegos en forma normal

Estos juegos, solo tienen un movimiento por jugador. Es decir, un turno. Cada jugador hace su acción sin tener conocimiento de qué acción tomaron los otros jugadores. Como si todos hicieran el movimiento simultáneamente.

# Componentes de un juego

*Jugadores*: los agentes que van a tomar decisiones. Un juego tiene n jugadores y estos se denotan con letras mayúsculas $A, B, C, \ldots$

*Acciones*: Cada jugador tiene un conjunto de acciones disponible. Una acción se denota con un nombre en minúsculas, por ejemplo: testificar, uno, etc.

*Función de recompensa*: función que indica la utilidad que recibe cada jugador para toda combinación de acciones que se puede dar en el juego.

# Matriz de recompensa

Cuando el juego tiene 2 jugadores, la función de recompensa puede representarse en una matriz de recompensa para cada jugador, pero usualmente las dos matrices se combinan en una sola así:

| | *O: one* | *O: two* |
|---|---|---|
| *E: one* | $E=+2, O=-2$ | $E=-3, O=+3$ |
| *E: two* | $E=-3, O=+3$ | $E=+4, O=-4$ |

Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]


# Juego de Morra de dos dedos

En este juego hay dos jugadores: el jugador par $E$ y el jugador impar $O$.

En cada turno, los jugadores deben hacer sus acciones simultáneamente. Las acciones disponibles son mostrar 1 dedo o mostrar 2. Dependiendo de si la suma es par o impar, se le da una utilidad al jugador correspondiente por una cantidad igual al número de dedos total. El jugador perdedor recibe la misma utilidad, pero negativa. Por eso es un juego de suma cero.

# Estrategias

Una estrategia es una o un conjunto de acciones a tomar por un jugador. Una **estrategia pura**, es una sola acción.

Una estrategia puede ser **mixta**, es decir **aleatoria**, en la que el jugador escoge acciones de acuerdo con una distribución de probabilidad.

# Estrategias mixtas o aleatorias

Una estrategia mixta se denota como:

$$[p: a; (1 - p): b]$$

y significa que el jugador toma la acción a con probabilidad $p$ y la acción $b$ con probabilidad $1 - p$. Por ejemplo:

$$[0.5: one; 0.5: two].$$

# Perfil de estrategia

Es la asignación de una estrategia a cada jugador, es decir, una combinación de estrategias del juego. Por ejemplo, en el juego de Morra los perfiles de estrategia disponibles son:

$$one, one$$
$$one, two$$
$$two, one$$
$$two, two$$

# Dilema del prisionero

En este juego, dos personas son arrestadas. Si ambas confiesan, cada una recibe una sentencia de 5 años de prisión. Si solo una confiesa, esa persona queda libre mientras que la otra recibe una sentencia de 10 años de prisión. Si ambas se niegan a confesar, cada una recibe solo un año de prisión.

# Dilema del prisionero

| | *Ali:testify* | *Ali:refuse* |
|---|---|---|
| *Bo:testify* | $A = -5, B = -5$ | $A = -10, B = 0$ |
| *Bo:refuse* | $A = 0, B = -10$ | $A = -1, B = -1$ |

Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

El razonamiento de $A$ es el siguiente:

- Si $B$ testifica entonces es mejor para mi testificar.
- Si $B$ se rehúsa, es mejor para mi testificar.

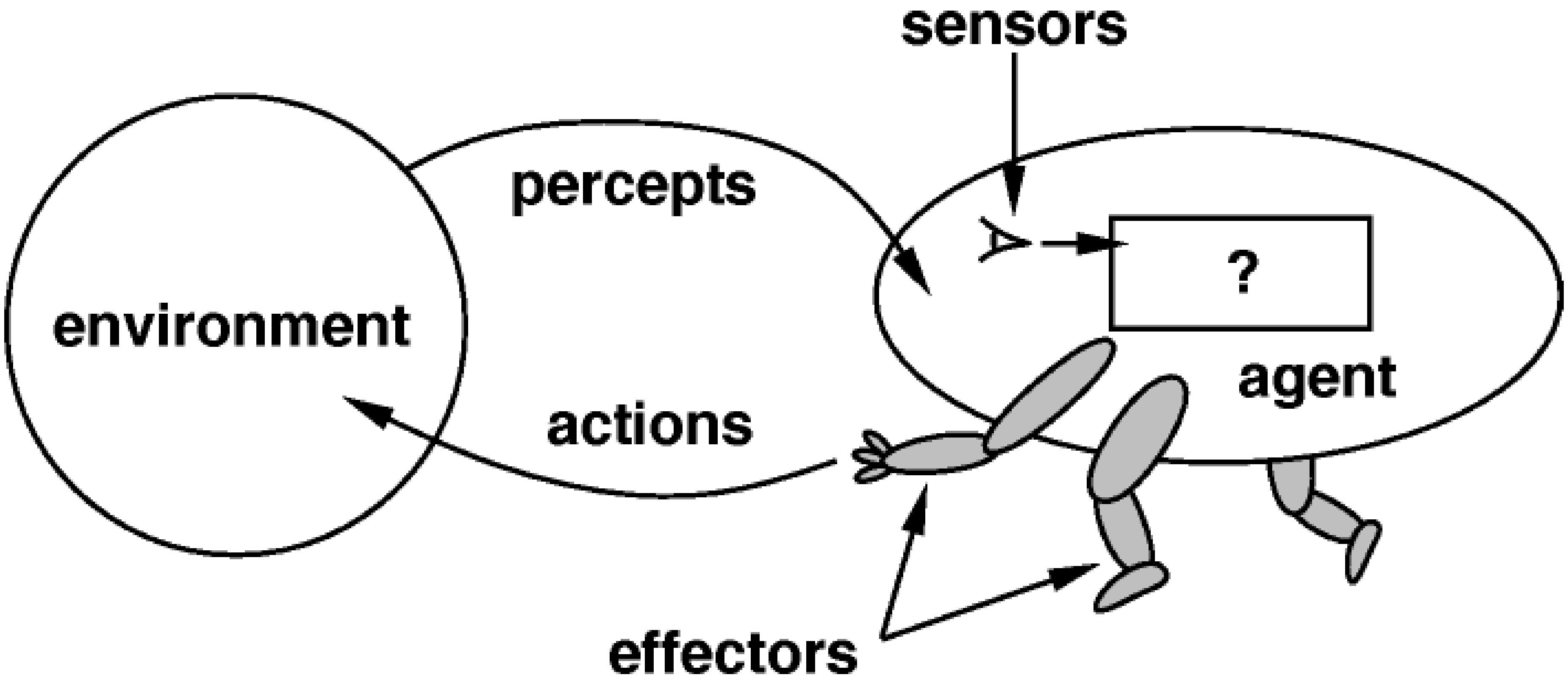
# Estrategia dominante

En este caso **testificar** es una **estrategia dominante**. Una estrategia $s$ para el jugador $P$ domina fuertemente la estrategia $s'$ si el resultado del juego de $s$ es mejor para $P$ que el resultado de $s'$. Domina débilmente si es mejor en al menos un perfil de estrategia, y en los otros el resultado es igual.

Una estrategia dominante es aquella que domina a todas las otras.

# Agente racional y estrategias dominantes

La teoría de juegos asume que un jugador es racional si siempre escoge una estrategia dominante y evita una estrategia dominada.




Tomado de: [Section 02 (massey.ac.nz)]

# Equilibrio de estrategia dominante

Si todos los jugadores escogen una estrategia dominante, entonces el resultado del juego es un equilibrio de estrategia dominante. Se llama equilibrio porque ningún jugador tiene un incentivo para cambiar de estrategia.

Recordemos que ninguno de los jugadores conoce la acción del otro.

PRISONER 2

| | Confess | Lie |
|---|---|---|
| onfess | -8 , -8 | 0 , -10 |
| Lie | -10 , 0 | -1 , -1 |

Tomado de: [D.5 Dominant strategies and Nash equilibrium | Game Theory - Microeconomics - YouTube]

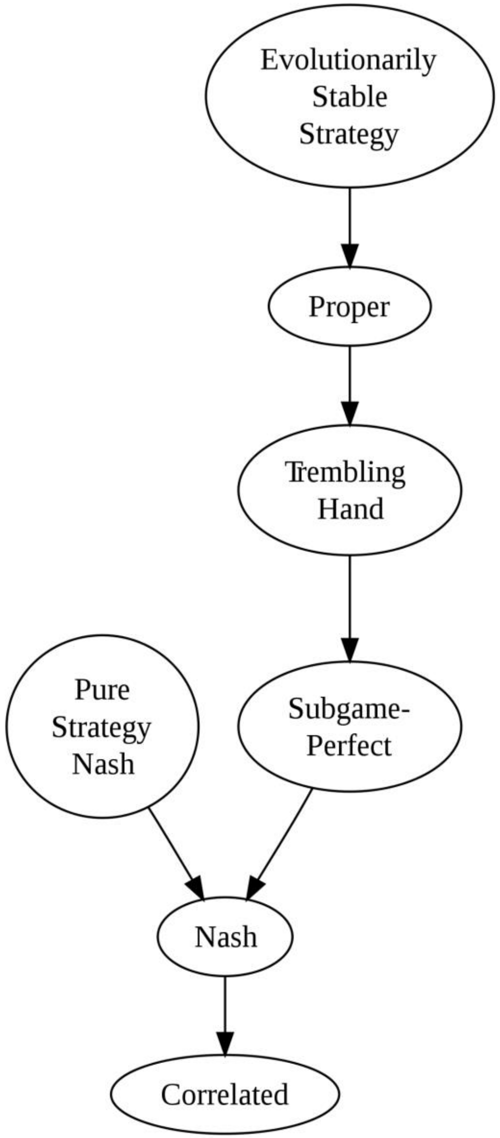
# Equilibrio de estrategia dominante

No todos los juegos tienen ese equilibrio ni estrategias dominantes. Es raro ver que una sola estrategia es la mejor respuesta a todas las posibles estrategias de los oponentes. Es decir, a todas las combinaciones o perfiles de estrategia.

# Equilibrio de Nash

Un equilibrio de Nash es un perfil de estrategia en el cuál ningún jugador podría cambiar su acción y recibir una recompensa mayor asumiendo que los otros jugadores conservan su decisión. Todo equilibrio de estrategia dominante es un equilibrio de Nash.

John Nash




Tomado de: [Solution concept - Wikipedia]

# Conceptos de solución

Los conceptos de solución que se mencionan en este capítulo son:

- Equilibrio de estrategia dominante
- Equilibrio de Nash
- Equilibrio de Nash perfecto de subjuegos (ENPS)
- Equilibrio Bayes-Nash
- Equilibrio Maximin

# Varios equilibrios de Nash

Un juego podría tener varios equilibrios de Nash. Este es un ejemplo de un juego que tiene dos equilibrios de Nash y que por tanto no tiene equilibrio de estrategia dominante.

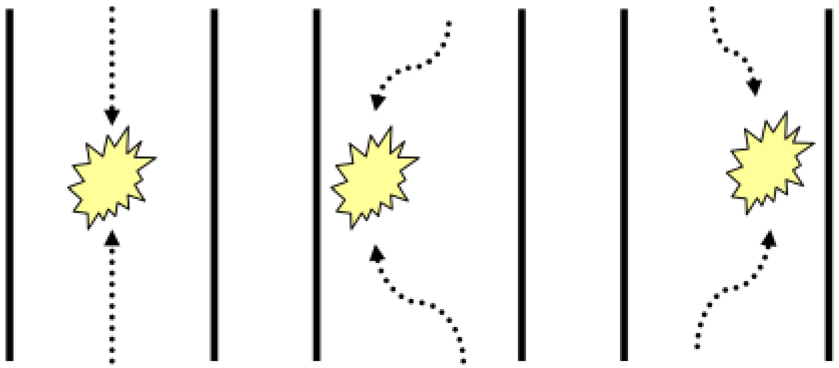
| | *Ali:l* | *Ali:r* |
|---|---|---|
| *Bo:t* | $A = 10, B = 10$ | $A = 0, B = 0$ |
| *Bo:b* | $A = 0, B = 0$ | $A = 1, B = 1$ |

Tomado de: Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)

# Punto focal

Un punto focal sería el equilibrio de Nash que más utilidad le dé a los jugadores. Se busca que los jugadores se coordinen de cierta manera sin tener conocimiento global del juego (ya que no es cooperativo) y lleguen a ese punto focal.

| | $Ali{:}l$ | $Ali{:}r$ |
|---|---|---|
| $Bo{:}t$ | $A=10, B=10$ | $A=0, B=0$ |
| $Bo{:}b$ | $A=0, B=0$ | $A=1, B=1$ |



Tomado de: Focal Points (or Schelling Points): How We Naturally Organize in Games of Coordination – Mind Your Decisions

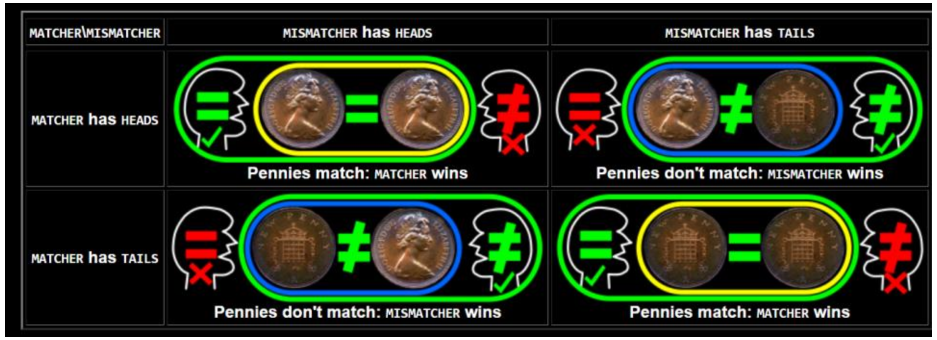
# Juego de los centavos

| | $Ali{:}heads$ | $Ali{:}tails$ |
|---|---|---|
| $Bo{:}heads$ | $A=1, B=-1$ | $A=-1, B=1$ |
| $Bo{:}tails$ | $A=-1, B=1$ | $A=1, B=-1$ |

Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

Es un ejemplo de juego que no tiene equilibrio de Nash con estrategias puras (determinísticas), sino aleatorias.

De hecho, Nash demostró que todo juego tiene al menos un equilibrio de Nash si se utilizan estrategias mixtas o aleatorias.




Tomado de: [Matching Pennies - The Matching Pennies Game]

# Ente benevolente y omnisciente

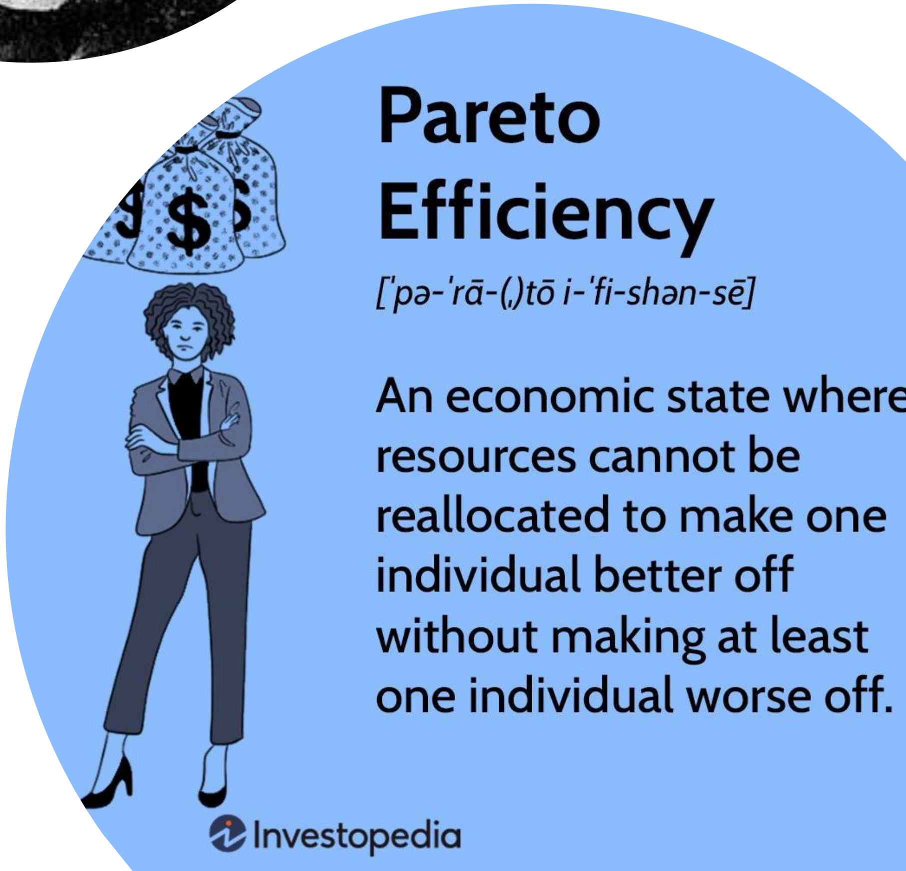
La teoría de juegos puede tener múltiples perspectivas. La más común es en la que los jugadores tratan de obtener los mejores resultados para sí mismos. Pero otra perspectiva de un ente benevolente y omnisciente que observa el juego. Como este ente es benevolente, quiere que se logre el mejor resultado promedio. Es decir, el resultado que es mejor para la sociedad como un todo.

# Bienestar social

El primer criterio es que se debe evitar desperdiciar utilidad, lo cual es cubierto por el concepto de **optimalidad de Pareto**. Un resultado es óptimo de Pareto si no hay otro resultado que beneficie a un jugador sin perjudicar a otro.

Cuando se escoge un resultado que no es óptimo de Pareto se desperdicia utilidad porque se está dejando de dar más utilidad a un jugador sin perjudicar a otros.

Vilfredo Pareto

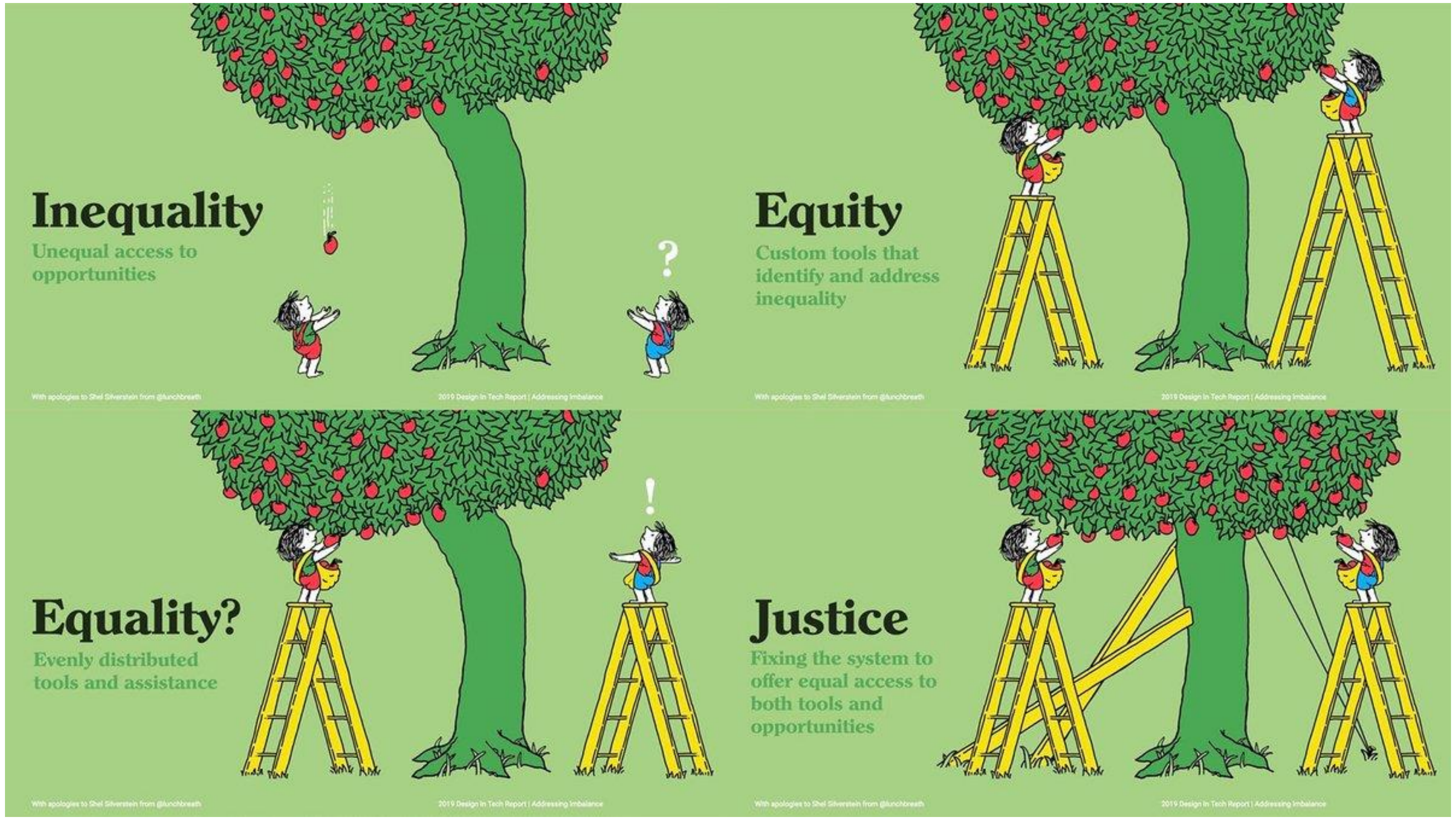


# Bienestar social utilitario

El utilitario es la suma de las utilidades que se le da a los jugadores, pero no tiene en cuenta la distribución de esas utilidades entre los jugadores lo que puede llevar a distribuciones desiguales si eso maximiza la utilidad. Por ejemplo, si un monstruo dice que ama las galletas miles de veces más que los demás, si se le dan todas, como si fuera una subasta, la utilidad reportada al final es máxima, pero con una distribución completamente desigual.

# Bienestar social igualitario

Se encarga de estudiar cómo se distribuyen esas utilidades entre los jugadores. Una propuesta es intentar maximizar la utilidad esperada del miembro de la sociedad más perjudicado. A esto se le llama criterio **maximin**. También está la métrica del **coeficiente de Gini** que indica que tan equitativamente se distribuyen las utilidades entre los jugadores. El problema con esto es que se pueden sacrificar gran parte del bienestar social por lograr esa distribución. Y, además, no es un esquema inmune al monstruo de la utilidad.

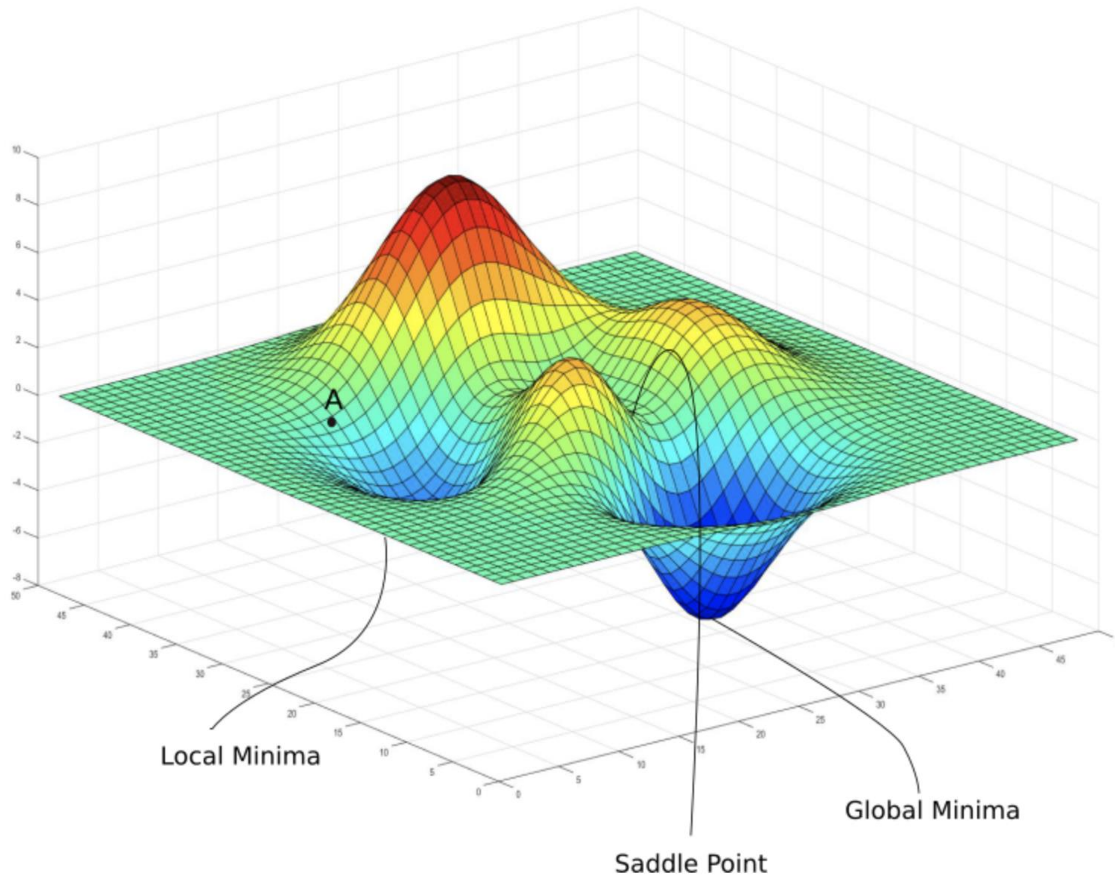


# Cálculo del equilibrio

El equilibrio podría calcularse de forma exhaustiva, iterando sobre todos los posibles perfiles de estrategia, es decir todas las combinaciones de acciones y verificar si ningún jugador tiene incentivo para cambiar de opción. Si es el caso, este perfil de estrategia es un equilibrio de Nash. Sin embargo, para un juego de $n$ jugadores cada uno con $m$ posibles acciones, se tienen $m^n$ posibilidades, lo que lo convierte en un problema **intratable**.

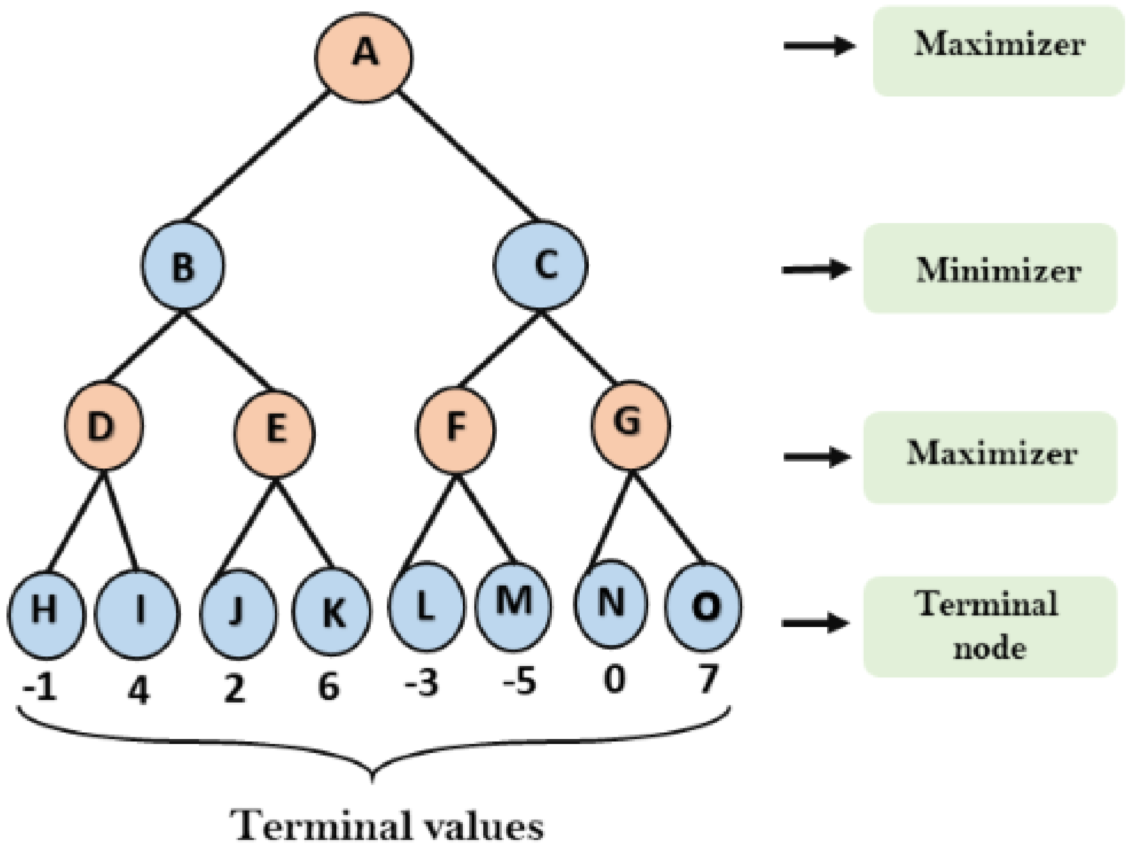
# Mejor respuesta miope

También llamada mejor respuesta iterada, consiste en escoger un perfil de estrategias aleatorio, y verificar si algún jugador no está jugando una acción óptima. Si es el caso, cambiarla por una óptima y repetir el proceso. El algoritmo puede converger al equilibrio de Nash de esta forma, pero no está garantizado para todos los juegos, pero para algunos sí.



# Maximin

Von Neumann propuso un algoritmo para juegos de suma cero de dos jugadores. Este método se llama **maximin**. Tomaremos como ejemplo el juego de Morra de dos dedos. Se escoge un jugador que será el maximizador. En este caso $E$, es decir el jugador que gana cuando el resultado es par. Suponemos que $E$ escoge una estrategia y la revela a $O$. Luego $O$ escoge una estrategia ya sabiendo la de $E$.

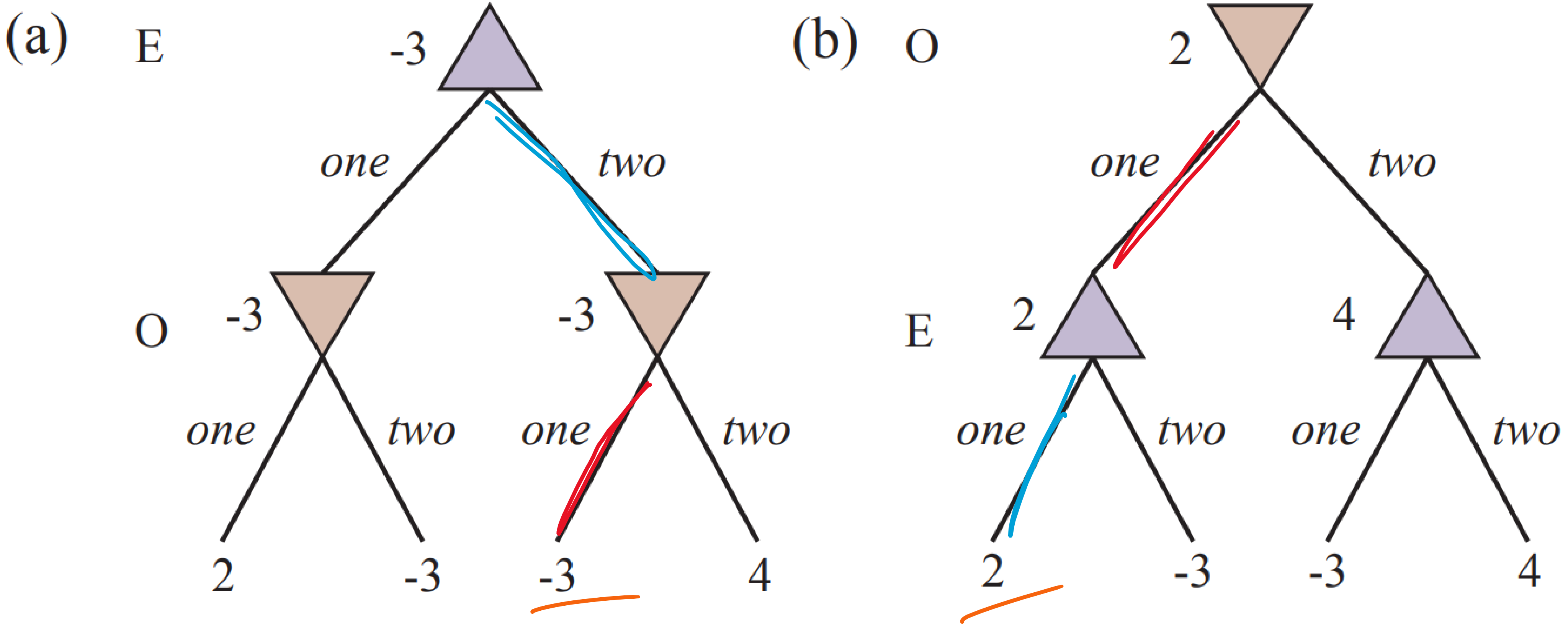



Tomado de: [Artificial Intelligence | Mini-Max Algorithm - Javatpoint]

# Maximin

De aquí se puede seguir un juego de a turnos en el que la utilidad esperada se puede calcular con el algoritmo de minimax (presentado en el capítulo 6 y que no se profundizará aquí), pero en términos simples, el algoritmo simula un juego en el que el jugador maximizador busca maximizar su utilidad y el otro busca minimizar la utilidad del primero. Si $E$ es el primer jugador y es maximizador va a escoger $two$. Como $O$ quiere minimizar la utilidad de $E$ escoge $two$ y entonces la utilidad de $E$ es $-3$. Al contrario, si $O$ empieza jugando, como es minimizador, quiere minimizar la utilidad de $E$ entonces escoge $one$ para que $E$ solo pueda escoger $one$ y la utilidad de $E$ sea $+2$.

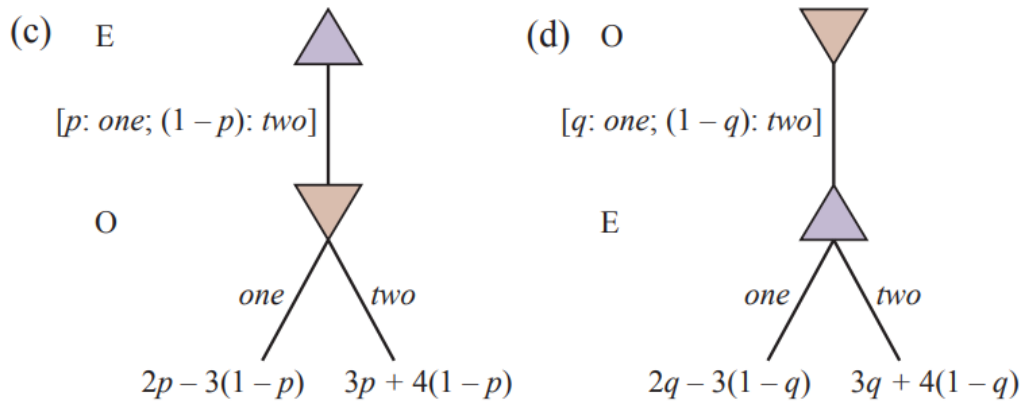
# Maximin




Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Maximin para estrategias aleatorias

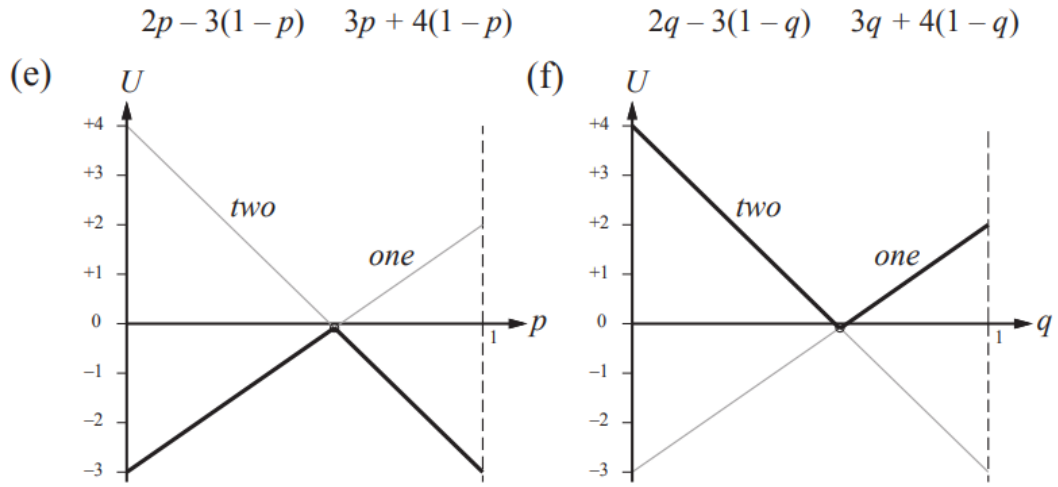
Para estrategias aleatorias, el primer jugador escogerá una acción de acuerdo con una distribución de probabilidad, así: $[p: one; (1-p): two]$. En este caso, el segundo jugador escogerá una estrategia pura o determinística pues no hay forma que pueda obtener mejor resultado usando una estrategia aleatoria también.




Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Maximin para estrategias aleatorias

La utilidad está dada por esta combinación lineal: $p \cdot U_{one} + (1 - p) \cdot U_{two}$. Los valores de $U_{one}$ y $U_{two}$ dependen de qué estrategia pura escoge él segundo jugador. Esas ecuaciones son rectas, una para cada elección del segundo jugador. El valor que el primer jugador debe escoger está en la intersección de ambas líneas.

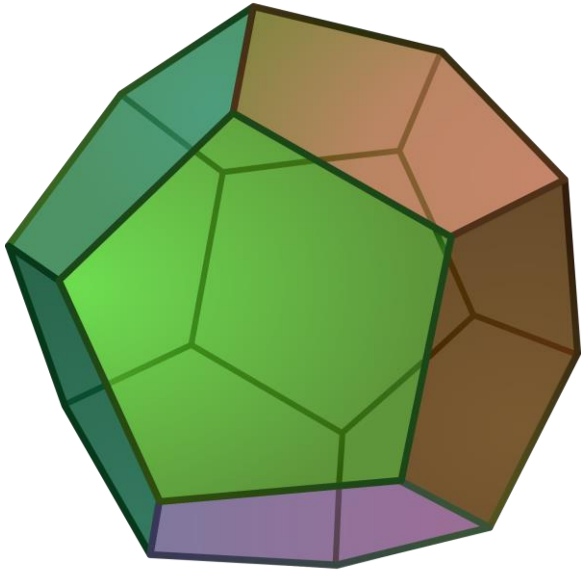



Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Maximin para estrategias aleatorias

En este ejemplo del juego de Morra de dos dedos ese valor de $p$ es $7/12$ y la utilidad esperada para el primer jugador es $-1/12$. De manera que la estrategia aleatoria que debe escoger el primer jugador es $[7/12\colon one; 5/12\colon two]$. Esto podría hacerse con un dado de doce caras. A esta estrategia se le llama maximin y es un equilibrio de Nash.

**Von Neumann** demostró que todo juego tiene un equilibrio maximin si se permiten estrategias aleatorias. Además, todo equilibrio de Nash es un maximin en un juego de suma cero de dos jugadores.

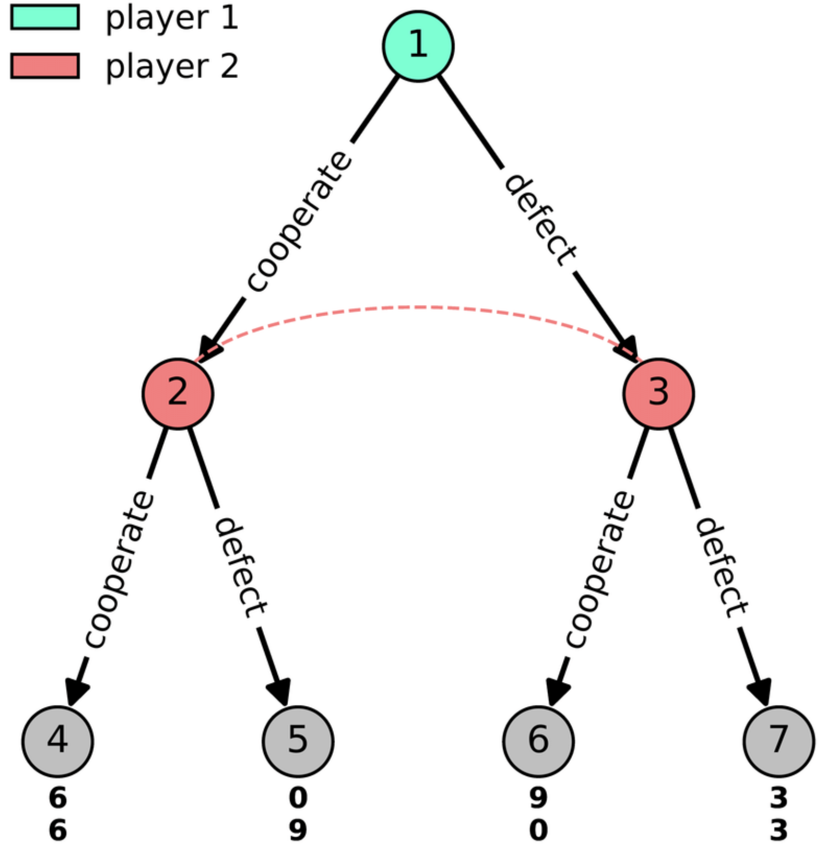


John Von Neumann

# Juegos repetidos o iterados

Los juegos de un solo turno se llaman **juegos de etapa**. Un juego repetido o iterado es un juego de múltiples turnos o movimientos. Es decir, como si se tuvieran varios juegos de un turno uno tras u otro.

En un juego repetido, una estrategia para un jugador ya no es una sola acción, sino un conjunto de acciones, una por turno.



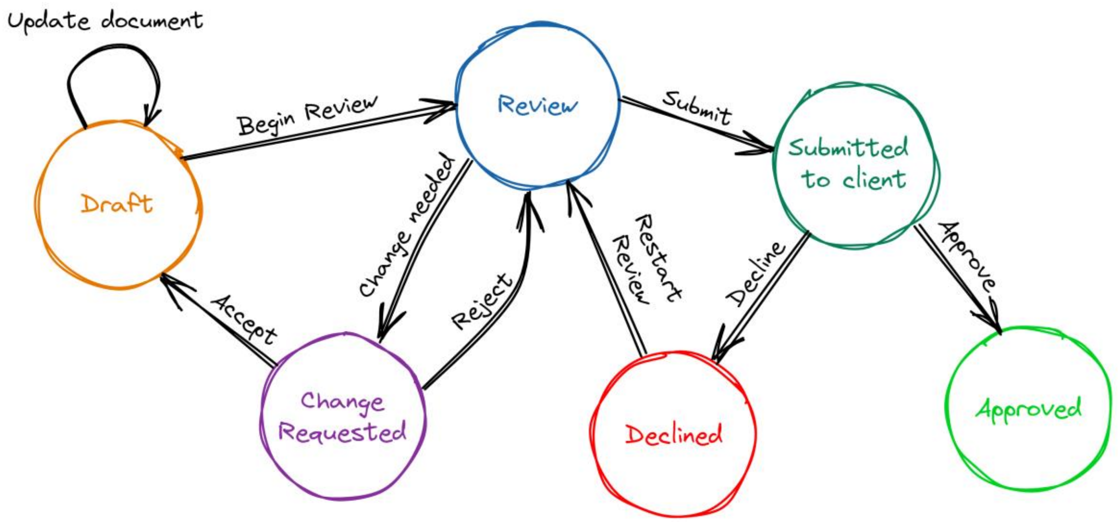

Tomado de: [A computational model of Ostrom's Institutional Analysis and Development framework - ScienceDirect]

# Inducción hacia atrás

Cuando el juego es finito, con un número fijo de etapas que es conocido por ambos jugadores, la estrategia dominante seguirá siendo la misma que en la del juego de etapa. Por ejemplo, en el dilema del prisionero, digamos que el juego tiene 100 turnos. El juego número 100 no es un juego repetido, por lo tanto, la mejor estrategia será la de testificar, pues no hay efectos en las siguientes etapas. Sin embargo, una vez que la ronda 100 está determinada, la 99 también y así sucesivamente hacia atrás. Este razonamiento se llama **inducción hacia a atrás**.

# Estrategias de máquinas de estado finito

Si el juego es infinito, la estrategia es diferente, en este caso es una función que indica una acción en el juego de etapa actual dado el historial del juego hasta el momento. Esto se puede representar con máquinas de estado finito.

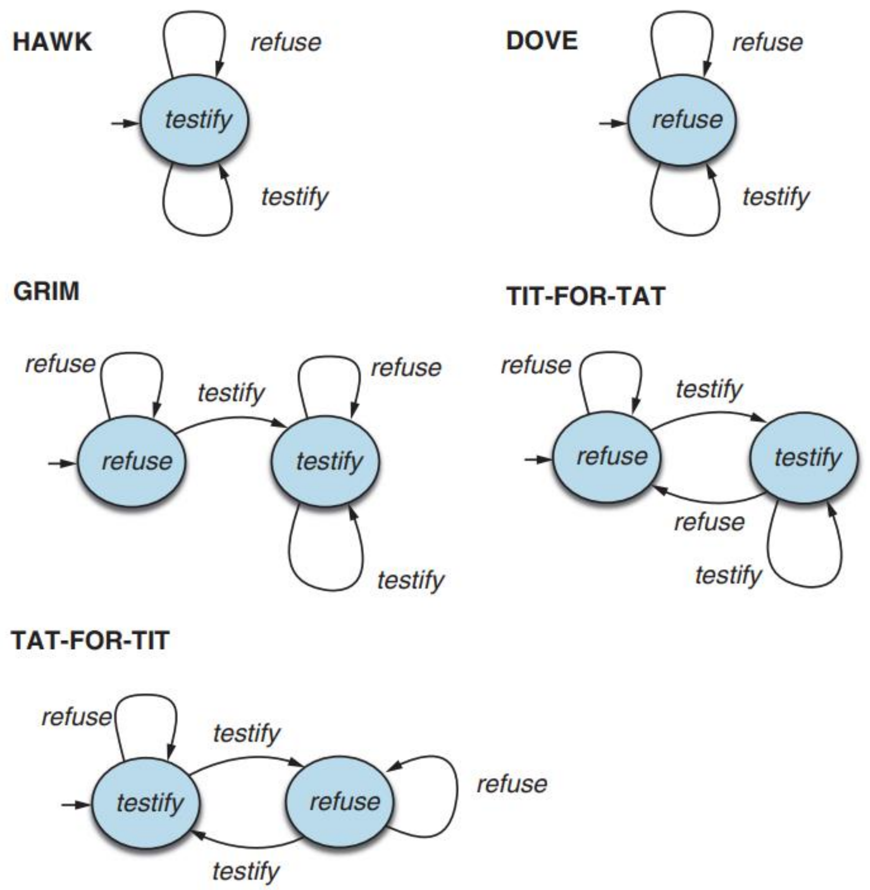



Tomado de: Modelling Workflows With Finite State Machines in .NET - Lloyd Atkinson

# Estrategias de máquinas de estado finito

En estos diagramas, los nodos o estados corresponden a la acción del jugador analizado y los arcos o transiciones a las acciones del oponente.

La estrategia TIT-FOR-TAT consiste en imitar al oponente. La estrategia HAWK consiste en testificar siempre. La DOVE, en rehusarse. La GRIM es como TIT-FOR-TAT con la excepción de que, si el oponente testifica, se convierte a HAWK, es decir sigue testificando.




Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Límite del promedio

Este límite converge cuando se usan estrategias de máquinas de estado finito. La razón es que en una máquina de estado finito se garantiza que en algún momento la secuencia de acciones de un jugador empezará a repetirse, sin importar qué tan elaborada sea. Entonces uno puede asumir que cada jugador escoge una estrategia, o es decir una máquina.

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} U_t$$

# Equilibrio de Nash se conserva en el caso infinito

| | 0 | 1 | 2 | 3 | 4 | 5 | … | |
|---|---|---|---|---|---|---|---|---|
| Ali: HAWK | *testify* | *testify* | *testify* | *testify* | *testify* | *testify* | … | utility = −5 |
| Bo: HAWK | *testify* | *testify* | *testify* | *testify* | *testify* | *testify* | … | utility = −5 |

Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Nuevos equilibrios de Nash en el caso infinito

| | 0 | 1 | 2 | 3 | 4 | 5 | ... | |
|---|---|---|---|---|---|---|---|---|
| Ali: GRIM | *refuse* | *refuse* | *refuse* | *refuse* | *refuse* | *refuse* | ... | utility $= -1$ |
| Bo: GRIM | *refuse* | *refuse* | *refuse* | *refuse* | *refuse* | *refuse* | ... | utility $= -1$ |

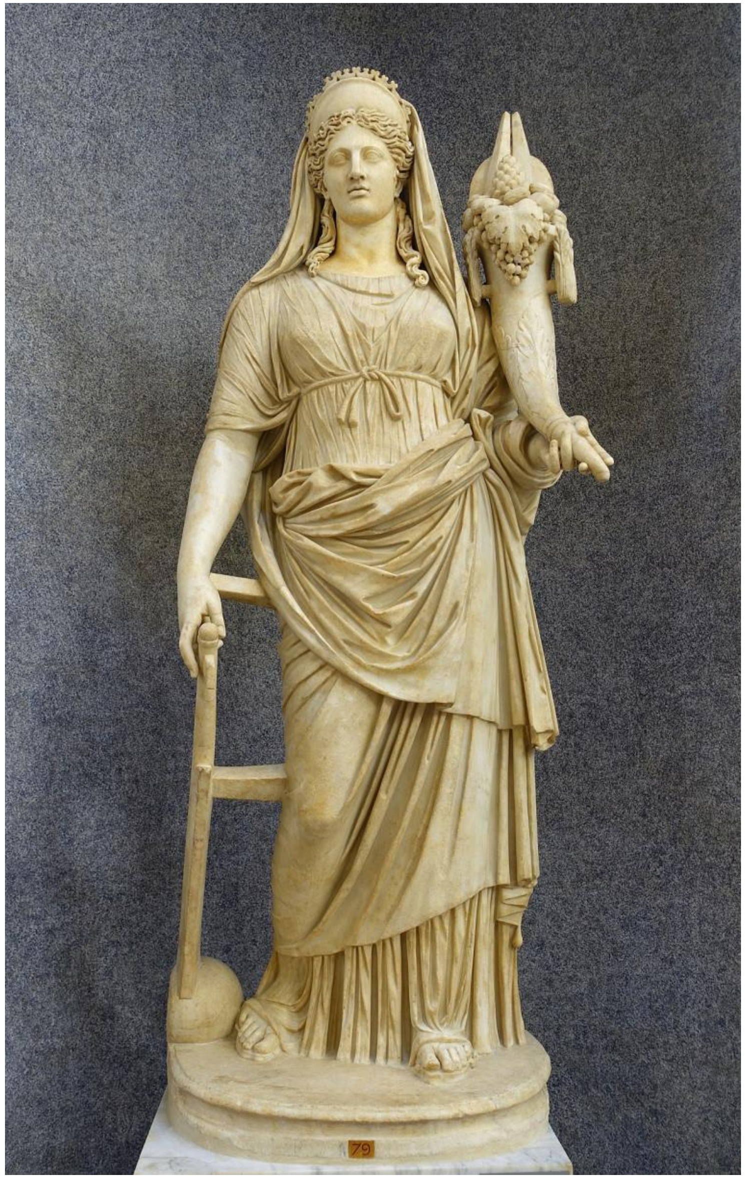
Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# La estrategia GRIM

La intuición detrás de la estrategia GRIM es que hay una amenaza mutua de que se castigará al oponente si este no cumple su parte para obtener el resultado deseado, que en este caso sería -1. Si alguno de los dos cambiara de estrategia y testificará, en ese caso el otro jugador pasará a modo castigo y testificará por siempre, haciendo que el otro jugador no pueda obtener una utilidad mayor a -5, que es peor de lo que obtendría si hubiera continuado con GRIM.

# El jugador azar

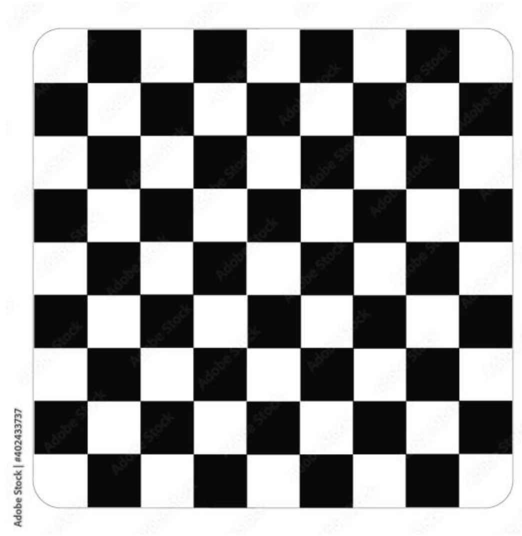
En juegos estocásticos se agrega un jugador adicional al juego llamado azar. Este jugador puede tomar acciones aleatorias de acuerdo con una distribución de probabilidad. La manera en la que esto se modela es haciendo que azar realice su acción aleatoria justo después de cada jugador realice su acción determinística. La estrategia de azar es parte de la definición del juego.



Tomado de: [Fortuna (mitología) - Wikipedia, la enciclopedia libre]

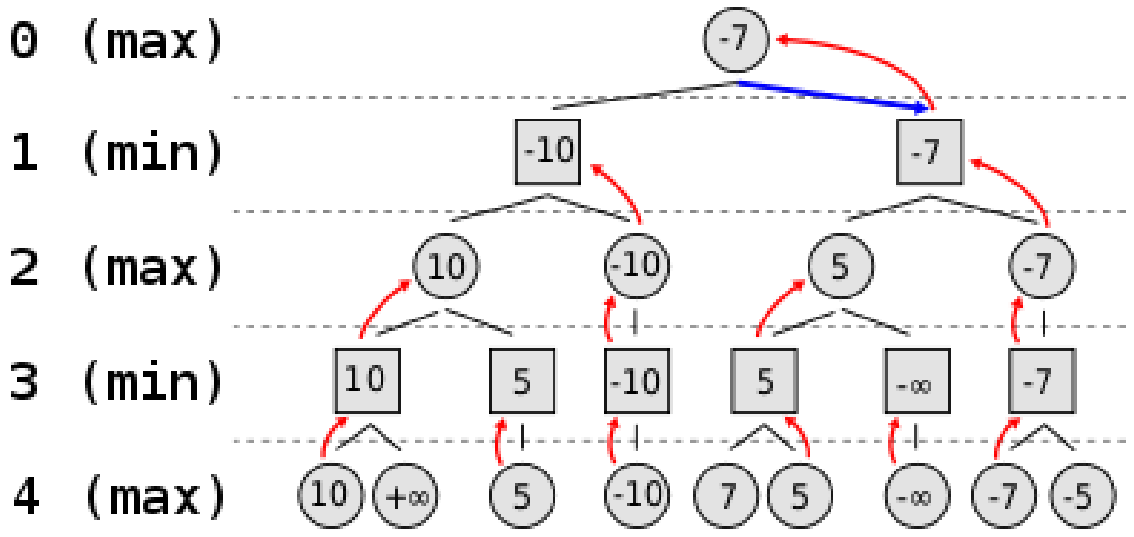
# Juegos de información perfecta e imperfecta

Los jugadores pueden tener información perfecta o imperfecta. Tener información perfecta significa que los jugadores siempre saben en qué parte del árbol del juego están, como ocurre en el ajedrez o en Go. Lo contrario ocurre en juegos como el póker.



# Cálculo del equilibrio en árboles del juego

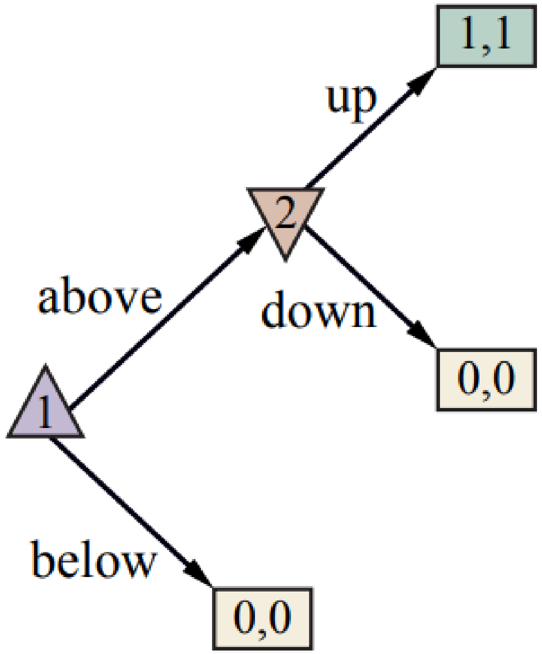
El "árbol del juego" es una representación que se denomina juego en forma extensiva. El equilibrio de Nash se puede calcular en estos árboles utilizando la técnica del "minimax con inducción hacia atrás", que se basa en la programación dinámica. El cálculo comienza desde los nodos terminales del árbol, donde se etiquetan con una estrategia que maximiza la recompensa del jugador en ese turno del juego. Luego, una vez que todos los nodos hijos han sido etiquetados, el proceso se repite en el nodo padre y así sucesivamente.




Tomado de: [Minimax - Wikipedia, la enciclopedia libre]

# Amenazas creíbles

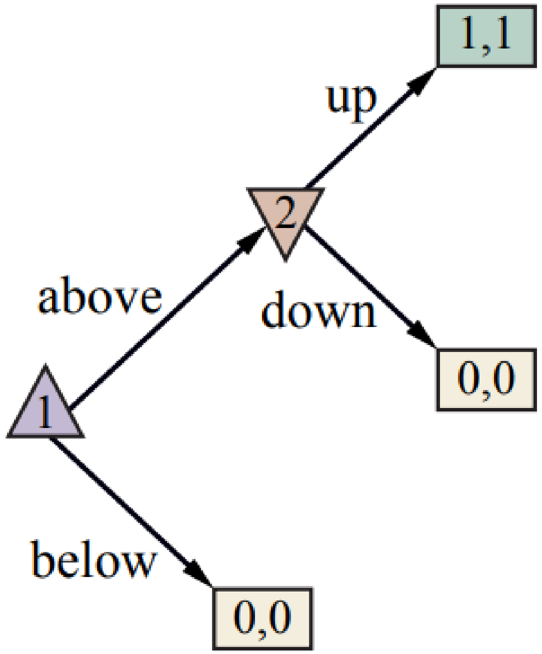
Al analizar estos árboles, a veces se encuentran equilibrios de Nash que pueden parecer contraintuitivos, como el que se presenta a continuación: la combinación "below, down" es un equilibrio de Nash. Parecería que el jugador 2 está amenazando al jugador 1 con moverse hacia abajo si el jugador 1 se mueve hacia arriba. Sin embargo, esta amenaza no es creíble, ya que lo mejor para el jugador 2 es moverse hacia arriba en el segundo turno.




Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Equilibrio de Nash perfecto de subjuegos

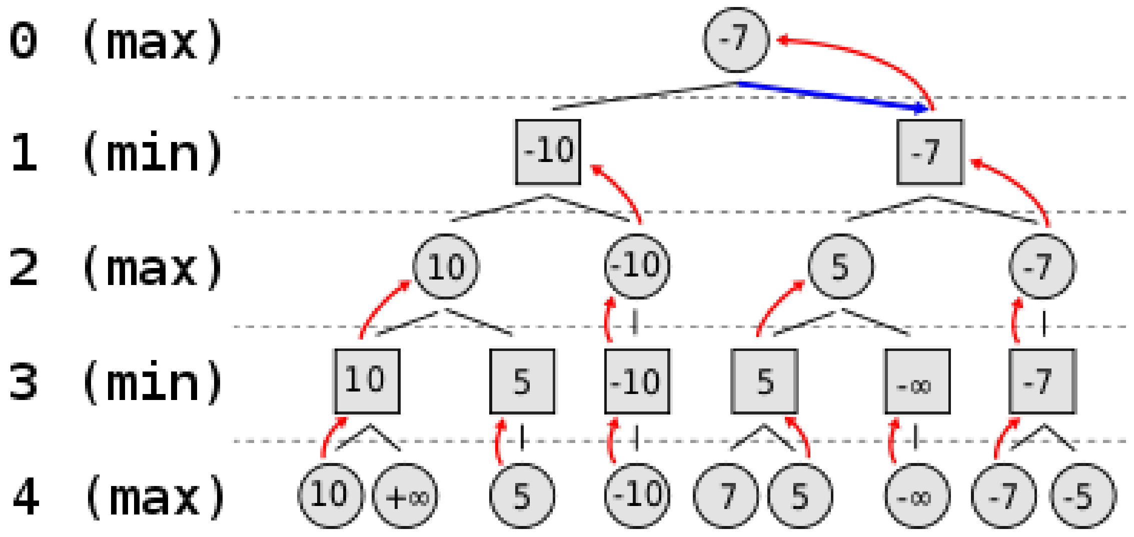
Debido a estas ambigüedades en el cálculo del equilibrio de Nash, surge un concepto más estricto llamado "equilibrio de Nash perfecto de subjuegos" (ENPS). Un "subjuego" se puede considerar como un subárbol en el árbol del juego. En la imagen de ejemplo, hay dos juegos. Un ENPS se alcanza cuando se logra un equilibrio en todos los subjuegos del juego. Con esta definición, “above, up" sería un EFPS. La ventaja es que todo equilibrio de Nash calculado mediante la inducción hacia atrás es un EFPS, y lo que es aún mejor, todo juego en forma extensiva tiene garantizado un EFPS.




Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Inducción hacia atrás en juegos de información imperfecta

La inducción hacia atrás no funciona con juegos con información imperfecta. En estos juegos no hay garantía de que un juego termine, y por esta razón la técnica de **minimax** no puede encontrar una solución.

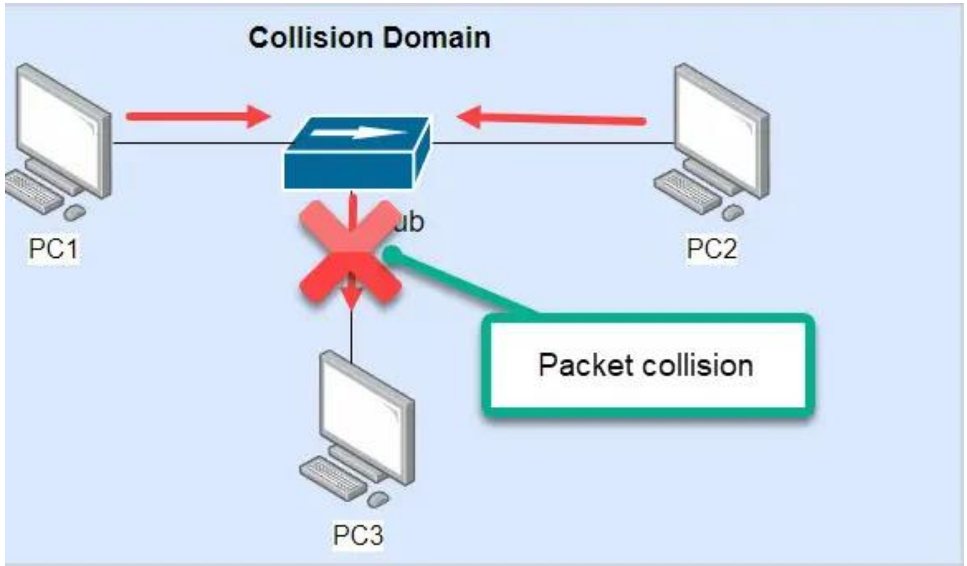



Tomado de: Minimax - Wikipedia, la enciclopedia libre

# Intersección de tráfico

Un ejemplo simple es el de un juego en el que dos jugadores están en un tablero 4x4 con movimientos simultáneos. Cuando un jugador alcanza un cuadro de salida, obtiene una recompensa. Sin embargo, si se define que los jugadores no pueden moverse si intentan moverse al mismo cuadro, como si fuera la intersección de tráfico, los jugadores podrían quedarse atascados por siempre si solo usan estrategias puras. Si en cambio usan **estrategias mixtas**, lograrían salir del atascamiento.

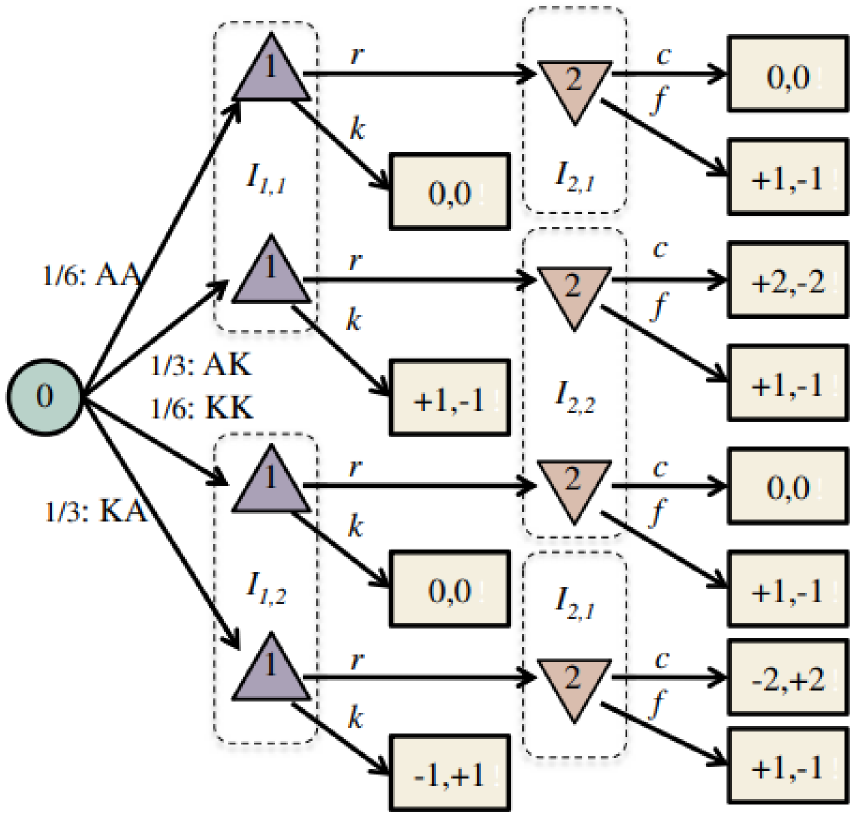
Esto se hace en resolución de paquetes en redes de Ethernet.




Tomado de: [Collision Domain Explained - NetworkVerge]

# Conjuntos de información

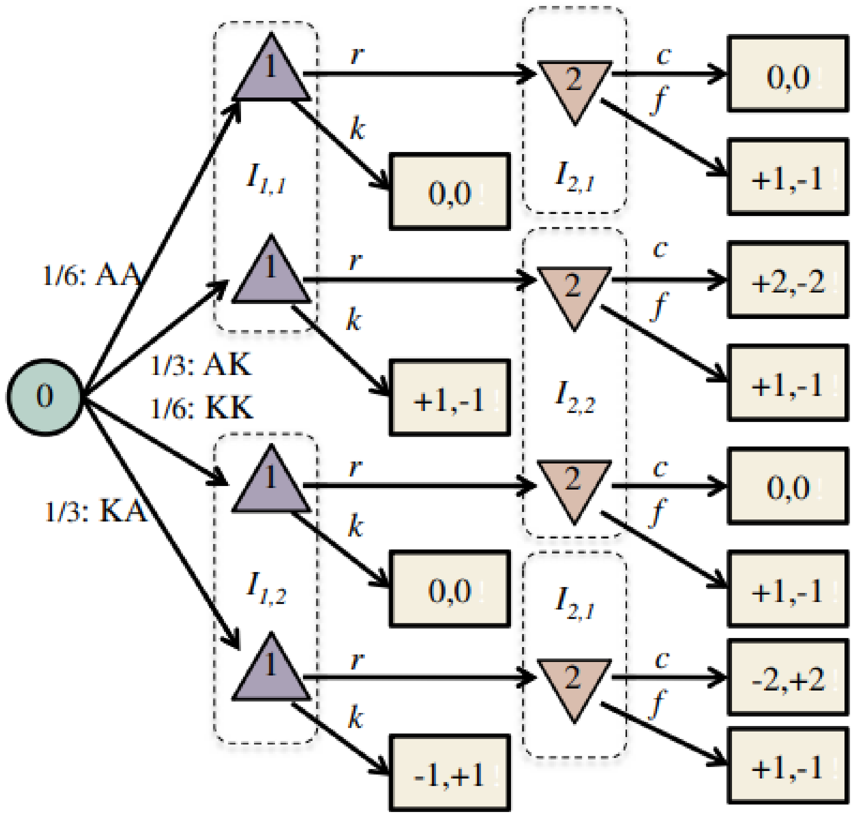
En juegos de información imperfecta, es posible resolverlos utilizando la representación en forma extensa. Esto se debe a que esta representación incluye los estados de información de todos los jugadores al mismo tiempo, lo que permite reducir el juego a una forma normal.




Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Conjuntos de información

El juego tiene solo un turno. El nodo raíz corresponde al movimiento del jugador suerte, que en este caso actúa de primeras repartiendo las cartas. A significa ace, y K significa rey. Las acciones de los jugadores 1 y 2 son subir (r), pasar (k), doblar (f) e igualar (c). Las secciones encerradas líneas punteadas son los conjuntos de información, o estados de creencias de los jugadores en cada turno. Por ejemplo, el $I_{1,1}$ es el estado de información en el que el jugador 1 sabe que tiene un as y no sabe que tiene el jugador 2. El $I_{2,1}$ es el estado en el que el jugador 2 sabe que él sabe que tiene un as y que el jugador 2 ha subido.




Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Matriz de recompensa cuando hay conjuntos de información

| | 2:*cc* | 2:*cf* | 2:*ff* | 2:*fc* |
|---|---|---|---|---|
| 1:*rr* | 0 | -1/6 | 1 | 7/6 |
| 1:*kr* | -1/3 | -1/6 | 5/6 | 2/3 |
| 1:*rk* | 1/3 | **0** | 1/6 | 1/2 |
| 1:*kk* | 0 | **0** | 0 | 0 |

Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

En estos juegos con información imperfecta, cada acción es una combinación por cada conjunto de información disponible en ese turno para ese jugador.

Por ejemplo, la estrategia $rk$ sería subir en $I_{1,2}$ es decir cuando sé que tengo un as y pasar en $I_{2,1}$ es decir cuando sé que tengo un rey.

Con la representación en matriz es posible encontrar los equilibrios fácilmente, que en este caso son $rk, cf$ y $kk, cf$.

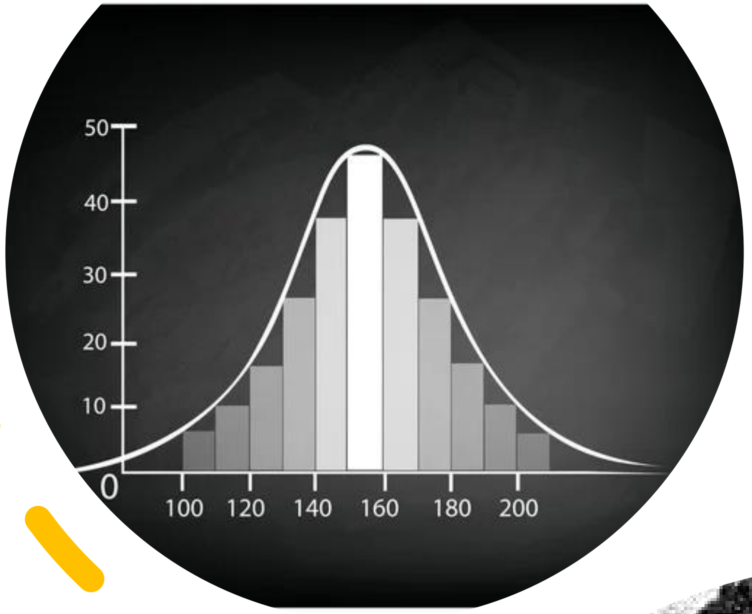
# Abstracciones

Cuando hay muchos estados, es posible encontrar soluciones aproximadas mediante abstracciones, como clasificar las manos como mejores, iguales o peores que la propia. Estas abstracciones forman parte de las estrategias utilizadas por Pluribus, un sistema que ha logrado derrotar a campeones mundiales en el póker Texas Hold'em.

Tomado de: [Remembering Pluribus: The Techniques that Facebook Used to Master World's Most Difficult Poker Game - KDnuggets]

# Equilibrio de Bayes-Nash

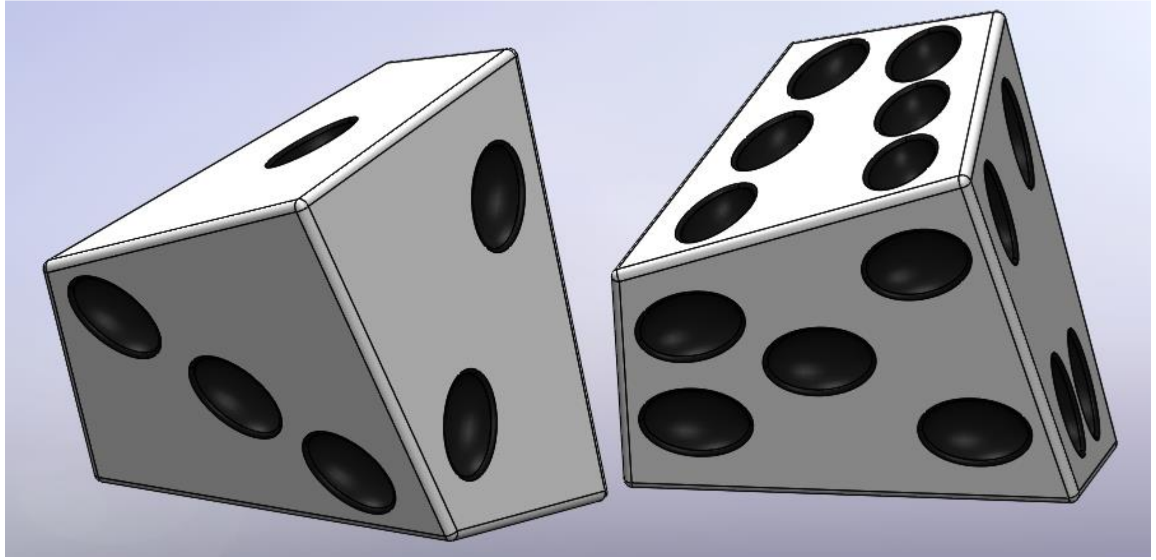
Cuando los jugadores no son completamente racionales, se puede utilizar el equilibrio de Bayes-Nash. Este equilibrio tiene en cuenta la distribución de probabilidad a priori de las estrategias de los jugadores, reflejando las creencias de los jugadores sobre las posibles estrategias de los demás. Es como aprender los patrones del otro jugador, incluso si no es completamente racional, y aprovechar esas creencias para obtener ventaja.




Thomas Bayes

# Acciones inciertas

En el caso de un juego que depende de un dado, simplemente se agrega un nodo adicional al árbol. Sin embargo, si el dado está sesgado, se agrega otro nodo justo antes para modelar esa posibilidad. Además, si el oponente conoce si el dado es justo o no, se puede agregar otro nodo antes, y así sucesivamente.

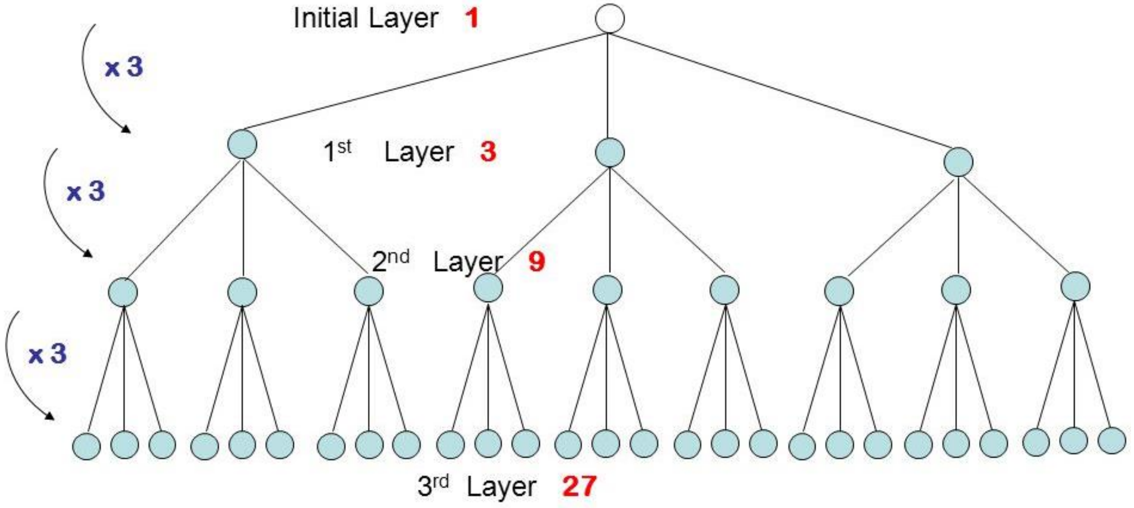


# Utilidades inciertas

Si no tenemos información sobre las utilidades del oponente o si no conocemos nuestras propias utilidades, como cuando no estamos seguros de si nos gustará la comida que pedimos en un restaurante, también se puede modelar agregando un nodo adicional al árbol.

# Incertidumbre en la teoría de juegos

En resumen, la teoría de juegos es capaz de manejar esta incertidumbre, pero cada pregunta adicional implica agregar más nodos al árbol, lo que puede hacer que algunos juegos sean muy complejos de tratar.




Tomado de: [8.5 Exponential Growth and 8.6 Exponential Decay FUNctions - ppt video online download (slideplayer.com)]

# Juegos de asistencia

Imagina un juego de asistencia entre Harriet, una humana, y Robbie, un robot. Harriet sigue sus preferencias, mientras que Robbie tiene una probabilidad a priori sobre las preferencias de Harriet. Ambos buscan una recompensa, thita, que es igual para los dos. Sus acciones incluyen enseñar, recompensar, corregir y dar recomendaciones, y ambos aprenden estas acciones en lugar de ser programadas.

# Juegos de asistencia

El objetivo es producir pisapapeles y cosedoras, con recompensas basadas en su valor de mercado. Por ejemplo, un pisapapeles puede valer $0.45 y una cosedora $0.55.

Robbie comienza con una distribución uniforme sobre las preferencias de Harriet. Harriet tiene tres opciones iniciales: hacer dos pisapapeles, dos cosedoras o uno de cada uno. Luego, Robbie decide su producción, y aquí entra la complejidad: Harriet elige basándose en cómo Robbie la interpretará, y Robbie elige en función de su interpretación de Harriet.

Para encontrar un equilibrio de Nash, utilizamos una estrategia de mejor respuesta miope: Harriet elige aleatoriamente, luego Robbie responde con su mejor estrategia, y así sucesivamente. El juego se equilibra cuando Harriet elige hacer uno de cada uno, y Robbie produce 50 de cada producto.

# Teoría de juegos cooperativos

# Juego cooperativo

En este tipo de juegos, los agentes pueden hacer acuerdos para cooperar. De esa forma, se benefician porque al final pueden obtener recompensas mayores que las que obtendrían si jugaran solos.

A continuación, se presentará un modelo de estos juegos. La idea principal es que cuando un grupo de agentes coopera, el grupo obtiene una utilidad, que luego puede ser repartida entre los miembros.

Un juego cooperativo se define como $G = (N, v)$ donde $N$ es un conjunto de $n$ jugadores $N = 1, 2, \dots, n$ y $v$ es una función característica que para cada subconjunto de jugadores $C \subseteq N$ indica el valor de recompensa que los jugadores obtendrían si jugarán juntos.

Por definición $v(\emptyset) = 0$, $v(C) > 0$ y para algunos juegos $v(\{i\}) = 0$.

# Coaliciones

Un conjunto de jugadores $C$ se llama coalición. El conjunto de todos los jugadores $N$ se llama gran coalición.

En este modelo, cada jugador debe escoger exactamente una coalición, de manera que todas las coaliciones son disyuntas y su unión conforma $N$, es decir, forman una partición de $N$. A esta partición se le conoce como estructura de coalición.

Una estructura de coalición sobre un conjunto de jugadores $N$ es un conjunto de coaliciones $\{C_1 \dots C_k\}$ que cumple estas propiedades:

$$C_i \neq \emptyset; C_i \subseteq N; C_i \cap C_j = \emptyset, \forall\, i \neq j; \bigcup_{i=1}^{k} C_i = N$$

Tomado de: Elecciones Colombia 2022: La coalición del aburrimiento | Opinión | EL PAÍS (elpais.com)

# Resultado de un juego cooperativo

El conjunto de todas las estructuras de coalición sobre un conjunto de jugadores $N$ se denota como $\mathbf{CS}(N)$ y $CS(i)$ es la coalición dentro de la estructura donde está el jugador $i$.

El resultado de un juego cooperativo es una pareja $(CS, \mathbf{x})$ que consiste en una estructura de coalición $CS$ y un vector de recompensa $\mathbf{x} = (x_1, \ldots, x_n)$ donde $x_i$ es la recompensa que obtiene el jugador $i$ del juego. Este vector de recompensa debe satisfacer:

$$\sum_{i \in C} x_i = v(C) \ \forall C \in CS$$

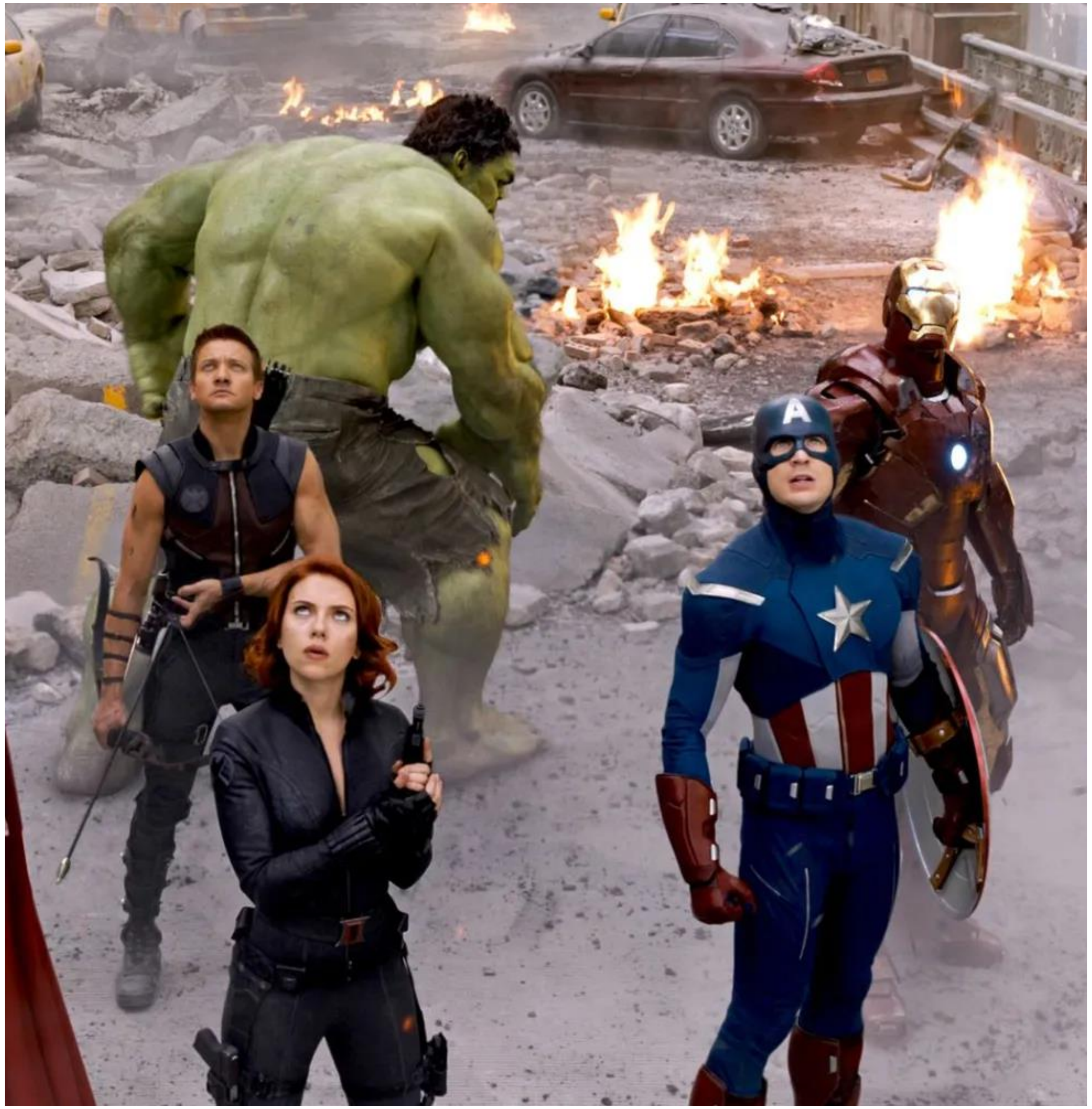
Es decir, que la suma de las recompensas en una coalición debe ser igual a la recompensa total de dicha coalición.

# Superaditividad

Algunos juegos tienen la propiedad de superaditividad, que consiste en que, si dos coaliciones se unen, obtendrían una recompensa no menor a la que obtendrían si no se unieran.

$$v(C \cup D) \geq v(C) + v(D)$$

En este tipo de juegos, la gran coalición recibiría una recompensa que sería igual de alta o más alta que el total recibido por cualquier otra coalición, pero rara vez este tipo de coalición terminan por conformarse.



# Imputación

Una imputación para un juego cooperativo $(N, v)$ es un vector de recompensa $\mathbf{x}$ que satisface dos condiciones:

- La imputación distribuye la utilidad de la gran coalición.
- Cada jugador no puede obtener menos utilidad que si trabajará solo (racionalidad individual).

Dada una imputación $\mathbf{x}$ y una coalición $C$, $x(C)$ es la cantidad total distribuida a $C$ por la imputación **x.**

# Núcleo de un juego cooperativo

El núcleo de un juego es el conjunto de todas las imputaciones que satisfacen $x(C) \geq v(C)$ para toda coalición en la estructura. Esto significa que si una imputación no está en el núcleo entonces existe alguna coalición para la cual $v(C) > x(C)$, lo que significa que los jugadores en $C$ se rehusarían a unirse a la gran coalición, pues estarían mejor trabajando en su equipo.

Es decir, que el núcleo de un juego es el conjunto de todas las posibles recompensas para las cuales ninguna coalición puede objetar unirse a la gran coalición. Por otro lado, si el núcleo está vacío, la gran coalición no se formará ya que siempre habrá alguna coalición más pequeña que se rehúse a unirse.

# Cálculo del núcleo

La cuestión computacional más importante en relación con el núcleo son el verificar si está vacío o no o si cierto vector de recompensa **x** hace parte del núcleo. Estos problemas se pueden resolver con programación lineal pero el número de ecuaciones crece exponencialmente, y a que el número de coaliciones posibles, es decir el tamaño de $\mathbf{CS}(N)$ crece exponencialmente con el número de jugadores. Para muchos juegos cooperativos este problema es intratable.

# Significado del núcleo

La noción del núcleo nos dice si una gran coalición puede formarse, pero no nos dice cómo se debe distribuir las utilidades entre sus miembros.

# Distribución de utilidades y el valor de Shapley

Para esto se puede utilizar el valor de Shapley, cuyo propósito es lograr una distribución justa de las utilidades $v(N)$ entre los miembros. Por justo se refiere a que las utilidades se distribuyen entre los jugadores dependiendo de qué tanto ha contribuido cada jugador en crear el valor $v(N)$. Para esto es útil entender la definición de contribución marginal.

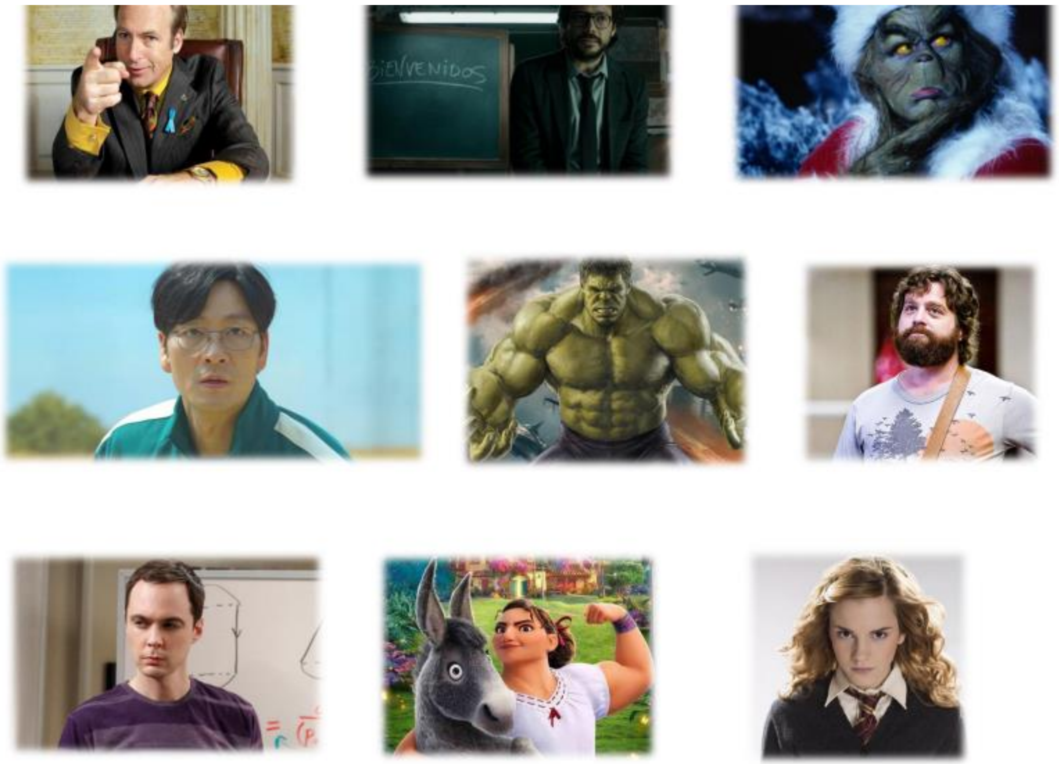


Imagen tomada de: [PROFI (unal.edu.co)]

# Contribución marginal

La contribución marginal que hace un jugador $i$ en una coalición $C$ es el valor que $i$ agrega si $i$ se une a la coalición. Esta contribución marginal podría ser negativa.

$$mc_i(C) = v(C \cup \{i\}) - v(C)$$

# Definición del valor de Shapley

Para determinar el valor con que se debe corresponder el jugador $i$, se debe hacer el cálculo del promedio de las contribuciones marginales que $i$ hace a los jugadores anteriores a él en todos los posibles ordenamientos (permutaciones) de $N$.

Digamos que $P$ es el conjunto de todas las permutaciones de $N$ y que $p_i$ es el conjunto de jugadores que preceden a $i$ en un ordenamiento $p \in P$, entonces el valor de Shapley de un juego $G$ es una imputación, es decir, un vector de recompensas $\phi(G) = (\phi_1(G), \dots, \phi_n(G))$ donde cada componente, es decir el valor que se le paga a cada jugador se define como:

$$\phi_i(G) = \frac{1}{n!} \sum_{p \in P} mc_i(p_i) = \frac{1}{n!} \sum_{p \in P} v(p_i \cup \{i\}) - v(p_i)$$

# Axiomas de la justicia

Lo interesante del valor de Shapley es que es la única solución que satisface los axiomas que caracterizan a una distribución justa, o los axiomas de la justicia:

- *Efficiency*: $\sum_{i \in N} \phi_i(G) = v(N)$. (All the value should be distributed.)
- *Dummy Player*: If $i$ is a dummy player in $G$ then $\phi_i(G) = 0$. (Players who never contribute anything should never receive anything.)
- *Symmetry*: If $i$ and $j$ are symmetric in $G$ then $\phi_i(G) = \phi_j(G)$. (Players who make identical contributions should receive identical payoffs.)
- *Additivity*: The value is additive over games: For all games $G = (N, v)$ and $G' = (N, v')$, and for all players $i \in N$, we have $\phi_i(G + G') = \phi_i(G) + \phi_i(G')$.

# Axiomas de la justicia

Estos axiomas son: eficiencia, es decir que no se debe desperdiciar utilidad, sino que toda debe ser repartida. Jugador ficticio, que dice que si un jugador no contribuye entonces no se le pagará nada. Simetría, que dice que si dos jugadores hacen contribuciones idénticas deben recibir la misma recompensa y aditividad, el cual tiene un valor más teórico que práctico.

# Cálculo del valor de Shapley

El valor de Shapley se calcula en tiempo polinomial utilizando redes de contribución marginal, aunque no todas las funciones características se pueden representar de esta manera. Sin embargo, en el peor de los casos, la red puede tener un número de nodos igual al número total de coaliciones, lo cual es exponencial en función del número de jugadores.

## Marginal Contribution Nets: A Compact Representation Scheme for Coalitional Games*

Samuel Ieong[†]
Computer Science Department
Stanford University
Stanford, CA 94305
sieong@stanford.edu

Yoav Shoham
Computer Science Department
Stanford University
Stanford, CA 94305
shoham@stanford.edu

**ABSTRACT**

We present a new approach to representing coalitional games based on rules that describe the marginal contributions of the agents. This representation scheme captures characteristics of the interactions among the agents in a natural and concise manner. We also develop efficient algorithms for two of the most important solution concepts, the Shapley value and the core, under this representation. The Shapley value can be computed in time linear in the size of the input. The emptiness of the core can be determined in time exponential only in the treewidth of a graphical interpretation of our

This payoff is intended to reflect the payoff the group of agents can secure for themselves regardless of the actions of the agents not in the group. These choices of primitives are in contrast to those of non-cooperative games, of which agents are modeled independently, and their payoffs depend critically on the actions chosen by the other agents.

### 1.1 Coalitional Games and E-Commerce

Coalitional games have appeared in the context of e-commerce. In [7], Kleinberg *et al.* use coalitional games to study recommendation systems. In their model, each individual

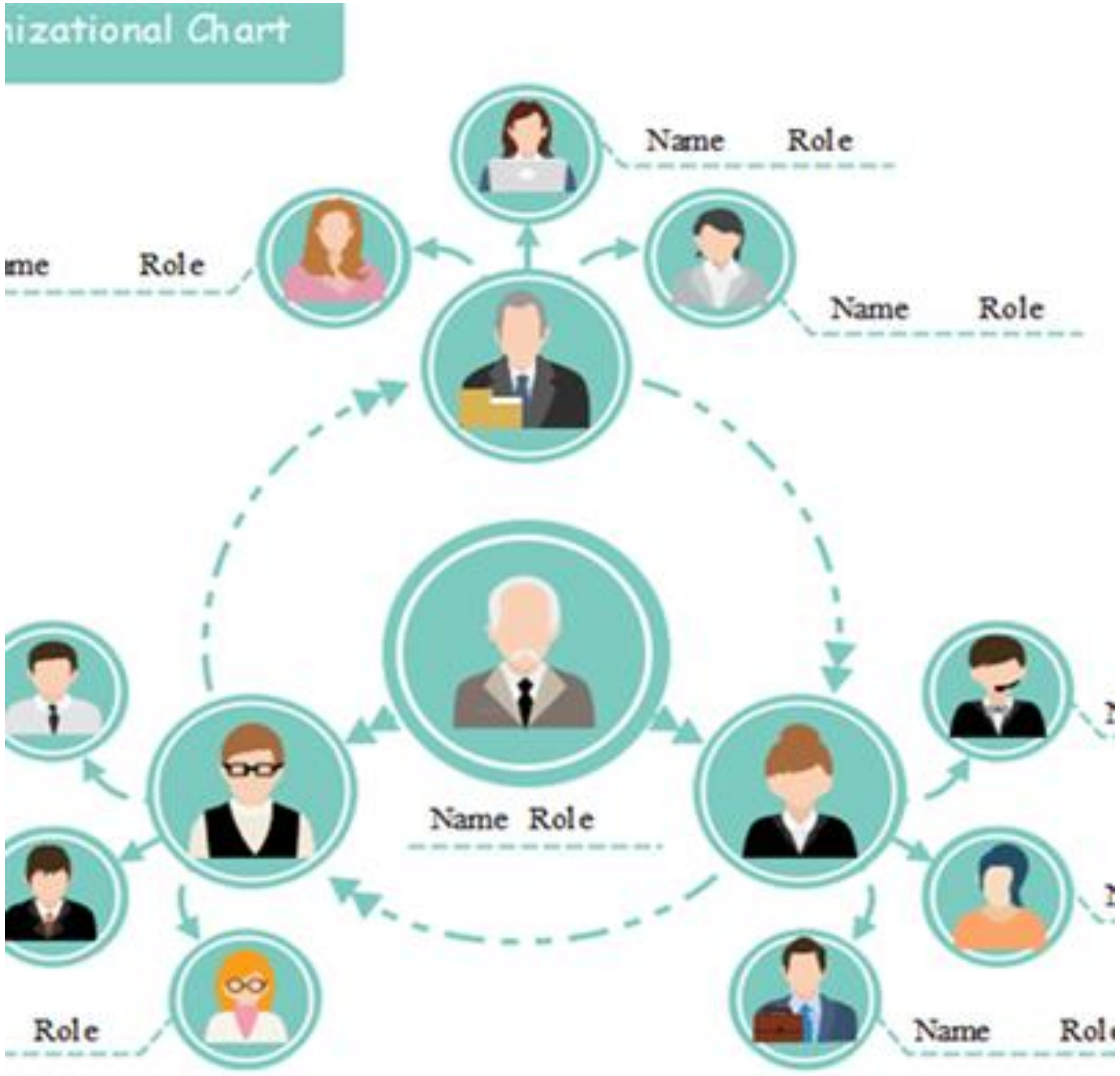
Tomado de: [Marginal contribution nets | Proceedings of the 6th ACM conference on Electronic commerce]

# Coaliciones para maximizar bienestar social

En el ejemplo de una empresa, estaríamos interesados en organizar a los jugadores en coaliciones que maximicen la productividad general.

El bienestar social de una estructura de coaliciones es la suma de las utilidades de todas las coaliciones en esa estructura:

$$sw(CS) = \sum_{C \in CS} v(C)$$




Tomado de: [plantillas para organigramas gratuitas (edrawsoft.com)]

# Estructura de coaliciones óptima

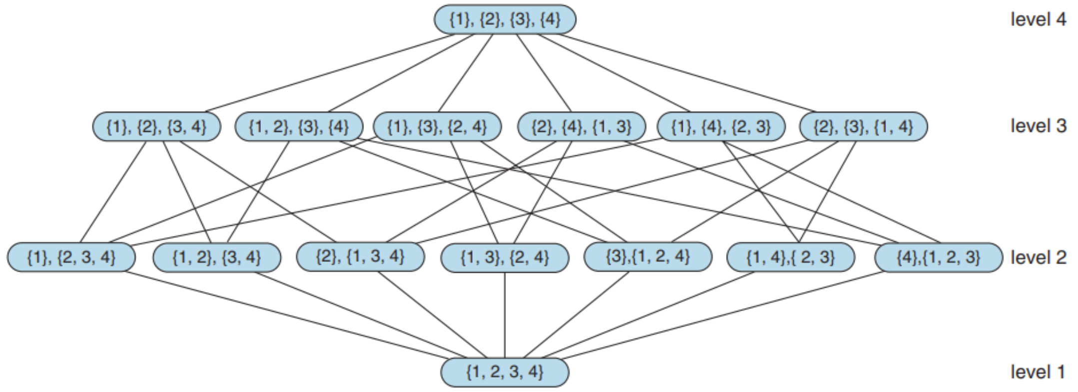
Una estructura de coaliciones óptima $CS^*$ es aquella que maximiza ese valor.

Este problema es muy conocido y se llama el problema del particionado de un conjunto. Este problema es NP-Hard ya que el número de estructuras es exponencial con el número de jugadores.

Tomado de: [Conditional (Partitioned) Probability — A Primer – Math ∩ Programming (jeremykun.com)](jeremykun.com)

# Estructura de coaliciones óptima

Para abordarlo, se pueden encontrar soluciones aproximadas explorando un grafo de estructuras de coaliciones. Este grafo crece exponencialmente, pero se organiza de manera que cada nivel tenga estructuras con el mismo número de coaliciones, lo que permite explorar solo algunos niveles para encontrar una solución $CS'$ cuyo valor de bienestar social es al menos una $1/n$ parte de la solución real $CS^*$, donde $n$ es el número de agentes.




Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# Técnicas especializadas de toma de decisiones multiagente

# Diseño de mecanismos

Esta sección se centra en el diseño de mecanismos, y no tanto en el diseño de agentes. El diseño de mecanismos es el problema de diseñar el juego correcto para que lo juegue una colección de agentes.

Un mecanismo consiste en:

- Un lenguaje para expresar un conjunto de estrategias que puede tomar cada agente
- Un jugador llamado centro que recolecta reportes de las estrategias escogidas por los jugadores: un ejemplo es un subastador.
- Una regla de resultado que conocen todos los agentes y que el centro usa para determinar las recompensas dadas sus estrategias escogidas.

# Red de contrato

El protocolo de red de contrato es la técnica más vieja y probablemente la más importante estudiada en inteligencia artificial. Es un protocolo para la compartición de tareas. Se inspira en la forma en la que las empresas usan contratos. El protocolo consiste en cuatro pasos:

1. Un agente identifica una tarea que debe realizarse colaborativamente debido a limitaciones individuales o porque la colaboración es más beneficiosa en términos de tiempo, costo o calidad.

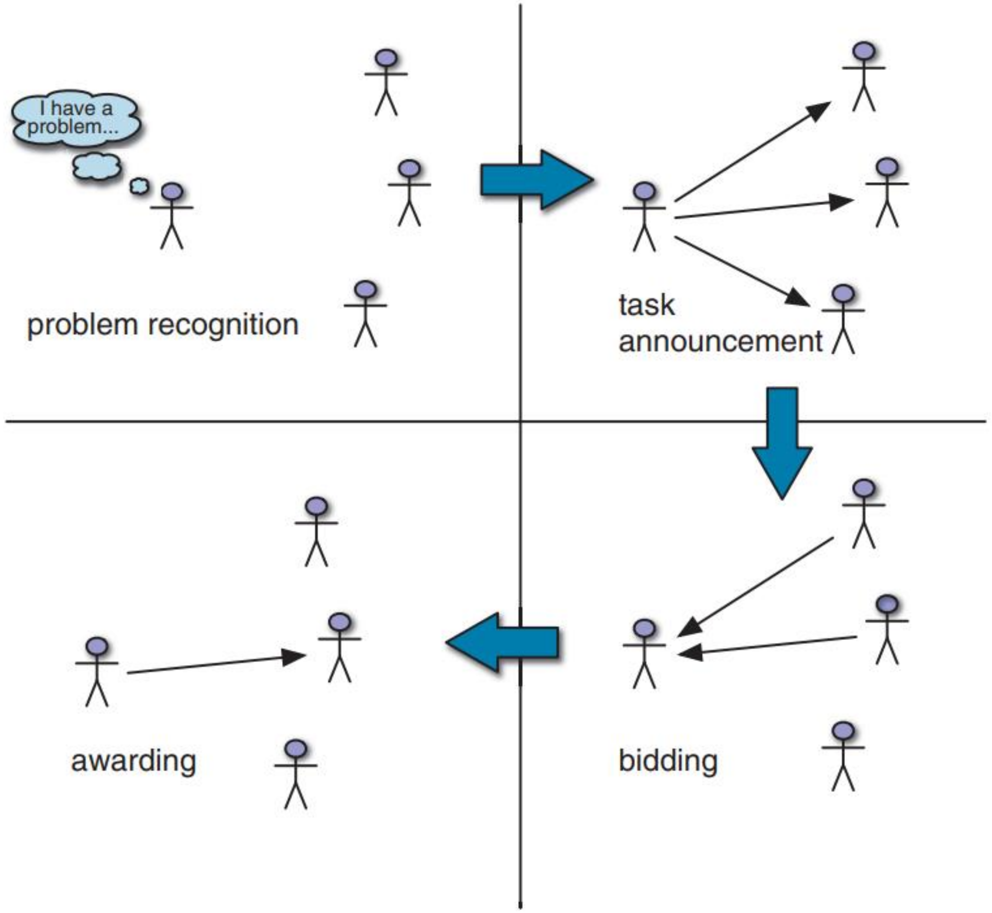
2. El agente anuncia la tarea y se convierte en el gerente. Este anuncio incluye detalles como plazos, requisitos de calidad y habilidades necesarias, similar a una oferta de trabajo.

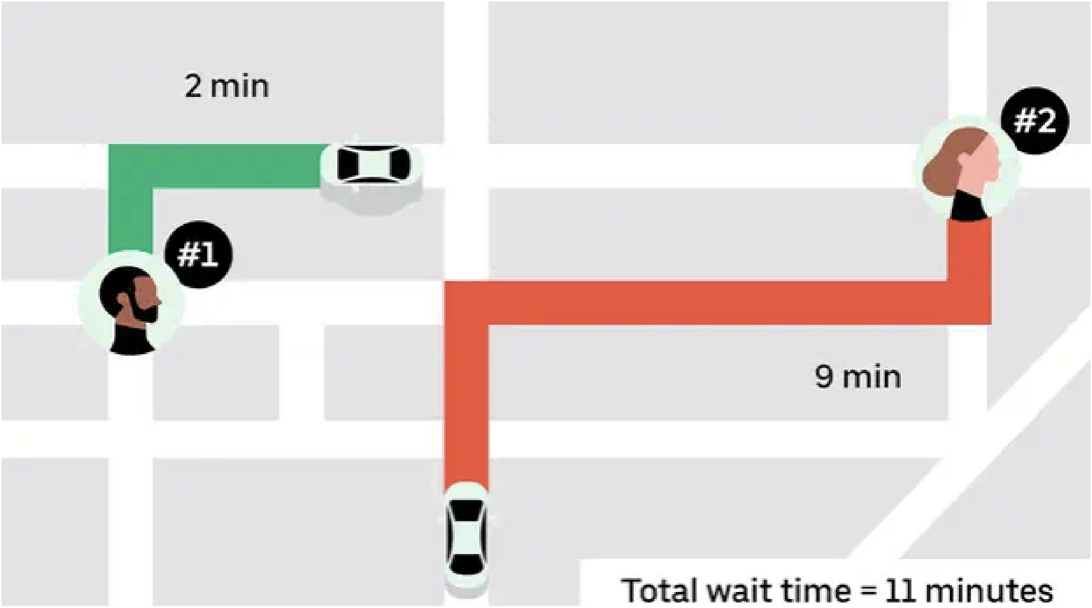
3. Los agentes que reciben el anuncio evalúan si desean postularse según sus capacidades y preferencias. Si están interesados, envían una solicitud que suele incluir información sobre sus habilidades relevantes y condiciones adicionales.

4. El gerente revisa las solicitudes y selecciona a los agentes más adecuados para llevar a cabo la tarea. Los agentes elegidos reciben una notificación y se convierten en contratistas hasta que se complete la tarea.

# Red de contrato

Tomado de: [Matching Market: Uber Pairing : Networks Course blog for INFO 2040/CS 2850/Econ 2040/SOC 2090 (cornell.edu)]




Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

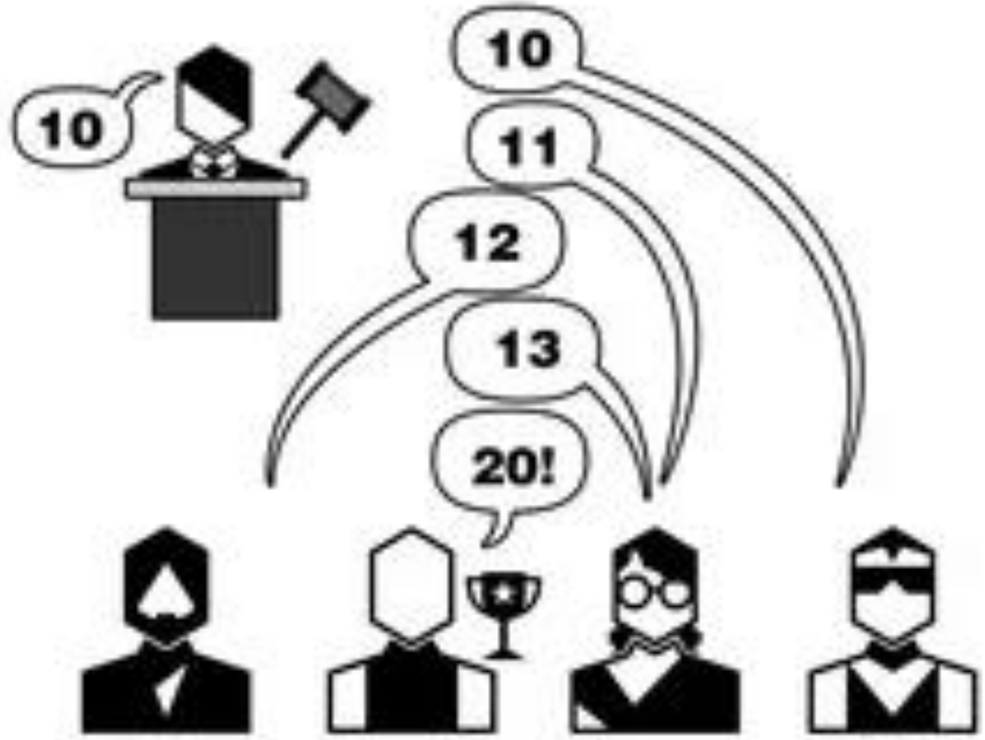
# Subastas

Las subastas son el mecanismo más importante para asignar recursos. La subasta más simple consiste en un recurso y un conjunto de postores. Cada postor tiene un valor de utilidad $v_i$ para el recurso. El recurso puede tener un valor privado para cada postor (valor sentimental, por ejemplo), o puede ser un valor común $X$ que es igual para todos pero que no es del todo conocido, y como tal termina siendo diferente para cada postor según la información que tenga. Un ejemplo es el de los tractos petroleros cuando se están subastando los derechos de perforación.

Dado el valor $v_i$ que tiene cada postor sobre el recurso, cada postor hace una oferta o puja $b_i$. La oferta más alta $b_{max}$ gana, pero el precio a pagar no tiene que ser $b_{max}$.

# Subasta inglesa o de oferta ascendente

Hay varios tipos de subasta y la más común es la inglesa o la subasta de oferta ascendente, la cual empieza con una oferta mínima definida por el centro y luego cada vez que un postor dice que está dispuesto a pagar esa cantidad, se sube a $b_{min} + d$ y así sucesivamente.



# La subasta es eficiente

La subasta es un mecanismo eficiente en el sentido de que se maximiza la utilidad del vendedor y también se maximiza la utilidad global porque el postor que más valora el recurso es el que más lo valora.

# Colusión

Una subasta debe evitar que los postores entren en colusión, que es un acuerdo ilegal entre dos o más postores para manipular precios. Un ejemplo es el de una subasta de espectro electromagnético que ocurrió en 1999 y dos compañías: Mannesmann y T-Mobile se aprovecharon de las reglas para terminar obteniendo el recurso por un valor muchísimo menor del que debía recibir el gobierno.



Referencia: [Low Price Equilibrium in Multi-Unit Auctions: The GSM Spectrum Auction in Germany (econstor.eu)]

# Subasta veraz

Para evitar la colusión, es importante que el juego de subasta tenga una estrategia dominante, de manera que los jugadores puedan hacer sus ofertas sin tener en cuenta el comportamiento de los otros. Este tipo de mecanismo se llama: a prueba de estrategia. Usualmente este tipo de mecanismo requiere que los postores revelen el su valor $v_i$ real. Ese tipo de subasta se llama: subasta veraz.

# La subasta inglesa es eficiente pero no veraz

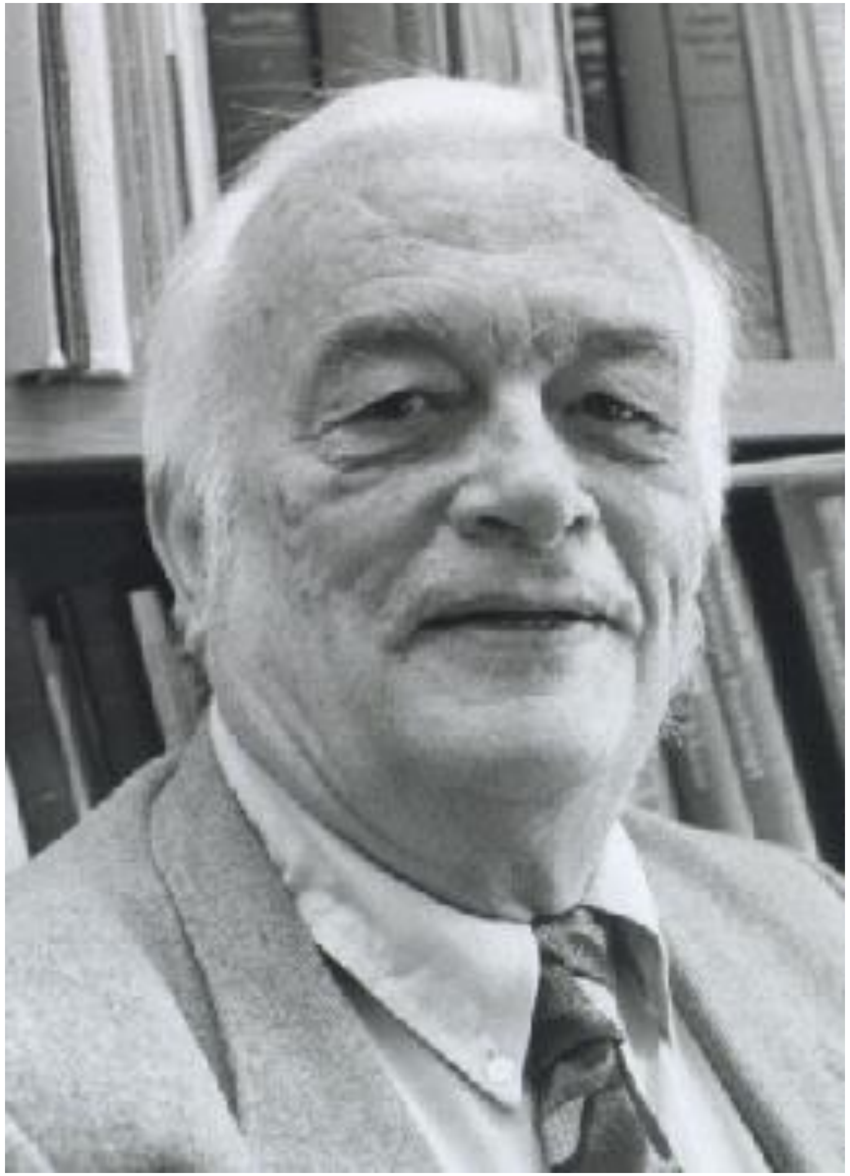
El postor que tiene el $v_i$ más alto es el que se lleva el recurso por un valor de $b_o + d$ donde $b_0$ es la oferta más alta y $d$ es el incremento del subastador. La estrategia dominante es: continuar ofertando siempre que $b_o + d \leq v_i$. El mecanismo no es completamente veraz porque lo único que saben los demás es que: $v_i \geq b_o + d$ pero no se revela el valor real $v_i$. Una desventaja de este tipo de subasta es que puede reducir la competencia haciendo que menos postores se unan o continúen en la subasta. **En términos de comunicación, es bastante costoso también.**

# Subasta de oferta sellada

Una alternativa es la subasta de oferta sellada. En este tipo de subasta cada postor hace su oferta y la comunica al subastador, pero nadie más las ve. En este caso, cada postor debe estimar el valor de utilidad de los demás competidores y asumir que el máximo será $b_0$. Luego oferta $b_0 + \epsilon < v_i$. El postor tiene que hacer más trabajo en este mecanismo. También es posible que la persona con el valor $v_i$ más alto no se lleve el recurso.

# Subasta de Vickrey

En este tipo de subasta, el ganador paga $b_0$ en lugar de pagar su propia oferta. La estrategia dominante en este caso es ofertar $v_i$, lo que lo hace un mecanismo veraz. Esto reduce completamente los costos computacionales y de comunicación y, además, se garantiza que quien se lleva el recurso es el que más lo valora, y que el subastador maximiza su utilidad. Este mecanismo se usa mucho en sistemas de inteligencia artificial distribuida.

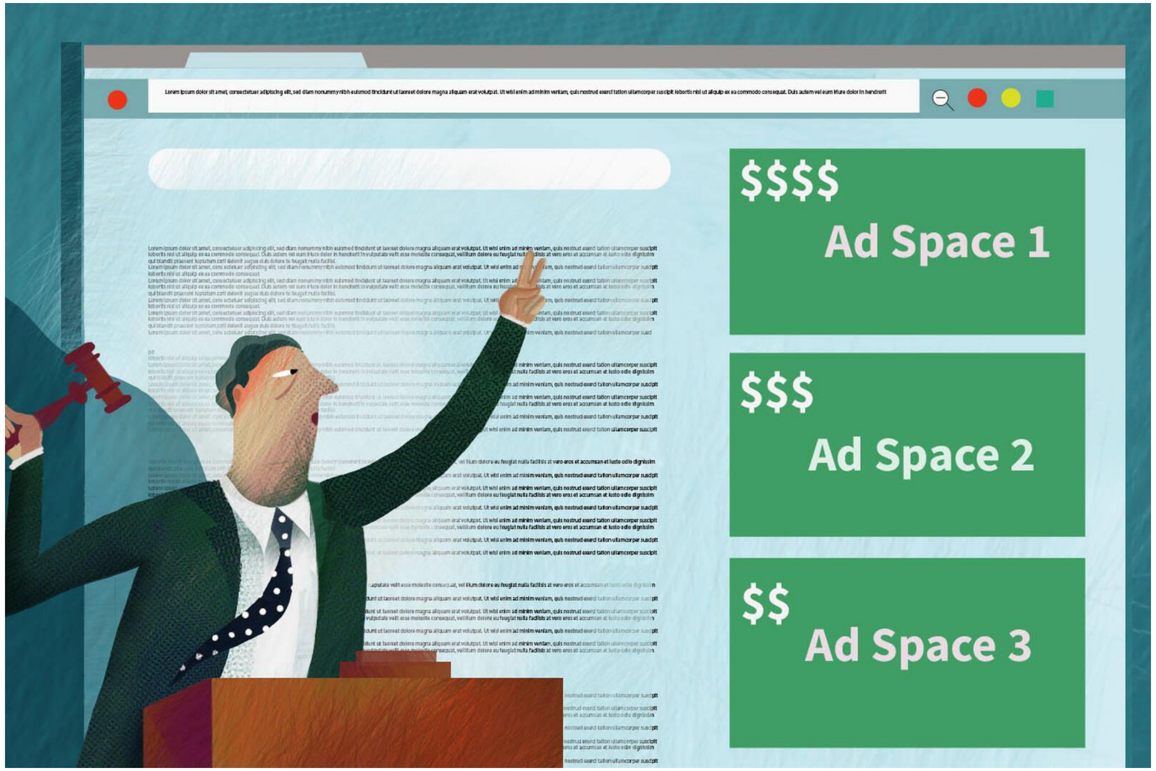


William Vickrey

# Motores de búsqueda

Los motores de búsqueda de internet hacen trillones de subastas cada año para vender anuncios y todos usan variantes de este mecanismo de subasta.

La subasta de Vickrey no es veraz cuando se subastan $n$ recursos con $n + 1$ postores. Por ejemplo, cuando se subastan lugares de anuncios en los motores de búsqueda. **Aggarwal et. al** propone una solución a este problema.



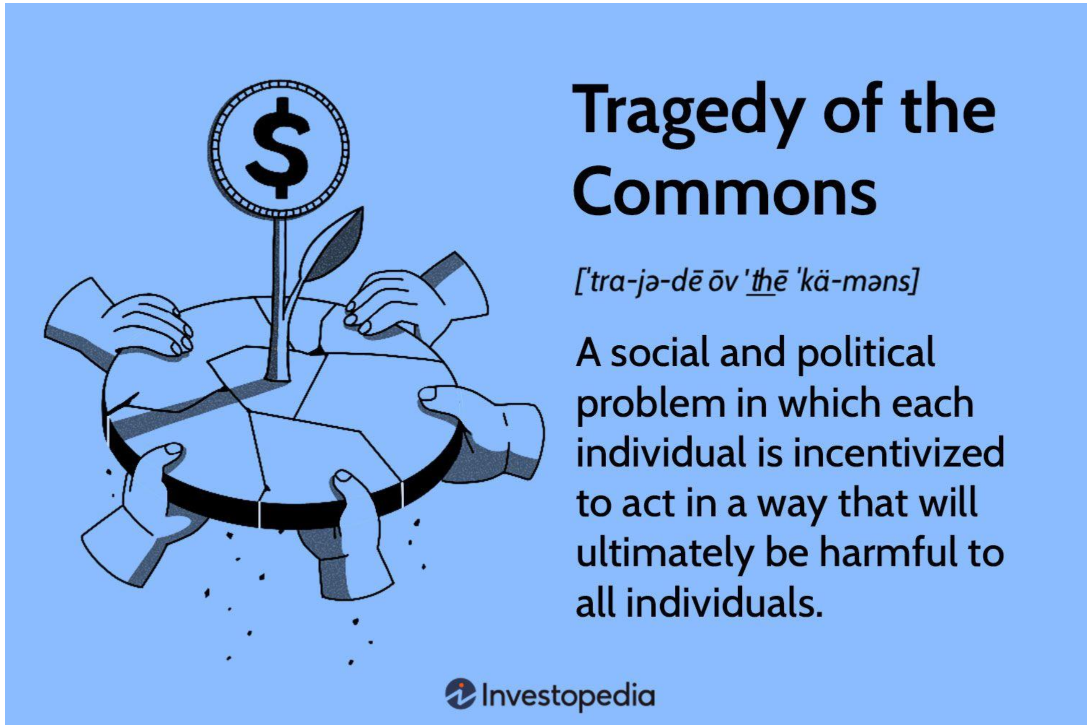

Tomado de: What Is "Auction Theory," and What Kinds of Questions Can It Answer? (northwestern.edu)

# La tragedia de los bienes comunales

Consideremos un juego en el que países tienen una política para controlar la contaminación del aire. Cada país tiene una elección: reducir la contaminación por un costo de $-10$ al implementar los cambios, o pueden seguir contaminando lo que les da un costo de $-5$ (por ejemplo, costos en salud) y además contribuye $-1$ a todos los otros países debido a que el aire es compartido.

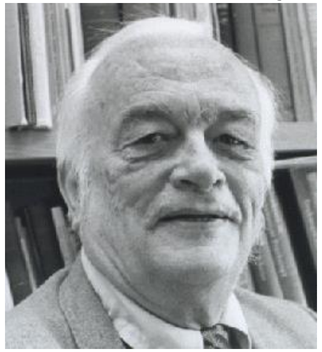
# La tragedia de los bienes comunes

La estrategia dominante es seguir contaminando, pero si hay 100 países cada país tiene una utilidad de $-104$. A esto se le conoce como la tragedia de los bienes comunales que dice que, si nadie tiene que pagar por usar un bien común, puede ser explotado de manera que les da menores utilidades a todos los agentes.



# Subasta Vickrey–Clarke–Groves (VCG)

William Vickrey



Edward H. Clarke

Theodore Groves

Un mecanismo que busca maximizar las utilidades es el VCG, que también es veraz, lo que significa que la estrategia dominante de los agentes es revelar su valor $v_i$. Puede explicarse de la siguiente manera: supongamos que una ciudad quiere instalar tres transceptores de internet inalámbrico gratuito en cinco barrios. La ciudad pregunta a cada barrio cuánto valora el transceptor.


Imágenes tomadas de: Edward H. Clarke ’62 | Princeton Alumni Weekly y Theodore Groves, Economics Dept, UCSD — Calisphere

# Subasta Vickrey–Clarke–Groves (VCG)

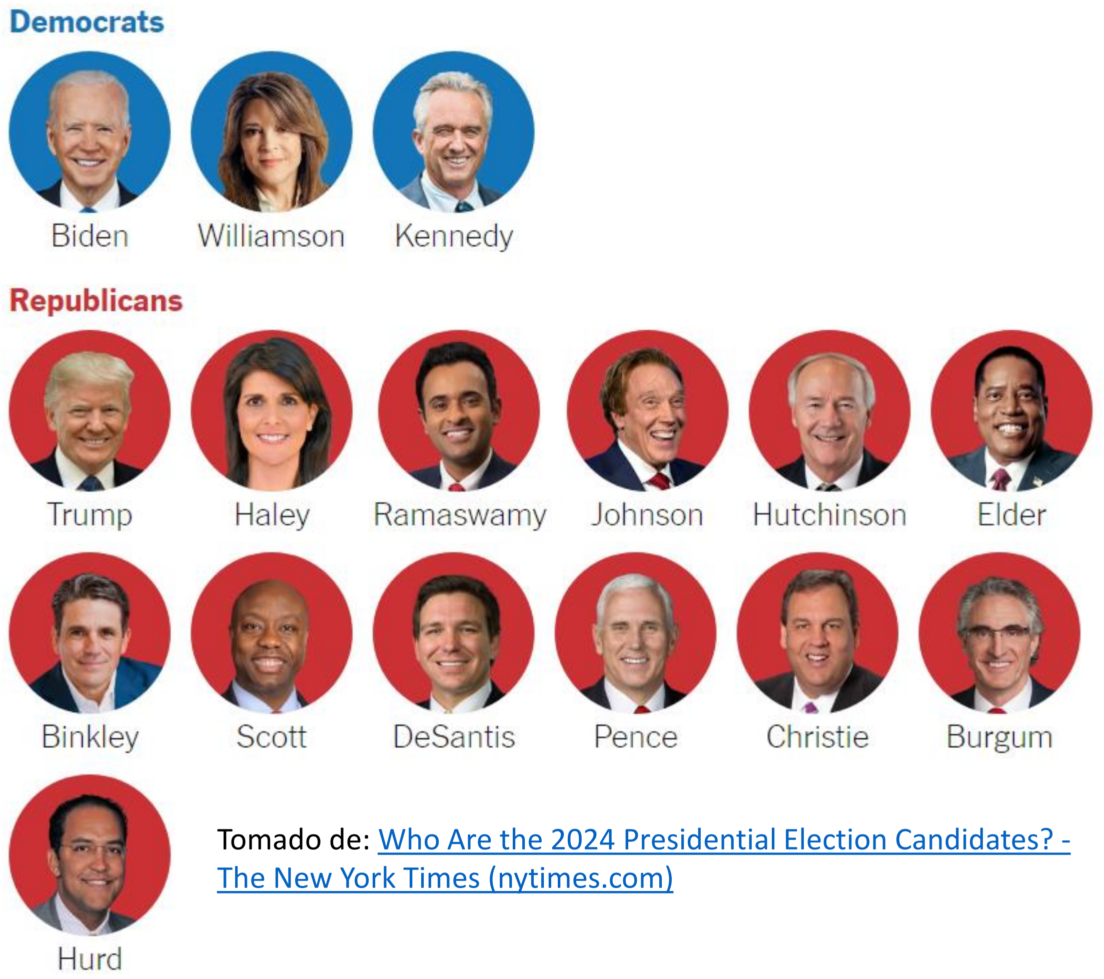
Para evitar que los barrios inflen sus valores, se les dice que los ganadores deben pagar un impuesto igual a la pérdida causada a los perdedores. Por ejemplo, si los ganadores tienen valores de 190, 100 y 50, y el cuarto tenía un valor de 20, entonces todos deben pagar un impuesto de 20 porque si alguno de los ganadores no hubiera ganado, el cuarto habría sido un ganador. Esto garantiza que ofertar más o menos que $v_i$ no es racional, ya que el valor del impuesto nunca está garantizado y podría ser mayor que el valor real.

# Votaciones

Las votaciones son mecanismos utilizados para la toma de decisiones política en sociedades democráticas. El estudio de las votaciones viene de la teoría de la elección social. La configuración es la siguiente. Se tiene un conjunto $N = \{1, 2, \dots, n\}$ de agentes que en esta sección llamaremos votantes. Estos votantes quieren tomar decisiones con respecto a un conjunto de posibles resultados $\Omega = \{\omega_1, \dots\}$. En una elección política cada elemento de $\Omega$ podría ser un candidato. Cada votante tiene preferencias sobre $\Omega$ y estas no son expresadas cuantitativamente sino cualitativamente, así:

$$\omega \succ_i \omega'$$

Lo que significa que el resultado $\omega$ tiene un ranking mayor para el votante $i$ que el candidato $\omega'$.




Tomado de: [Who Are the 2024 Presidential Election Candidates? - The New York Times (nytimes.com)]

# Orden de preferencia social

El problema fundamental de la teoría de la elección social es el de encontrar un orden de preferencia social, es decir, encontrar un orden cualitativo o ranking de esos candidatos desde el más preferido al menos preferido usando una función de bienestar social.

En muchos casos, lo que se quiere es un resultado social, que significa, el resultado más preferido de todo el grupo de votantes.

Para un orden de preferencia social escribimos:

$$\omega \succ^* \omega'$$

Para decir que $\omega$ es preferido sobre $\omega'$ en ese orden.

# Paradoja de Condorcet

Las sociedades democráticas quieren un resultado social que refleje las preferencias de los votantes, pero esto no siempre es sencillo. Un ejemplo es la paradoja de Condorcet, en el que se tienen tres candidatos y tres votantes. Cada votante establece un orden diferente de los candidatos. Sea cual sea el candidato que se escoja, 2/3 de los votantes (es decir, la mayoría), estará inconforme con el resultado.

Sean los candidatos $\Omega = \{\omega_a, \omega_b, \omega_c\}$ y los votantes $N = \{1, 2, 3\}$. Si las preferencias de los votantes son:

$$\omega_a \succ_1 \omega_b \succ_1 \omega_c$$
$$\omega_b \succ_3 \omega_c \succ_3 \omega_a$$
$$\omega_c \succ_2 \omega_a \succ_2 \omega_b$$

Obtenemos que, cualquiera sea el ganador: $2/3$ votantes prefieren $\omega_c$ sobre $\omega_a$, $2/3$ votantes prefieren $\omega_a$ sobre $\omega_b$ y $2/3$ votantes prefieren $\omega_b$ sobre $\omega_c$

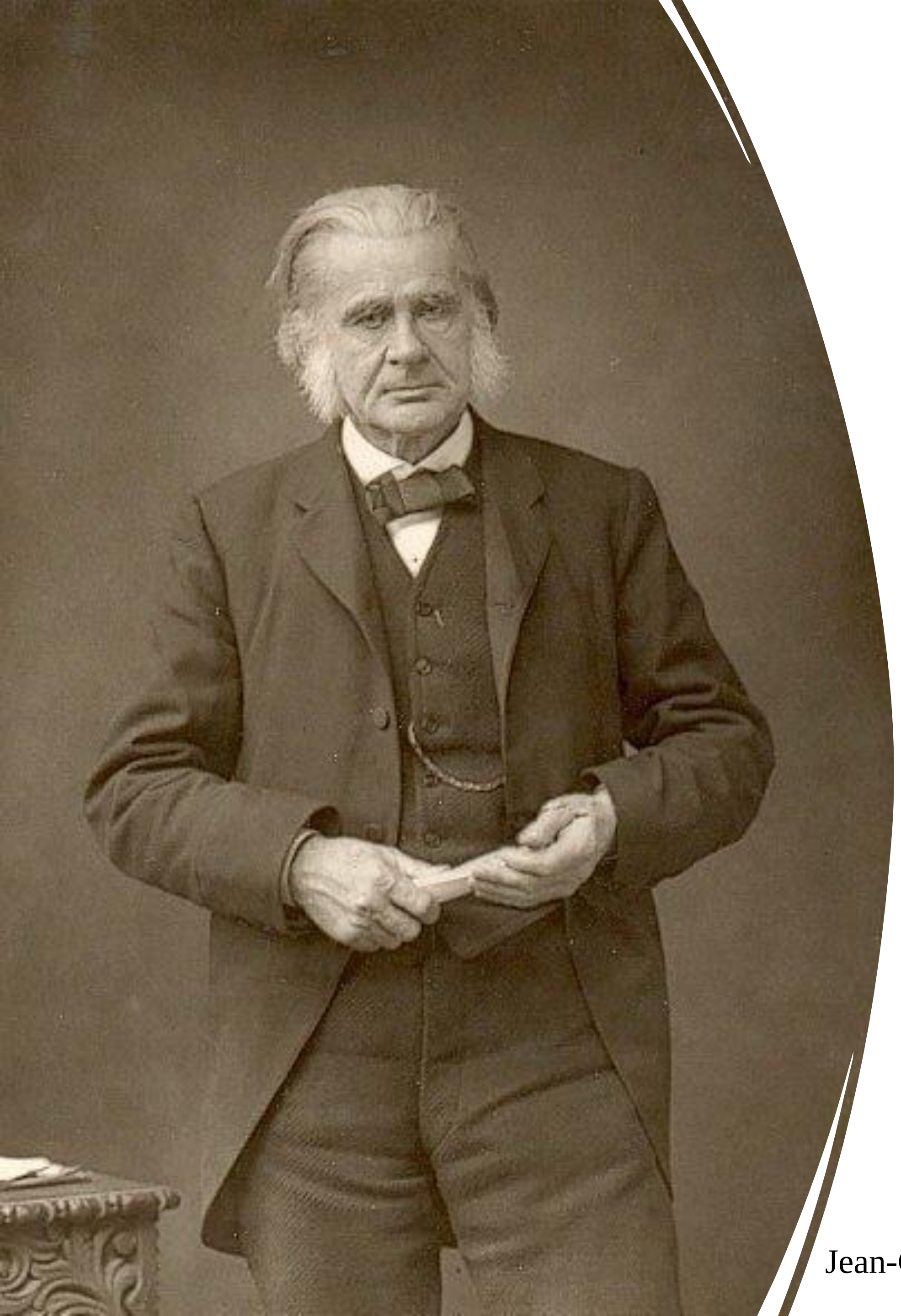
# Condiciones de una buena función de bienestar social

- *Condición de Pareto*: Si todos los votantes prefieren a $\omega$ sobre $\omega'$ entonces $\omega \succ^* \omega'$
- *Condición de ganador de Condorcet*: Un candidato es ganador de Condorcet si la mayoría de los votantes lo prefieren sobre todos los demás candidatos. Es decir, el ganador de Condorcet es aquel que ganaría a todos los demás en competiciones dos a dos.
- *Independencia de alternativas irrelevantes*: Si un votante cambia el orden de sus preferencias, pero $\omega \succ_i \omega'$ se mantiene, entonces $\omega \succ^* \omega'$ debe mantenerse.
- *No dictaduras*: Una función de bienestar social no puede dar como resultado el orden de preferencia de un votante e ignorar todos los otros.

# Teorema de imposibilidad de Arrow

Kenneth Arrow

Aunque estas cuatro condiciones suenan razonables, el teorema de Arrow dice que es imposible satisfacer todas las cuatro condiciones a la vez cuando se tienen más de dos candidatos.



Jean-Charles de Borda

# Procedimientos de votación

- *Mayoría simple*: Para dos candidatos, se pregunta a cada votante cuál de los dos candidatos prefieren y se escoge el candidato con más votos.
- *Votación plural*: Para más de dos candidatos, se pregunta a cada votante cuál es el candidato que más prefieren y se selecciona al candidato o los candidatos (si hay un empate) con más votos.
- *Cuenta de Borda*: Es un sistema que toma en cuenta toda la información del ordenamiento de preferencias del votante. Por ejemplo, si se tienen $k$ candidatos, en el ordenamiento de cada votante el candidato que más prefiere tendrá un puntaje de $k$ el segundo $k-1$ y así sucesivamente. Luego se suman todos los puntajes para todos los candidatos y así se establece el orden social para elegir el ganador o los ganadores.


# Procedimientos de votación

- *Votación de aprobación*: Los votantes envían un subconjunto de los candidatos que ellos aprueban. Los ganadores son los candidatos que más veces fueron aprobados por los votantes.
- *Segunda vuelta instantánea*: Los votantes ordenan todos sus candidatos. Si un candidato tiene la mayoría de los votos en primer lugar, se declara ganador. Si no, el candidato con menor número de votos en primer lugar es eliminado, y se repite el proceso.
- *Verdadera mayoría*: El ganador es el candidato que vence a todos los demás candidatos en competiciones por pares. Se dice que un candidato $\omega$ vence a un candidato $\omega'$ en este sistema si más votantes establecieron $\omega \succ \omega'$ que $\omega' \succ \omega$. Lo bueno de este sistema es que la mayoría siempre estará conforme, pero el problema es que no siempre se puede decidir una votación con este sistema, por ejemplo, si ningún candidato gana una mayoría.

# La votación no es un procedimiento veraz

Las votaciones a menudo no garantizan que los votantes expresen sus preferencias de manera veraz. Por ejemplo, en las votaciones plurales, los votantes pueden cambiar sus preferencias para beneficiar a un candidato que tenga más posibilidades de ganar, en lugar de votar por su candidato favorito.

Imagen tomada de: [Sergio Fajardo obtuvo menos votos que en 2018 y se quedó sin segunda vuelta - Infobae]

# Teorema de Gibbard–Satterthwaite

Cualquier función de elección social que satisfaga la condición de Pareto para dos o más candidatos, es manipulable o una dictadura

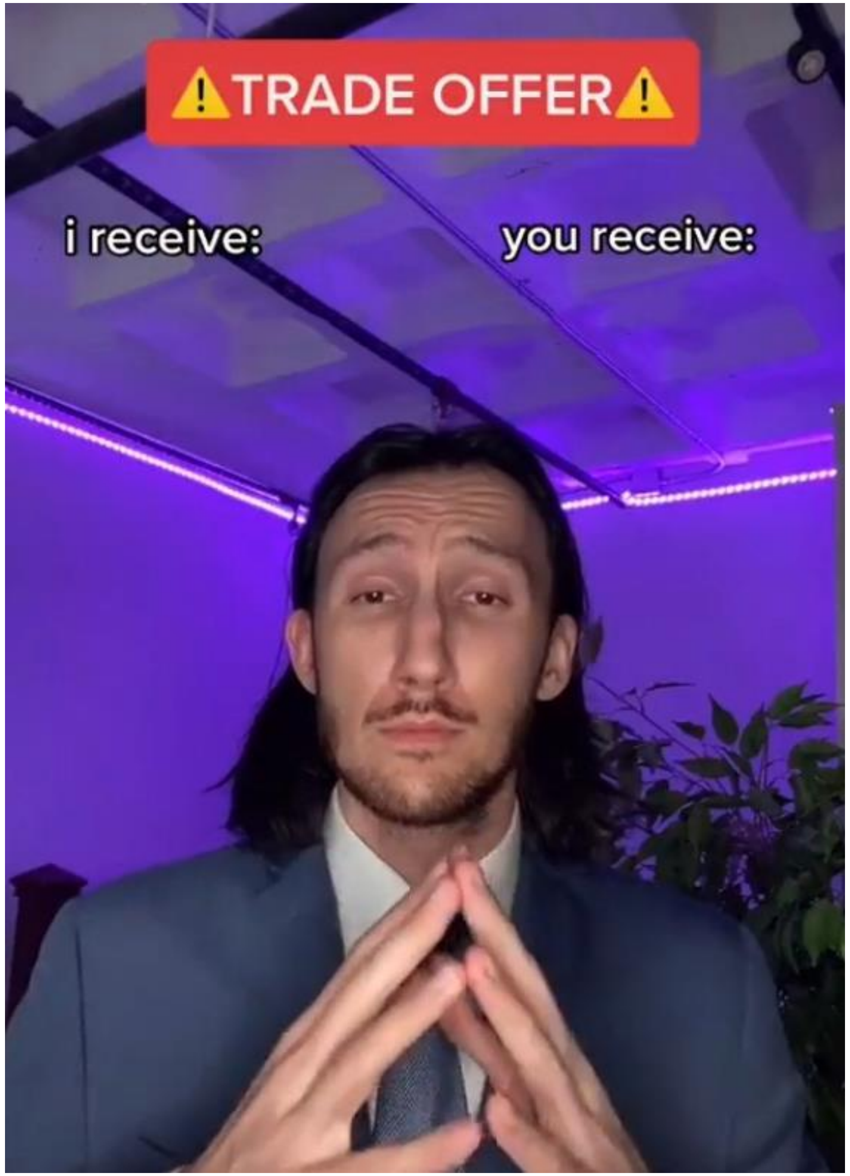

Tomado de: [Gibbard's Theorem vs Stable Matching | by Chris Smith | Medium]


# Negociación

La negociación es una herramienta que se utiliza cuando los agentes necesitan llegar a un acuerdo sobre un tema de interés común. En este proceso, los agentes hacen ofertas o propuestas entre sí siguiendo protocolos específicos y toman decisiones sobre si aceptar o rechazar dichas ofertas.

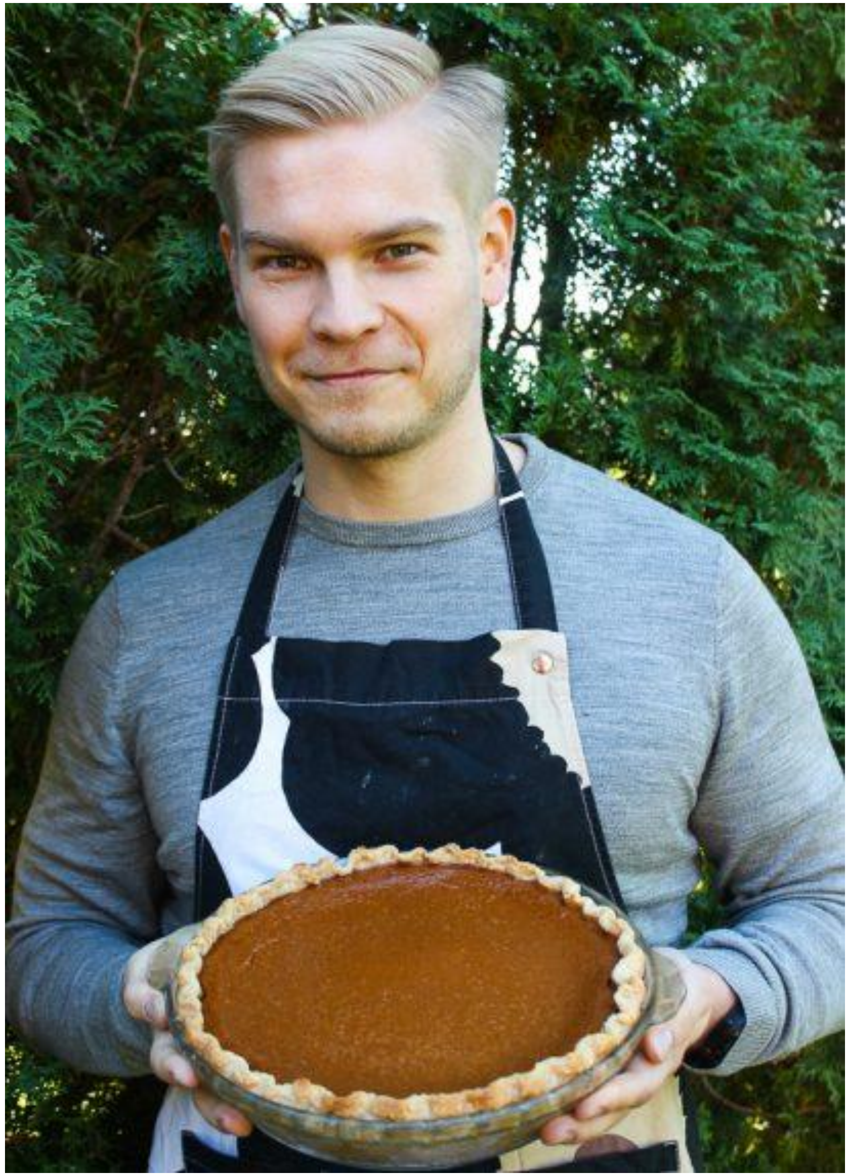
# Protocolo de ofertas alternantes

Supongamos que tenemos dos agentes $A_1, A_2$. La negociación ocurre en una serie de rondas. En la ronda 0, $A_1$ hace una oferta. Si $A_2$ acepta, se acaba la negociación, si $A_2$ rechaza se sigue a la siguiente ronda. En la siguiente ronda $A_2$ hace una oferta y ahora $A_1$ decide si acepta o rechaza y así sucesivamente. El trato conflicto es cuando siempre se rechaza la oferta y los agentes siguen negociando por siempre. Se asume que los agentes desean evitar esto y quieren terminar la negociación en un tiempo finito, por definición.



# Dividir el pastel en una ronda

Por esa razón, si la negociación es dividir un pastel, una oferta sería $(x, 1 - x)$. Digamos que el juego tiene una ronda. En ese caso el primer jugador hace la oferta $(1, 0)$. Como esta es la última ronda, el segundo jugador debe aceptar, de otro modo, se entraría en el trato conflicto. Como el primer jugador no puede obtener nada mejor que 1 y como el segundo jugador prefiere obtener 0 a entrar en el trato conflicto, este es un equilibrio de Nash.



# Dividir el pastel en un número finito de rondas

Si el juego tuviera dos rondas, el segundo jugador rechazaría en la primera y en la segunda ronda haría la oferta $(1, 0)$. Nuevamente, este es un equilibrio de Nash. Pero, lo interesante, es que cuando el número de rondas es fijo, cualquiera que tenga su movimiento de oferta de últimas tiene todo el poder, y se lleva todo el pastel.

# Dividir el pastel en un número infinito de rondas

Cuando no hay un límite fijo de rondas, y el primer agente tiene como estrategia ofrecer $(1, 0)$ y rechazar cualquier otra oferta, lo mejor que puede hacer el segundo jugador es aceptar para evitar seguir negociando por siempre.

Pero eso aplica para cualquier oferta que haga el jugador 1 dentro del conjunto de negociación.

Por tanto, esto nos dice que hay un número infinito de equilibrios de Nash.

# Los agentes son impacientes


Para cualquier resultado $x$ y rondas $t_1, t_2$ con $t_1 < t_2$ ambos agentes prefieren un resultado $x$ en la ronda $t_1$ y no en la ronda $t_2$.

# Factor de paciencia

Para evitarlo, se puede incluir un factor de descuento $\gamma_i$ para cada agente con un valor entre 0 y 1. El valor de la tajada del pastel en la ronda $t$ para el agente $i$ es:

$$\gamma_i^t x$$

En la ronda 0 es $\gamma_i^0 x = x$ y cada vez ese valor irá disminuyendo con el paso de las rondas, porque el valor del factor es entre 0 y 1. Entre más grande sea este valor del factor para un agente, significa que tendrá más paciencia. Con este cambio, en un juego de dos rondas, el valor del pastel para el segundo jugador será menor, de manera que el primer jugador puede aprovechar eso para ofrecer en la primera ronda $(1 - \gamma_2, \gamma_2)$ y este sería un equilibrio de Nash. Los agentes con más paciencia obtendrán mayores porciones.

# Negociación de tareas

Supongamos que hay dos agentes que deben hacer un conjunto de tareas en dos máquinas diferentes. Sin embargo, configurar las máquinas es costoso. De manera que los agentes pueden llegar a un acuerdo de que cada uno usará solo una máquina y hará las tareas del otro que necesiten esa máquina. Así ambos se ahorran una parte del costo de la configuración.

Formalmente el conjunto de tareas es $T$ y la asignación inicial es $(T_1^0, T_2^0)$. El costo de un conjunto de tareas es $c(T_i)$ y el costo de no hacer nada es cero: $c(\emptyset) = 0$. La función costo es monotónica, de manera que, a más tareas, más costo.

Una oferta $(T_1, T_2)$ tiene utilidad para el agente $i$ si:

$$U\big((T_1, T_2)\big) = c\big(T_i^0\big) - c(T_i) \geq 0$$

# Protocolo de concesión monotónica

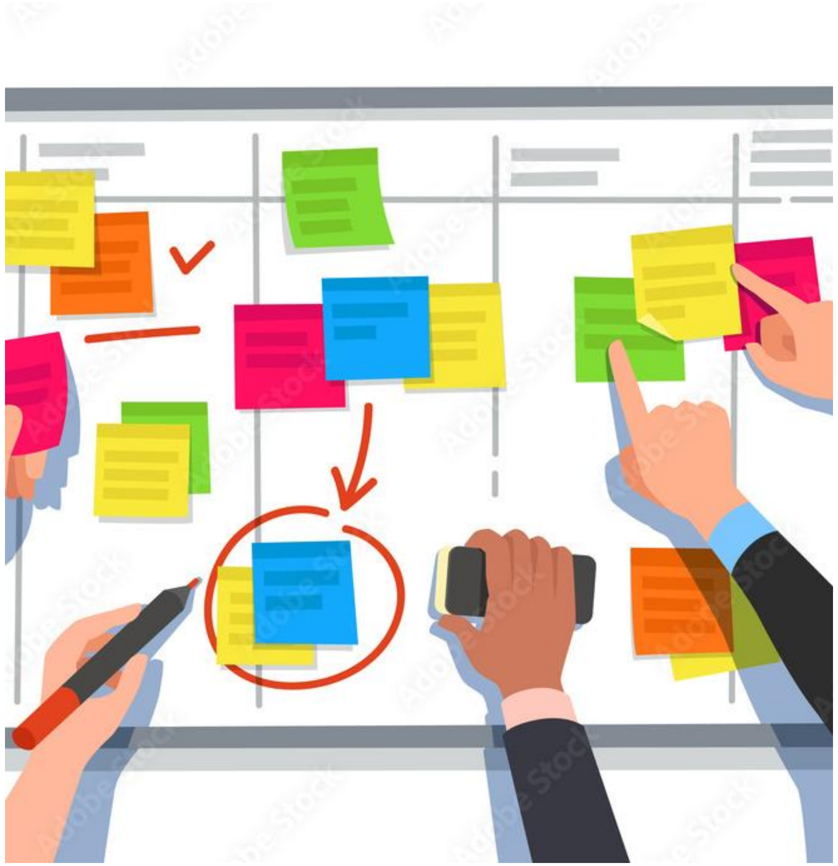
Las reglas son las siguientes:

- La negociación se hace en rondas.
- En la primera ronda, ambos agentes hacen una oferta simultáneamente: $D_i = (T_1, T_2)$
- Un acuerdo se alcanza si los dos agentes proponen sus ofertas de manera que:

$$U_1(D_2) \geq U(D_1) \ o \ U_2(D_1) \geq U(D_2)$$

Es decir, si alguno de los agentes encuentra que la propuesta de su compañero es igual o mejor que la suya. Si ambas ofertas son iguales o mejores que las otras, se selecciona una al azar.

Si no hay acuerdo, la negociación continúa hacia una siguiente ronda de propuestas simultaneas. En la ronda $t + 1$ cada agente debe repetir su oferta o hacer una concesión, es decir, una propuesta que es preferida por el otro agente. Si ningún agente hace una concesión, la negociación termina y ambos agentes implementan el trato de conflicto que en este caso es hacer las tareas que se les asignaron inicialmente.



# Estrategia Zeuthen y voluntad de riesgo

Busca medir la voluntad que tienen los agentes de arriesgar el conflicto. Intuitivamente un agente estará más dispuesto a arriesgar conflicto si la diferencia entre la utilidad de la propuesta actual y la del trato de conflicto es baja, pues tiene poco que perder. Si la diferencia es alta, el agente tiene más que perder y estaría menos dispuesto a arriesgar el conflicto y más dispuesto a conceder (es decir hacer concesiones).

# Cálculo de la voluntad de arriesgar

La voluntad de arriesgar el conflicto de un agente $i$ en una ronda $t$ se define como la utilidad que perdería al conceder y aceptar la oferta sobre la utilidad que perdería al no conceder y causar conflicto (es decir hacer las tareas que se le asignaron inicialmente). Un valor más cercano a 1 significa que hay menos que perder y como tal se estará más dispuesto a arriesgar.

$$risk_i^t = \frac{\text{utility } i \text{ loses by conceding and accepting } j\text{'s offer}}{\text{utility } i \text{ loses by not conceding and causing conflict}}.$$

Tomado de: [Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)]

# La estrategia Zeuthen

La estrategia dice que la primera propuesta de los agentes debe ser aquella que maximiza su propia utilidad. Luego de eso, el agente que debe conceder en la ronda $t$ es aquel que tenga menor de voluntad de riesgo, es decir aquel que tiene más que perder si se llega al conflicto. La cantidad a conceder debe ser la más pequeña posible que haga que el otro jugador conceda en la siguiente ronda.

Si ambos agentes tienen el mismo riesgo, entonces se debe lanzar una moneda para decidir quién concede, haciendo que se logre un equilibrio de Nash.

Frederik Zeuthen

Tomado de: A Personal note by Karl Vind – University of Copenhagen (ku.dk)

# Referencias


- Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu)
- Investopedia
- Partially ordered set - Wikipedia
- Social Choice Theory (Stanford Encyclopedia of Philosophy)
- Marginal contribution nets | Proceedings of the 6th ACM conference on Electronic commerce
- Low Price Equilibrium in Multi-Unit Auctions: The GSM Spectrum Auction in Germany (econstor.eu)