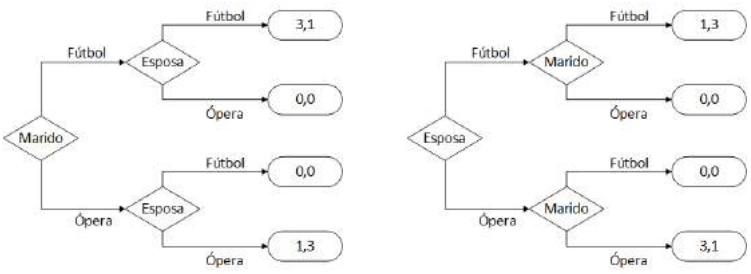
Tema 30

Juegos no cooperativos con información completa dinámicos

30.1. Representación de juegos en forma extensiva

La forma extensiva de un juego es una representación gráfica que adquiere forma de árbol. Comienza con un nodo que representa el comienzo del juego y de él salen varias ramas con las posibles elecciones del jugador que toma una decisión. Estas nuevas ramas terminan en un nodo de decisión que da lugar a nuevas ramas o en un nodo final etiquetado con los pagos de los jugadores (en el mismo orden que mueven). También se pueden introducir nodos de oportunidad en los que las ramas representan varios escenarios posibles que dependen de factores ajenos al juego (cuando los jugadores eligen saben cuál es el escenario en el que se encuentran).

Ejemplo 30.1 *Si en la batalla de los sexos suponemos que los esposos deciden ir a la opera o al fútbol de forma secuencial y uno decide primero tenemos (el primer pago del par corresponde al que elige primero):*



Si el marido decide primero y propone ir al fútbol su mujer puede optar por acompañarle o no. Como prefiere estar con él a irse sola, decidirá acompañarle. Si la esposa decide primero y propone ir a la ópera, su marido responderá de la misma forma e irá a la ópera. De esta forma, si la decisión es secuencial el que decide primero tiene ventaja y solo es posible un resultado. Si la decisión es simultánea las dos soluciones son posibles a este juego y los dos van juntos al fútbol o los dos van juntos a la ópera (hay otro equilibrio en estrategias mixtas con pagos iguales para ambos). ♣

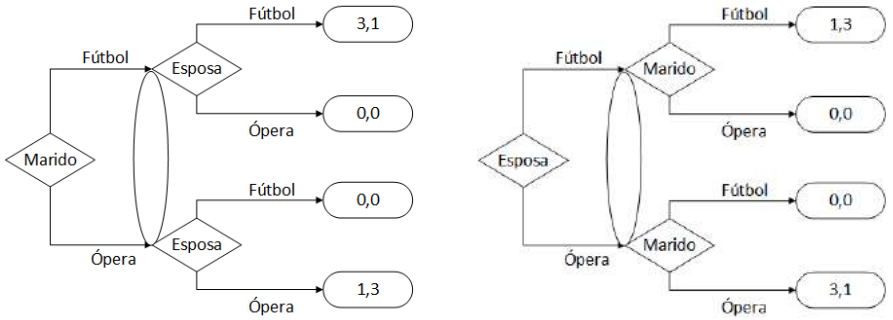
Definición 30.2 *La forma extensiva de un juego es su representación en forma de árbol (árbol del juego).*

- *Comienza con el nodo raíz (representa el comienzo del juego)*
- *Del nodo raíz salen varias ramas (representan las opciones del jugador que toma una decisión)*
- *Cada rama termina en un nodo de uno de estos dos tipos:*
  - *nodo terminal etiquetado con los pagos (corresponde al final del juego).*
  - *nodo de decisión (un jugador hace una elección que da lugar a nuevas ramas).* ♣

La forma extensiva del juego muestra perfectamente quiénes juegan, cómo juegan, cuándo juegan y cuánto ganan o pierden. Así, un jugador es el que tiene un nodo de decisión en el juego y las posibles elecciones disponibles en ese momento del juego son las ramas que parten del nodo de decisión (un nodo de decisión sólo se alcanza cuando las elecciones previas se han realizado).

Para representar juegos con movimientos simultáneos de los jugadores se considera que un jugador actúa igual si elige simultáneamente a sus oponentes que si elige después pero desconoce la elección de éstos. Así, sustituiremos los nodos en los que un jugador no puede distinguir la elección realizada por un óvalo que los envuelve, que recibe el nombre de **conjunto de información**.

Ejemplo 30.3 *La batalla de los sexos clásica se representa en forma extensiva de dos formas equivalentes en las que al hacer su elección no saben que ha hecho su pareja (equivale a una elección simultánea):*



♣

Para que un árbol represente un juego se tienen que ser consistente y no puede haber ambigüedad acerca de cómo se desarrolla el juego. En particular, para saber dónde comienza el juego debe haber sólo un nodo raíz, para que no lleguemos a un punto muerto durante el juego las ramas del árbol no pueden crear un ciclo y para que solo haya una forma de procede distintas ramas no pueden conducir a un mismo nodo.

Nota Para garantizar la consistencia del árbol del juego se imponen las siguientes restricciones sobre los predecesores de un nodo, que son los nodos desde los que se puede llegar a él a través de sus ramas:

- Un nodo no puede ser un predecesor de si mismo.
- Si un nodo A es un predecesor de un nodo B y B es un predecesor del nodo C entonces A es también un predecesor de C (el predecesor de un predecesor de un nodo es también su predecesor).
- S dos nodos A y B son predecesores de un nodo C o bien A es un predecesor de B o bien B es un predecesor de As (los nodos pueden ordenarse).
- Si dos nodos A y B no son predecesores uno del otro debe haber un tercer nodo C que preceda a ambos (antecesor común). ♣

Al igual que en un juego en forma estratégica, en un juego en forma extensiva una estrategia es una regla de decisión que le dice a un jugador cómo actuar en cada momento. En cada nodo de decisión o, en su caso, en cada conjunto de información, las estrategias dan lugar a una acción. Si tenemos una y sólo una estrategia para cada jugador podemos determinar como se desarrollará el juego exactamente. De este modo, la estrategia de un jugador es un plan condicional y completo de acción: en todos y cada uno de los nodo de decisión del árbol le dice lo que debe decidir si el juego llega al nodo de decisión (que rama seguir). Una distribución de probabilidad sobre estas estrategias da lugar a una estrategia mixta.

Nota La *forma estratégica de un juego en forma extensiva* se obtiene enumerando la lista de jugadores, sus estrategias respectivas y los pagos asociados a cada perfil de estrategias. ♣

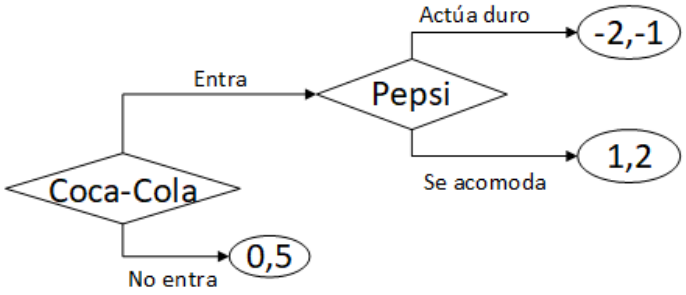
30.2. Juegos con información perfecta

Definición 30.4 *Los juegos de información perfecta (perfect information games) son juegos en forma extensiva con exactamente un nodo en cada conjunto de información.* ♣

Nota En un juegos con información perfecta no existe ambigüedad y cada vez que un jugador juega conoce exactamente la historia del juego (no hay movimientos simultáneos). ♣

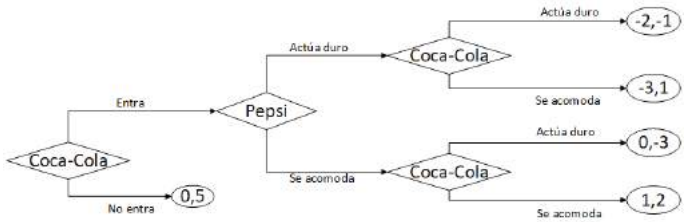
Ejemplo 30.5 *Juego de entrada I*

Coca-Cola tiene que decidir si entra o no en un nuevo mercado dominado por Pepsi. La rentabilidad del mercado depende de la reacción de Pepsi. Si Pepsi contraataca y actúa duro ambas tienen pérdidas. Por el contrario, si Pepsi no contraataca y se acomoda tienen beneficios. Los pagos que obtienen aparecen en la forma extensiva del juego, donde el primer pago de cada par es el de Coca-Cola (actúa primero).



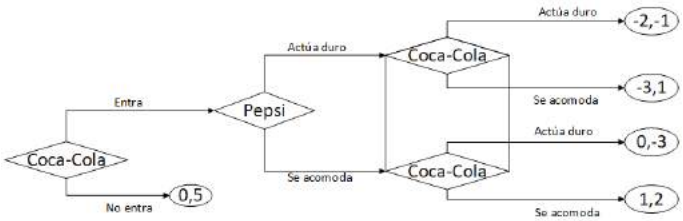
Ejemplo 30.6 *Juego de entrada II*

Coca-Cola tiene que decidir si reacciona contraatacando y actúa duro o no lo hace y se acomoda después de observar la respuesta de Pepsi.



Ejemplo 30.7 *Juego de entrada III*

Coca-Cola tiene que decidir si reacciona contraatacando y actúa duro o no lo hace y se acomoda sin conocer la reacción de Pepsi *(no es un juego con información perfecta).*



En un juego en forma extensiva una predicción razonable sobre el juego tiene que incluir la racionalidad secuencial. La racionalidad implica que un jugador escogerá la mejor respuesta a su disposición en un nodo de decisión teniendo en cuenta su percepción sobre el futuro del juego y la secuencialidad implica que en el futuro los jugadores razonarán de la misma manera y en los nodos posteriores escogerán la mejor respuesta teniendo en cuenta su percepción sobre lo que resta del futuro del juego.

Ejemplo 30.8 *Juego de entrada I (continuación del ejemplo 30.5)*

Comprobar, utilizando la forma estratégica del juego, que el juego de entrada I tiene como equilibrios de Nash (E, A) *y* (O, T) *pero el único equilibrio razonable es* (E, A)*, en el que Coca-Cola entra en el mercado y Pepsi no contraataca (denotamos E=enter, O=out, T=tough y A=accommodate).* ♣

Ejemplo 30.9 *Juego de entrada II (continuación del ejemplo 30.6) En este juego cada estrategia de Coca-Cola tiene que tener tres componentes: la primera le dice a Coca-Cola si entrar o no en el mercado, la segunda que hacer si Pepsi actúa duro y la tercera que hacer si Pepsi se acomoda.*

Hay esencialmente tres equilibrios de Nash en estrategias puras:

- Equilibrios de Nash en los que Pepsi actúa duro (T) y Coca-Cola juega cualquiera de las cuatro estrategias en las que se mantiene fuera ($O.T.T$, $O.T.A$, $O.A.T$ y $O.A.A$)
- (ETA, A) con el resultado de que Coca-Cola entra y se acomodan ambas empresas.
- (EAA, A) con el mismo resultado que en el segundo equilibrio.

Coca-Cola ↓ \ Pepsi →	T	A
ETT	-2, -1	0, -3
ETA	-2, -1	1, 2
EAT	-3, 1	0, -3
EAA	-3, 1	1, 2
OTT	0, 5	0, 5
OTA	0, 5	0, 5
OAT	0, 5	0, 5
OAA	0, 5	0, 5

La única estrategia secuencialmente racional para Coca-Cola es $E.TA$ y, por tanto, el segundo es el único equilibrio de Nash secuencialmente racional. ♣

En un juego con información perfecta en un nodo terminal la única predicción razonable es aquella en la que el jugador toma la acción que maximiza sus ganancias, ya que esta decisión termina el juego. En el penúltimo nodo de decisión el jugador conoce la consecuencia exacta de cada una de sus elecciones porque sabe la decisión que se tomará en el nodo terminal y puede calcular el pago exacto de cada una de sus decisiones y elegir la mejor opción. Análogamente, en el nodo anterior el jugador que toma la decisión conoce la consecuencia exacta de sus elecciones, ya que conoce las elecciones que se tomarán en el penúltimo nodo de decisión y en el nodo terminal. La **inducción hacia atrás** es la repetición de este proceso hasta llegar al nodo inicial y permite obtener una solución secuencialmente razonable siempre y cuando haya un último nodo desde el que empezar.

Proposición 30.10 *(Teorema de Kuhn) Todo juego de información perfecta con un número finito de nodos tiene solución por inducción hacia atrás y es única si todos los pagos de cada jugador son distintos.* ♣

La inducción hacia atrás y la eliminación de estrategias dominadas están íntimamente relacionadas, ya que si dos estrategias son idénticos en todo excepto en la decisión en un nodo terminal, la estrategia con la mejor decisión en este nodo domina la estrategia alternativa y cualquier otra estrategia que no tome esta decisión en ese nodo. Si dos estrategias son idénticos en todas partes excepto en un penúltimo nodo de decisión en el que hay una mejor decisión, ya que sabemos cuál va a ser la decisión en el nodo posterior. En

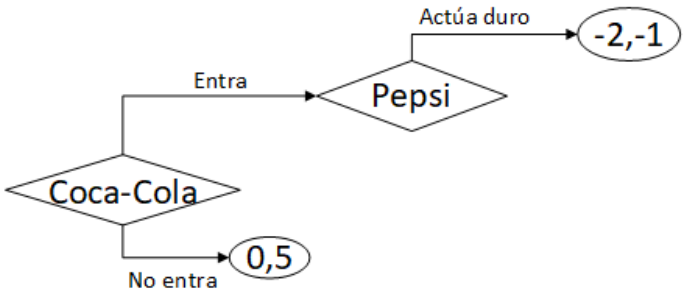
este caso, la estrategia que contiene la mejor decisión en este nodo domina a las otras (todas las estrategias son eliminadas excepto la que tiene la mejor decisión en el penúltimo nodo). Al continuar el proceso obtenemos la solución por inducción hacia atrás en la forma extensiva y la solución por eliminación iterada de estrategias dominadas en la forma estratégica.

Ejercicio 30.11 *Comprobar que el resultado por IEDS es el mismo que obtuvimos por inducción hacia atrás en los juego de entrada I y II (ejemplos 30.5 y 30.8 y ejemplos 30.6 y 30.9 respectivamente).*

Nota (El poder del compromiso) Aunque parece que es peor tener menos opciones, hay juegos en los que tener menos opciones puede suponer tener más pagos. Si un jugador tiene más opciones se puede comportar de manera muy diferente en el futuro. Este comportamiento afecta al juego actual y a los demás jugadores y puede ser beneficioso o perjudicial para el jugador con mayores opciones. ♣

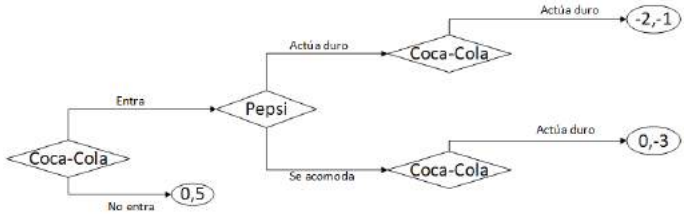
Ejemplo 30.12 *Juego de entrada I' (variante del juego de entrada I del ejemplo 30.5).*

Si en el juego de entrada I reducimos las opciones de Pepsi después de la entrada de Coca-Cola eliminando acomodar, Pepsi sólo tiene la opción de competir y jugar duro.



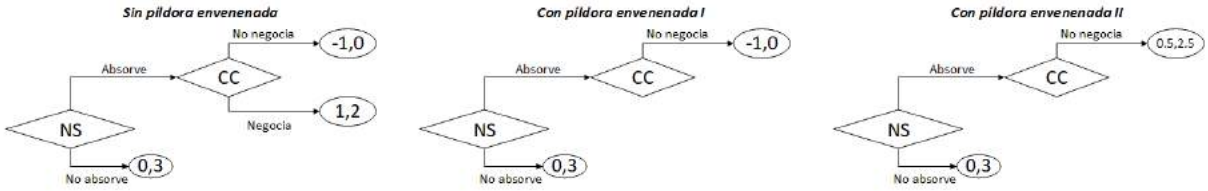
Ejemplo 30.13 *Juego de entrada II' (variante del juego de entrada II del ejemplo 30.6).*

Si en el juego de entrada II reducimos las opciones de Coca-Cola después de su entrada y suponemos que Coca-Cola no se puede acomodar, Coca-Cola sólo podrá jugar duro.



Nota (Disuasores en caso de absorción) Las **píldoras envenenadas** son defensas legales o económicas que hacen poco atractiva la absorción de una empresa por otra cuando la operación es hostil y actúan como un instrumento de compromiso de dos maneras diferentes: o bien pueden desalentar a la empresa agresora de la entrada o bien pueden cambiar los términos de los acuerdos de adquisición. ♣

Ejercicio 30.14 *(Píldoras envenenadas) La compañía de ferrocarriles Norfolk Southern planea la absorción de la compañía de ferrocarriles CSX-Conrail. Si Norfolk Southern se decide a hacer una oferta CSX-Conrail pueden negarse a negociar, imponiendo sus propios términos, o puede llegara algún tipo de acuerdo. Otra opción para CSX-Conrail es incluir una píldora envenenada de forma que se vea comprometida a luchar. Dispone de dos opciones, la primera le compromete a luchar manteniendo los pagos que se obtienen (píldora I). La segunda también le compromete a luchar pero además cambia los pagos en caso de absorción (píldora II). Analizar la situación y, en particular, estudiar la decisión de CSX-Conrail de armarse o no con cada una de las píldoras envenenadas (los pagos se incluyen en su forma extensiva):*



♣

Ejercicio 30.15 *En el juego de entrada II suponemos que la decisión de Coca-Cola sobre su entrada al mercado es reversible en el sentido de que después de entrar y tras la decisión de Pepsi entre actuar duro y acomodarse, además de poder elegir entre actuar duro y acomodarse, puede salir del mercado con pérdidas para ella de -1 y ganancias para Pepsi de 3 si actuó duro y de 4 si se acomodó.*

Representar el juego en forma extensiva y resolverlo por inducción hacia atrás y mediante eliminación iterada de estrategias dominadas, explicando la conexión entre los pasos de eliminación iterada e inducción hacia atrás y cualquier relación con el poder de compromiso. ♣

30.3. Juegos con información imperfecta

En los juegos de información imperfecta también tenemos que determinar la credibilidad de los comportamientos futuros (amenazas y/o promesas). El determinar cuáles de los equilibrios de Nash son razonables va a motivar el concepto de equilibrio perfecto en subjuegos.

Definición 30.16 *Un **subjuego** es una parte de un juego en forma extensiva formada por una colección de nodos que satisface:*

- *El subjuego comienza en un nodo de decisión simple (empieza con un único nodo).*
- *El subjuego contiene todos los sucesores del nodo inicial (no podemos salir del subjuego).*
- *Si el subjuego contiene parte de un conjunto de información lo contiene completo (los jugadores poseen información completa sobre la historia del juego)*

Como una estrategia especifica que hacer en cualquier momento del juego también lo hace en el subjuego y, por tanto, es una estrategia para el subjuego (si s_i es una estrategia del juego original la estrategia para el subjuego g se denota por $s_i(g)$).

Definición 30.17 *Un perfil de estrategias (s_i, s_{-i}) en un juego en forma extensiva G es un **Equilibrio Perfecto en Subjuegos** (EPS) si especifica un equilibrio de Nash en cada subjuego:*

$$(s_i(g), s_{-i}(g)) \text{ es un equilibrio de Nash de } g \text{ para cada subjuego } g \subseteq G. \quad \clubsuit$$

Proposición 30.18 *En un juego con información perfecta los EPS son las soluciones por inducción hacia atrás (si es única hay un único EPS).* ♣

Nota Todo EPS es un equilibrio de Nash (pero no al revés). ♣

Ejercicio 30.19 *Juego de entrada III (ejemplo 30.7) Coca-Cola tiene que decidir si entrar o no en un nuevo mercado dominado por su rival Pepsi (ex Unión Soviética). La decisión de Coca-Cola se guía por la rentabilidad potencial de este nuevo mercado, que depende de la reacción de Pepsi. Si Pepsi contraataca y actúa duro Coca-Cola tiene pérdidas. Por el contrario, si Pepsi no contraataca y se acomoda lo que tiene son ganancias. En este ejemplo de juego de entrada Coca-Cola tiene que decidir si reacciona contraatacando sin conocer la reacción de Pepsi (denotamos E=enter, O=out, T=tough y A=accommodate).*

Juego en forma normal

	T	*A*
ET	*-2,-1*	*0,-3*
EA	*-3,1*	*1,2*
OT	*0,5*	*0,5*
OA	*0,5*	*0,5*

Juego post-entrada

	T	*A*
T	*-2,-1*	*0,-3*
A	*-3,1*	*1,2*

En el juego post-entrada tenemos dos equilibrios puros, (T,T) *y* (A,A) *y un equilibrio en estrategias mixtas* $(1/3,1/2)$*, donde estas probabilidades son las probabilidades de* T*. Demostrar que los únicos equilibrios perfectos en subjuegos son* $(E.A,A)$*,* $(O.T,T)$ *y* $(O.1/3,1/2)$*.* ♣

Ejercicio 30.20 *Se supone que en el juego entrada III la decisión de entrar o no de Coca-Cola se toma al mismo tiempo que la decisión de Pepsi entre actuar duro o acomodarse y que si Coca-Cola decide entrar además tiene que elegir entre actuar duro o acomodarse. Representar el juego en forma extensiva y resolverlo por inducción hacía atrás, indicando los subjuegos que tiene el juego.*

Considerar los siguientes casos:

a) Los pagos en cada nodo terminal son los mismos que en el juego de entrada III, con la salvedad de que si Coca-Cola se queda fuera los pagos son (0, -1) si Pepsi actúa duro y (0, 0) si se acomoda.

b) Los pagos en cada nodo terminal son los mismos que en el juego de entrada III, con la salvedad de que en este caso si Coca-Cola se queda fuera y Pepsi actúa duro los pagos son (0, 1) y si se acomoda (0, 0).

Ejercicio 30.21 *Se supone como en el ejercicio 30.20 que la decisión de entrada de Coca-Cola y la decisión de Pepsi entre actuar duro o acomodarse son simultáneas y que, además, cuando Coca-Cola entra y hace su elección entre actuar duro o acomodarse no conoce la elección de Pepsi.*

a) Representar el juego en forma extensiva considerando los mismos pagos que en el ejercicio 7 (a y b), indicando los subjuegos y las estrategias que tiene Coca Cola.

b) Explicar por qué esta estructura de tiempo es equivalente a que Coca-Cola y Pepsi tomen una sola decisión simultánea, indicando cuáles son las estrategias involucradas y cuales son los equilibrios de Nash del juego.

Ejercicio 30.22 *Dos ejércitos están enfrentados y pueden luchar o no. Los resultados dependen del nivel alerta y preparación, que puede ser alto o bajo (L=Low) . El ejército 1 está en disposición de declarar su nivel de alerta ante el segundo, que siempre está en alerta. Tras declarar el nivel de alerta se puede o no producir el enfrentamiento. Determinar la forma extensiva del juego y analizar sus EPS si las matrices de pago dependen del estado de alerta del ejército 1 y son:*

Con un nivel de preparación alto

Ejército 1\Ejército 2	*Luchar*	*No luchar*
Luchar	$-\frac{1}{2},-\frac{1}{2}$	*1,-1*
No luchar	*-1,1*	*0,0*

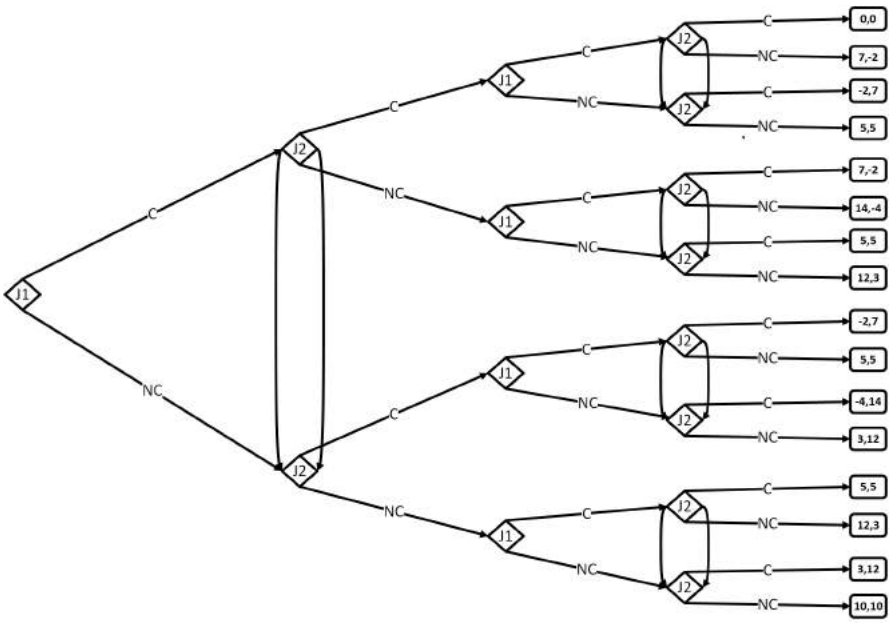
Con un nivel de preparación bajo

Ejército 1\Ejército 2	*Luchar*	*No luchar*
Luchar	$-\frac{3}{4},-\frac{1}{4}$	*0,0*
No luchar	*-2,2*	*1,3*

♣

30.4. Juegos repetidos finitamente

Ejemplo 30.23 *(El dilema del prisionero en dos etapas) En una primera aproximación al dilema del prisionero iterado se considerara que tras la primera interacción entre los prisioneros, y una vez reveladas sus decisiones, vuelven a interactuar y los pagos que obtienen son la suma de los pagos correspondientes:*



♣

Definición 30.24 *En todo juego repetido hay un componente del juego en forma estratégica que recibe el nombre de juego de etapa (stage game) y se juega en periodos discretos de tiempo* $t = 1, 2 \ldots, T$*:*

$$G = \{S_i, \pi_i; i = 1, 2 \ldots, N\},$$

con S_i *el conjunto de estrategias del jugador* i *y* π_i *su función de pagos (depende de* $s_1, s_2 \ldots, s_N$*).*

Los pagos del juego completo son:

- *Para juegos repetidos finitamente* ($T < +\infty$) *la suma de los pagos de cada etapa:*

$$\pi_{i0} + \pi_{i1} + \pi_{i2} + \cdots + \pi_{iT}$$

- *Para juegos repetidos infinitamente* ($T = +\infty$) *la suma de los pagos de cada etapa descontados según un factor de descuento* $\delta : 0 < \delta < 1$

$$\pi_{i0} + \delta\pi_{i1} + \delta^2\pi_{i2} + \cdots + \pi_{it} + \cdots$$

El factor de descuento δ *se puede interpretar como un factor de descuento asociado a un interés* ($\delta = \frac{1}{1+i}$) *o como la probabilidad de que el juego no termine.* ♣

Nota En cada etapa del juego los jugadores conocen la historia del proceso y las estrategias se pueden expresar como acciones a realizar dependiendo de la historia del juego hasta la etapa correspondiente. ♣

Nota (juegos repetidos con detección imperfecta) En algunos casos un jugador sabe que le corresponde mover pero no sabe exactamente qué es lo que ha pasado en etapas anteriores del juego. Cuando la ignorancia de un jugador se limita a lo que ha pasado en el juego pero la estructura del juego es de conocimiento común, hablamos de un juego de información completa con detección imperfecta. ♣

Ejemplo 30.25 *En el dilema del prisionero en* T *etapas el juego de etapa tiene un único equilibrio* (C, C)*. Como los jugadores pueden jugar en todo momento como si les quedara una única etapa el juego completo tiene un único equilibrio: los dos jugadores confiesan siempre y en cualquier contingencia.* ♣

Ejemplo 30.26 *En el dilema del prisionero modificado los prisioneros tienen también la opción de confesar parcialmente con la matriz de pagos*

1\ 2	*Confesar*	*No confesar*	*Confesar parcialmente*
Confesar	*0,0*	*7,-2*	*3,-1*
No confesar	*-2,7*	*5,5*	*0,6*
Confesar parcialmente	*-1,3*	*6,0*	*3,3*

En el dilema del prisionero modificado en T etapas el proceso se repite T veces y los pagos son la suma de los pagos de cada etapa. El juego de etapa tiene dos equilibrios (C,C) con pagos $(0,0)$ y (P,P) con pagos $(3,3)$ y un EPS se obtiene jugando siempre el mismo equilibrio de etapa.

Para analizar otros equilibrios que aparecen y mostrar que el comportamiento de los jugadores en la última etapa depende del comportamiento observado en etapas anteriores comenzamos considerando dos etapas.

♦ *En el dilema del prisionero modificado en 2 etapas aparece varios equilibrios perfectos en subjuegos. En particular, tenemos un EPS si ambos juegan con la siguiente estrategia*

- *Jugar N en la primera etapa.*

  Si en la primera etapa se ha jugado (N,N) jugar P y en caso contrario jugar C.

♦ *En el dilema del prisionero modificado en T etapas aparece un EPS si ambos juegan con la estrategia:*

- *Jugar N en la primera etapa.*

  Si en las etapas anteriores todos han jugado N jugar N, salvo en la última etapa que jugamos P. En caso contrario jugar C. ♣

Ejercicio 30.27 *Comprobar que en el dilema del prisionero modificado en 2 etapas aparece un EPS si*

- *El primer prisionero juega con la estrategia:*

  En la primera etapa jugar N.

  Si en la primera etapa se ha jugado (N,P) jugar P y en caso contrario jugar C.

- *El segundo prisionero juega con la estrategia:*

  En la primera etapa jugar P.

  Si en la primera etapa se ha jugado (N, P) *jugar P y en caso contrario jugar C.* ♣

Si hay más de un equilibrio de Nash siempre existe la posibilidad de mantener un buen comportamiento en las primeras etapas de interacción. El buen comportamiento en las interacciones tempranas puede ser recompensado por el juego de mejores equilibrios de Nash en subjuegos futuros y cualquier desviación de esta conducta puede ser castigado por el juego de malos equilibrios de Nash en subjuegos futuras. A un buen comportamiento en la primera etapa le corresponde un buen comportamiento en la siguiente y que a un mal comportamiento le corresponde un mal comportamiento en la siguiente. Si los jugadores observan cooperación en la primera etapa son propensos a pensar que ésta seguirá e incluso pueden jugar estrategias que no corresponden a equilibrio en un juego de etapa.

En resumen, en un juego repetido finitamente siempre se obtiene un EPS si los jugadores juegan repetidamente un equilibrio del juego de etapa. Si éste tiene un único equilibrio de Nash, el único EPS es que cada jugador juegue su estrategia de equilibrio en toda circunstancia. Sin embargo, si el juego de etapa tiene más de un equilibrio de Nash hay más de un EPS. Estos equilibrios pueden incluir estrategias que no correspondan con un equilibrio de etapa y se sustentan en la promesa de beneficios futuros.

Ejercicio 30.28 *Representar la forma extensiva de la batalla de los sexos en dos etapas y hacer un esquema del juego repetido T veces. Demostrar que en el juego repetido T veces un equilibrio perfecto en subjuegos es que ambos jugadores jueguen fútbol seguido de ópera cada dos etapas, independientemente de lo que se haya jugado en las etapas anteriores. Demostrar que también es un EPS jugar ópera seguido de fútbol cada dos etapas, independientemente de lo que se haya jugado en las etapas anteriores. Explicar por qué en un juego finitamente repetido con más de un equilibrio de Nash en el juego de etapa un EPS es alternar entre estos equilibrios.* ♣

30.5. Juegos repetidos infinitamente

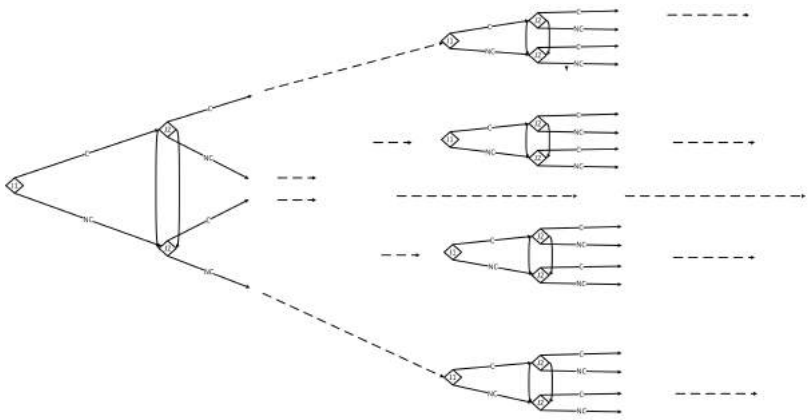
El dilema del prisionero iterado infinitamente surge cuando tras cada interacción entre los prisioneros, y una vez reveladas sus decisiones, vuelven a interactuar indefinidamente. En este caso no podemos consi-

derar que los pagos que obtienen son la suma de los pagos correspondientes a cada etapa y consideramos la suma de los pagos de cada etapa descontados según un factor de descuento δ que puede interpretarse también como la probabilidad de que el juego continúe.

En un juego repetido finitamente si el juego de etapa tiene un único equilibrio de Nash sólo se obtiene un EPS en el que los jugadores juegan repetidamente el único equilibrio del juego de etapa. Con el mismo juego de etapa pero repetido infinitamente hay otros EPS con estrategias que en alguna etapa no correspondan a un equilibrio. Estas estrategias se basan en la promesa de beneficios futuros y se sustentan mediante amenazas. En particular, si en el dilema del prisionero iterado la interacción es indefinida ambos presos pueden no confesar y encontrarse en un EPS en el que ambos utilizan la estrategia del disparador severo en la que los presos no confiesan siempre y cuando en las etapas anteriores no haya confesado ninguno.

- Estrategia del disparador severo
  - Jugar N en la primera etapa.
  - Si en las etapas anteriores los jugadores han jugado siempre N entonces jugar N.
  - En caso contrario jugar C de ahora en adelante.

Proposición 30.29 *En el dilema del prisionero iterado infinitamente tenemos un equilibrio perfecto en subjuegos si ambos jugadores juegan con la estrategia del disparador severo y el factor de descuento es lo suficientemente alto (el futuro importa):*



Si dividimos los subjuegos de una etapa en los subjuegos que siguen a un juego (N, N) en todas las etapas y el resto de subjuegos en los que se ha confesado alguna vez, en estos últimos la estrategia especifica jugar

siempre C y, al ser un equilibrio de Nash, y ningún jugador incrementa sus pagos jugando N contra C. Sin embargo, en los juegos que siguen a un juego (N, N) en todas las etapas los incentivos para desviarse y jugar C frente al otro jugando N dependen del pago:

- Si juega C obtiene un pago de 7 pero sus pagos futuros desaparecen y su pago total aumenta solo 7.
- Si juega N obtiene un pago de 5 y sus pagos futuros se mantienen con un incremento del pago total:

$$5 + \delta 5 + \delta^2 5 + \cdots = \frac{5}{1-\delta}$$

Así, el jugador no tiene incentivos para desviarse si $\frac{5}{1-\delta} > 7$ o equivalentemente si $\delta > \frac{2}{7}$. ♣

Nota En la estrategia del disparador con perdón se juega N en la primera etapa y se sigue jugando N mientras los jugadores sigan jugado siempre N. Sin embargo si algún jugador ha jugado C se juega C en las T etapas siguientes para jugar N en la etapa $T + 1$ con la misma estrategia en mente. Si ambos jugadores juegan con esta estrategia tenemos un EPS cuyos pagos medios son mayores que con la estrategia del disparador severo (δ próximo a 1):

- Si juega C obtiene un pago de 7 y T etapas después (en las que obtiene un pago de 0) obtiene un flujo de pagos futuros de 5 con lo que el incremento de su pago será

$$7 + 0 + \cdots + 0 + 5\delta^{T+1} + 5\delta^{T+2} + + \cdots = 7 + \frac{5\delta^T}{1-\delta}$$

- Si juega N obtiene un pago de 5 y un flujo de pagos futuros de 5 con lo que su pago se incrementa en

$$5 + 5\delta + 5\delta^2 + \cdots = \frac{5}{1-\delta}$$

No tiene incentivos para desviarse si $\frac{5}{1-\delta} > 7 + \frac{5\delta^T}{1-\delta}$ Si δ está próximo a 1 es equivalente a $5(T + 1) > 7$, con lo que incluso con un periodo de castigo es suficiente para obtener un EPS. ♣

Definición 30.30 *Un ciclo de comportamiento (behavior cycle) es una sucesión de acciones que se repite tras T=T1+T2+T3+T4 etapas en las que se juega*

- *(N,N) T1 etapas*
- *(C,C) T2 etapas*
- *(N,C) T3 etapas*
- *(C,N) T4 etapas,*

Este ciclo de comportamiento es individualmente racional si cada jugador obtiene pagos estrictamente positivos dentro del ciclo. ♣

Proposición 30.31 *(Teorema de tradición oral o teorema "folk") cualquier ciclo de comportamiento individualmente racional es factible como EPS y se sustenta en una estrategia del disparador severo (siempre que el factor de descuento δ sea cercano a uno).*

En esta estrategia comenzamos con el ciclo de comportamiento deseado y continuamos con él siempre y cuando no se desvíe ningún jugador (si alguno se desvía se confiesa a partir de entonces). ♣

Ejercicio 30.32 *(Modelo de competencia de precios) En un mercado la curva de demanda es $p = 6 - Q$, donde p es el precio al que se vende el producto y Q la cantidad de producto vendida (se vende todo lo que se produce). La empresa con menor precio obtiene todo el mercado, salvo que las dos empresas fijen el mismo precio, ya que en este caso se reparten el mercado (cada una obtiene la mitad). Los precios solo pueden fijarse en unidades de dólar y no hay costes de producción. Las empresas compiten en precios y la competencia continúa indefinidamente pues las dos firmas consideran que existe una cierta probabilidad, δ de que vuelvan a competir.*

a) Determinar la forma extensiva del juego, identificar sus subjuegos y las estrategias de las empresas, y definir un equilibrio perfecto en subjuegos.

b) Considerar la estrategia de fijar un precio de 2 dólares y continuar con él si fue el precio que ambas empresas fijaron en el pasado, pero si alguna empresa no fijó este precio cambiarlo a un dólar. ¿Para qué valores de δ es esta estrategia un equilibrio perfecto en subjuegos?

c) Demostrar que hay un equilibrio perfecto en subjuegos en el que el precio siempre es de 2 dólares que se sostiene con una estrategia del disparador con perdón (explícitar la naturaleza de esta estrategia).

d) ¿Cuál es el precio máximo que puede surgir en un equilibrio perfecto en subjuegos de este modelo si la probabilidad de que vuelvan a competir es del 90 % ($\delta = 0.9$)?.

e) Enunciar una versión del teorema folk para este modelo y argumentar su validez. ♣

30.6. La tragedia de los comunes (juegos markovianos)

Uno de los inconvenientes de la estructura de los juegos repetidos es que se supone que, literalmente, se juega el mismo juego de etapa en cada interacción. Aunque las estrategias de los jugadores son dinámicas y cambiantes, el fondo del problema es estático e inmutable. Los juegos markovianos abordan esta deficiencia, ya que se trata de una clase de juegos en los que tanto el problema subyacente como las estrategias de los jugadores son dinámicos. La diferencia con los juegos repetidos es que el juego se desarrolla en un entorno dinámico que puede cambiar de un período a otro y afecta a las ganancias en el juego de etapa de cualquier período (entorno del juego). El entorno puede cambiar por razones ajenas a la voluntad de los jugadores, o puede cambiar a causa de lo que hacen los jugadores. El problema de cómo usar un recurso de propiedad común permite introducir los juegos markovianos considerando que el entorno del juego es el tamaño del recurso, que evolucionará a través del tiempo de acuerdo con el patrón de uso en el pasado y afectará a los pagos de los jugadores en cada etapa del juego.

Vamos a comenzar nuestro análisis con un modelo sencillo en el que la principal conclusión que extraeremos será que en el equilibrio de Nash hay un uso excesivo del recurso comparado con el resultado socialmente deseable. El fenómeno se conoce con el nombre de la tragedia de los comunes debido a este uso excesivo, que es aún más marcado en poblaciones grandes (cuantas más personas utilizan el recurso o cuanto más intensamente es utilizado menos hay en el futuro).

La razón por la que nos preocupamos sobre los recursos de propiedad común es precisamente porque se pueden usar durante un período de tiempo largo y durante generaciones, siempre que se usen de forma que se regeneren. Su uso hoy marca el uso futuro y la pregunta clave es cuánto aprovechamiento supone un "mejor" uso del recurso, teniendo en cuenta que el carácter de acceso común hace que haya dos fuentes de externalidad: la utilización individual puede disminuir lo que usan los otros jugadores (externalidad actual) y la utilización conjunta puede afectar a la condición futura del recurso (externalidad futura). De hecho, en problemas relacionados con el medio ambiente la externalidad realmente mportante es la externalidad futura. En este sentido, este problema difiere del duopolio, ya que la sobreproducción de una empresa genera una externalidad a los demás jugadores reduciendo el precio de mercado pero no se traslada a futuras interacciones. En nuestro análisis vamos a considerar que la utilidad que produce el consumo del recurso es logarítmica (la mayor parte de los resultados son ciertos para cualquier función de utilidad cóncava).

Ejemplo 30.33 *(Modelo simple) Tenemos un recurso de propiedad común de tamaño $y > 0$ del que dos jugadores puede retirar una cantidad no negativa como consumo, siempre y cuando la cantidad total no sobrepase el tamaño del recurso, ya que el intento de consumir en exceso se traduce en que la cantidad disponible se reparte entre ellos. Cuando el consumo total es menor que el tamaño del recurso la cantidad sobrante es la futura base del recurso y, por tanto, del consumo futuro. Suponiendo en una primera aproximación que hay solo dos períodos de tiempo en el modelo:*

a) Determinar el patrón de consumo cuando cada jugador maximiza su utilidad.

b) Determinar el patrón de consumo cuando ambos jugadores maximizan la utilidad conjunta.

c) Comparar el patrón de consumo en el equilibrio de Nash con el consumo socialmente óptimo.

Solución En ambos casos denotamos por c_1 y c_2 las cantidades consumidas por los jugadores en el primer periodo y nos situamos en el último período (período 2) para proceder por inducción hacia atrás.
(a) En el último periodo cada jugador decide la cantidad a consumir del recurso disponible $y - (c_1 + c_2)$. Como no hay más períodos, no hay ninguna razón para guardar algo y cada jugador consumirá tanto como sea posible. Este intento de consumo hace que el recurso que queda se divida entre ellos y cada uno consume

$$(y - (c_1 + c_2))/2$$

En el primer periodo los jugadores tienen que determinar la cantidad a consumir del recurso y en este caso cada jugador quiere maximizar la utilidad total de ambos periodos. Como el consumo total del periodo determina el tamaño de los recursos que quedarán para el segundo periodo, la utilidad para un jugador depende de la cantidad que el otro jugador decida consumir. Por tanto el jugador 1 resuelve el problema

$$\max_{c_1} \left[\ln(c_1) + \ln\left(\frac{y - (c_1 + c_2)}{2} \right) \right]$$

La situación para el segundo jugador es análoga. Por las condiciones de primer orden las funciones de mejor respuesta de los jugadores son

$$R_1(c_2) = \frac{y - c_2}{2y} \qquad R_2(c_1) = \frac{y - c_1}{2y}$$

El equilibrio de Nash viene dado por los niveles de consumo c_1^* y c_2^* tales que $R_1(c_2^*) = R_2(c_1^*)$, con lo que, sustituyendo en las funciones de reacción, el equilibrio de Nash es $c_1^* = c_2^* = y/3$.

En el equilibrio de Nash cada jugador consume y/3 en el primer período y dejan un total de y/3 para el segundo período. Esta cantidad se divide por la mitad en el segundo período con un consumo de y/6 cada uno. La utilidad para cada jugador en este modelo es $\ln(y/3) + \ln(y/6)$.

(b) Para determinar el patrón de consumo de los recursos que constituye un uso socialmente óptimo suponemos que los dos jugadores constituyen una sociedad y se reúnen para decidir cuánto debe consumir cada uno para que la utilidad conjunta sea la máxima posible. ("bien común"). Por lo tanto, un patrón de consumo, (c_1^*, c_2^*), es socialmente óptimo si se maximiza la suma de la utilidad de los dos jugadores:

$$\max_{(c_1,c_2)} \left[\ln c_1 + \ln c_2 + 2 \ln \left(\frac{y - (c_1 + c_2)}{2} \right) \right]$$

Este procedimiento conduce a una solución socialmente óptima en la que cada jugador consume un cuarto del recurso en cada período, $c_1^* = c_2^* = y/4$.

(c) En la solución socialmente óptima exactamente la mitad del recurso se utiliza en el primer período, en contraste con el equilibrio de Nash donde se utiliza hasta dos terceras partes del recurso. En este sentido, tenemos un exceso de extracción del recurso y por eso decimos que tenemos una tragedia en el uso de los bienes comunes. Las razones del exceso de extracción son las externalidades del uso individual del recurso. Si un jugador recorta su consumo en el primer periodo en una unidad hay una unidad más disponible para el consumo de ambos en el segundo. Sin embargo, este jugador solo será capaz de recuperar la mitad de esta unidad adicional en el segundo período, debido a que la otra mitad queda para el jugador 2, y tiende a consumir en exceso. Este fenómeno no se da en la solución socialmente óptima en la que nos preocupamos por la utilidad para ambos. Una unidad de consumo dejada para mañana por cualquiera de los jugadores sigue siendo en el segundo periodo una unidad de consumo para la sociedad en su conjunto y, por tanto, no hay externalidad que distorsione el consumo del primer periodo. ♣

Ejemplo 30.34 *Determinar el patrón de consumo en el que cada jugador intenta maximizar su propia utilidad (equilibrio de Nash) y el patrón de consumo en el que ambos jugadores intentan maximizar la utilidad conjunta (consumo socialmente óptimo) cuando el número de jugadores que usan el recurso aumenta y comparar ambos patrones de consumo.*

Solución (denotamos por N el número de jugadores)

El consumo que maximiza la utilidad individual de un jugador, por ejemplo el jugador 1, depende de lo que los otros consumen en el primer período y se determina a partir del problema de optimización:

$$\max_{c_1} \left[\ln(c_1) + \ln\left(\frac{y - (c_1 + c_2 + \cdots + c_N)}{N} \right) \right]$$

con función de respuesta $R_1(c_2, \ldots, c_N) = \dfrac{y - (c_2 + \cdots + c_N)}{Ny}$

Vamos a simplificar el problema suponiendo que en el equilibrio de Nash cada jugador consume la misma cantidad (equilibrio simétrico).

$$c^* = R_1(c^*, \ldots, c^*) \Longleftrightarrow c^* = \frac{y - (N-1)c^*}{Ny} \Longrightarrow c^* = y/(N+1).$$

El consumo total, por tanto, es $yN/(N+1)$ y la cantidad que queda después del primer período es $y/(N+1)$. Si N es grande una extremadamente pequeña cantidad del recurso alcanza el segundo periodo y, por tanto, la tragedia de los comunes se agrava en las grandes poblaciones.

Para determinar el patrón de consumo socialmente óptimo se maximiza la suma de la utilidad de todos los jugadores:

$$\max_{(c_1,\ldots,c_N)} \left[\ln c_1 + \cdots + \ln c_N + N \ln\left(\frac{y - (c_1 + \cdots + c_N)}{N} \right) \right]$$

Este procedimiento conduce a una solución socialmente óptima en la que cada jugador consume en el primer período, $c_1^* = \cdots = c_N^* = y/2N$. Por tanto, siempre queda la mitad del recurso para el futuro, no importa cuántos jugadores haya.

Una unidad de consumo retenido para consumo social en la actualidad está siempre disponible como una unidad de consumo en el futuro, independientemente del número de miembros que haya en la sociedad. Así que la conclusión de que la mitad del total de recursos debe ser retenido para consumo social en el futuro es un resultado que es independiente del número de jugadores. Por el contrario, en el uso individualista del recurso, a medida que el número de jugadores aumenta, la tragedia aún más grave. La razón es que si un jugador renuncia a una unidad de consumo en el primer periodo es capaz de recuperar sólo una fracción de esta unidad en el próximo período $(1/N)$ y esto hace que el jugador esté aún menos dispuesto a guardar cualquier recurso para el futuro. ♣

Hasta ahora hemos considerado que los recursos son agotables, con lo que la cantidad del recurso que queda al final del primer período es exactamente la cantidad disponible para consumo en el segundo período. En el siguiente ejemplo vamos a considerar un modelo de recursos renovable en el que la cantidad sobrante aumenta al inicio del segundo periodo (sigue existiendo externalidad futura porque al sacrificar una unidad de consumo se obtiene solo una fracción del incremento en el segundo periodo).

Ejemplo 30.35 *(Modelo simple con recursos renovables) En esta variante del modelo del ejemplo 30.33 lo que no se consume en el primer período se regenera y la cantidad de recurso obtenida es una función, $f(x) = \sqrt{x}$, de la cantidad de recurso disponible al final del primer periodo, x.*

a) Determinar el patrón de consumo cuando cada jugador maximiza su utilidad (equilibrio de Nash).

b) Determinar el patrón de consumo socialmente óptimo.

c) Comparar los patrones de consumo, entre ellos y con los de consumo de recursos no renovables.

d) Analizar el problema de los recursos renovables cuando el número de jugadores aumenta. ♣

La tragedia de los comunes con interacción continua y recursos renovables es un ejemplo de juego dinámico markoviano y nos va a permitir introducir su formulación genérica.

Ejemplo 30.36 *(Tragedia de los comunes con interacción continua y recursos renovables) Consideramos dos jugadores que tienen acceson a un recurso común que les produce una utilidad logarítmica. El tamaño del recurso limita el consumo total y cambia con el consumo de los jugadores (entorno del juego). La cantidad del recurso no extraído en cada periodo es la base de inversión que genera el recurso futuro a través de una función de producción que supondremos e (en esta función cuanto más alta es la base de inversión mayores son las existencias pero con un aumento cada vez menor)*

$$f(x) = 10\sqrt{x}.$$

a) Determinar el patrón de consumo cuando cada jugador maximiza su utilidad (equilibrio de Nash).

b) Determinar el patrón de consumo cuando ambos jugadores maximizan la utilidad conjunta (consumo socialmente óptimo).

c) *En ambos casos, determinar la evolución del stock y estudiar si hay un stock sostenible en el tiempo. Comparar los patrones de consumo y analizar si la interacción estratégica lleva a la sobreexplotación.*

Solución Denotamos por $y(t)$ el tamaño de los recursos en el período t y por $c_i(t)$ el consumo del jugador i en el mismo periodo. Se tiene en cada periodo

- Tanto el tamaño del recurso como el consumo de los jugadores son positivos: $y(t) \geq 0$ y $c_i(t) \geq 0$.
- El consumo total esta limitado por las existencias: $c_1(t) + c_2(t) \leq y(t)$.
- La base de inversión que genera el recurso futuro es la cantidad no consumida

$$x(t) = y(t) - (c_1(t) + c_2(t))$$

- El tamaño del recurso del período siguiente es

$$y(t+1) = 10\sqrt{x(t)} = 10\sqrt{y(t) - (c_1(t) + c_2(t))}$$

Si no hay cantidad de recurso que invertir, $x(t) = 0$, en el período siguiente no hay stock de recurso y el juego acaba, $y(t+1) = 0$. Un horizonte natural para el juego es que continue siempre y cuando exista un nivel de recursos positivos. Por tanto, potencialmente, el juego puede continuar eternamente.

(a) Supongamos que los dos jugadores extraen el recurso unilateralmente.

En este caso, cada jugador sólo considera su propia utilidad y trata de maximizarla. El equilibrio del juego puede ser obtenido por inducción hacia atrás suponiendo que si los consumos deseados son mayores que el stock disponible, este stock se divide a partes iguales entre los dos jugadores.

► Para empezar, supongamos que hay exactamente un periodo de uso y estamos en él con stock y.

Dado que la utilidad aumenta con la cantidad consumida y no hay más períodos después, cada jugador intentará consumir cuanto más mejor. Por lo tanto, en este periodo cada jugador consume $c_1 = c_2 = y/2$. En consecuencia, la utilidad de equilibrio del jugador 1 viene dada por

$$W_1^{(1)}(y) = \ln(y/2) = \ln(y) + B(1)$$

- Si hay dos períodos de uso, en el primero los jugadores van a consumir una fracción del recurso

$$c_1 = \theta_1 y \qquad c_2 = \theta_2 y$$

El jugador 1, por ejemplo, se enfrenta al siguiente problema

$$\max_{c_1} \left[\ln(c_1) + \delta W_1^{(1)}\left(10\sqrt{y - c_1 - c_2}\right)\right]$$

donde el segundo sumando corresponde a la utilidad que le produce el resultado de la regeneración del recurso sobrante en el siguiente periodo (esta utilidad se descuenta un periodo mediante el factor de descuento δ).

Sustituyendo $c_1 = \theta_1 y$, $c_2 = \theta_2 y$ y suprimiendo todas las constantes irrelevantes el problema puede ser escrito de forma equivalente como

$$\max_{\theta_1} \left[\ln(\theta_1) + \frac{\delta}{2}\ln(1 - (\theta_1 + \theta_2))\right]$$

donde la mejor respuesta de consumo es

$$\theta_1 = b(\theta_2) = \frac{1 - \theta_2}{1 + \frac{\delta}{2}}$$

En equilibrio, $\theta_1 = b(\theta_2)$ y $\theta_2 = b(\theta_1)$, cada jugador consume con la misma tasa de extracción (equilibrio simétrico)

$$\theta_1 = \theta_2 = \frac{1}{2 + \frac{\delta}{2}}$$

La utilidad de equilibrio cuando quedan dos períodos es

$$W_1^{(2)}(y) = \ln\left(\frac{y}{2+\frac{\delta}{2}}\right) + \delta\, W_1^{(2)}\left(10\sqrt{y - \left(\frac{y}{2+\frac{\delta}{2}}\right) - \left(\frac{y}{2+\frac{\delta}{2}}\right)}\right)$$

que se puede escribir como

$$W_1^{(2)}(y) = (1 + \frac{\delta}{2})\ln(y) + B(2)$$

► Supongamos que quedan tres períodos de uso. En el primero de estos períodos tenemos que resolver el problema:

$$\max_{c_1}\left[\ln c_1+\delta W_1^{(2)}\left(10\sqrt{y-c_1-c_2}\right)\right]$$

Sustituyendo $c_1=\theta_1 y$, $c_2=\theta_2 y$ y suprimiendo todas las constantes irrelevantes el problema puede ser escrito de forma equivalente como:

$$\max_{c_1}\ln\theta_1+\frac{\delta}{2}\left(1+\frac{\delta}{2}\right)\ln\left(1-\theta_1-\theta_2\right)$$

La tasa de extracción de equilibrio es

$$\theta_1=\theta_2=\frac{1}{2+\frac{\delta}{2}+\left(\frac{\delta}{2}\right)^2}$$

► Si suponemos que quedan T períodos podemos ver un patrón y hacer una conjetura:

La tasa de extracción de equilibrio es

$$\theta_1=\theta_2=\frac{1}{2+\frac{\delta}{2}+\cdots+\left(\frac{\delta}{2}\right)^{T-1}}$$

En el modelo con infinitos períodos, siguiendo la conjetura, el consumo en equilibrio, $c*(y)=\theta^* y$, vendrá dada por el límite del consumo de equilibrio cuando T tiende a infinito

$$c^*(y)=\frac{1}{2+\frac{\delta}{2}+\cdots+\left(\frac{\delta}{2}\right)^{T-1}+\ldots}y=\frac{1-\frac{\delta}{2}}{2-\frac{\delta}{2}}y=\frac{2-\delta}{4-\delta}y$$

La base de inversión es

$$x(t)=y-2c^*(y)=\frac{\delta}{4-\delta}y.$$

Por tanto, la fracción de la inversión óptima asociada es $\frac{\delta}{4-\delta}$ y la ecuación que mide la evolución del recurso es

$$y(t+1)=10\sqrt{x(t)}=10\sqrt{\frac{\delta}{4-\delta}y(t)}$$

cuyo punto fijo es el stock del recurso sostenible en equilibrio, $\frac{100\delta}{4-\delta}$.

(b) Para obtener la solución socialmente óptima debemos maximizar la suma de las utilidades de los dos jugadores para lo que vamos a proceder a través de la inducción hacia atrás.

► Si solo hay un periodo en el que el stock es y, como en este último periodo nunca queda ningún stock sin uso, se consumirá todo y cada jygador consumirá la mitad $c_1 = y/2$, $c_2 = y/2$. En consecuencia, la utilidad social óptima de cada jugador cuando solo queda una etapa corresponde al stock disponible y está dada por

$$V^1(y) = \ln\left(\frac{y}{2}\right) = \ln y - \ln 2 = \ln y + A_1$$

donde A_1 es una abreviatura para la constante.

► Si solo hay dos períodos de uso, en el primero la extracción social óptima se encuentra al resolver el problema:

$$\max_{c1+c2\leq y} \left[\ln c_1 + \ln c_2 + 2\,\delta\, V_1^1\left(10\sqrt{y-c1-c2}\right)\right]$$

donde la utilidad del segundo periodo se descuenta un periodo mediante el factor de descuento δ.

Si al igual que en el caso anterior consideramos que los jugadores van a consumir una fracción del recurso, $c_1 = \theta_1 y$, $c_2 = \theta_2 y$, podemos simplificar el problema:

$$\max_{\theta_1+\theta_2\leq 1} \left[\ln\theta_1 + \ln\theta_2 + \delta\, 2\, \ln(1-\theta_1-\theta_2)\right]$$

En este caso,

$$c_1 = c_2 = \frac{y}{2+\delta}$$

donde la utilidad social óptima por jugador cuando quedan dos etapas depende del stock disponible y viene dada por

$$V_1^{(2)}(y) = \ln\left(\frac{y}{2+\delta}\right) + 2\delta V_1^{(1)}\left(10\sqrt{y-\left(\frac{y}{2+\delta}\right)-\left(\frac{y}{2+\delta}\right)}\right) = \left(1+\frac{\delta}{2}\right)\ln y + A_2$$

► Si solo quedan tres períodos de uso, en el primero de ellos tenemos que resolver el problema:

$$\max_{c_1+c_2\leq y} \left[\ln c_1 + \ln c_2 + 2\,\delta\, V_1^{(2)}\left(10\sqrt{y-c1-c2}\right)\right]$$

Análogamente al caso anterior, el problema podemos reescribirlo

$$\max_{\theta_1+\theta_2\leq 1}\left[\ln\theta_1+\ln\theta_2+\delta\left(1+\frac{\delta}{2}\right)\ln\left(1-\theta_1-\theta_2\right)\right]$$

El consumo social óptimo es

$$c1_{=}c_2=\frac{y}{2\left(1+\frac{\delta}{2}+\frac{\delta^2}{4}\right)}$$

al que le corresponde una utilidad social óptima per cápita de

$$V^3(y)=\left(1+\frac{\delta}{2}+\frac{\delta^2}{4}\right)\ln y+A_3$$

► Cuando el número de períodos restantes es T, podemos ver un patrón y hacer una conjetura en la que el consumo social óptimo es

$$c1=c2=\frac{y}{2\left(1+\frac{\delta}{2}+\frac{\delta^2}{4}+\cdots+\left(\frac{\delta}{2}\right)^{T-1}\right)}$$

En el modelo de infinitos períodos en cada etapa hay exactamente el mismo número de períodos restantes y la fracción de consumo de cada etapa es idéntica y viene dada por el límite del consumo óptimo cuando T tiende a infinito:

$$c(y)=\frac{y}{2\left(1+\frac{\delta}{2}+\frac{\delta^2}{4}+\cdots+\left(\frac{\delta}{2}\right)^{T-1}+\ldots\right)}=\frac{y}{2\left(\frac{1}{1-\frac{\delta}{2}}\right)}=\left(1-\frac{\delta}{2}\right)\frac{y}{2}$$

La base de inversión es

$$x(t)=y(t)-2c(y(t))=\frac{\delta}{2}y(t).$$

Por tanto, la fracción de la inversión óptima asociada es $\delta/2$ y la ecuación que marca la evolución del recurso es

$$y(t+1)=10\sqrt{x(t)}=10\sqrt{\frac{\delta}{2}y(t)}$$

con un stock de recurso sostenible socialmente óptimo que corresponde al punto fijo de la ecuación y es $100\delta/2$

La utilidad socialmente óptima viene dada por

$$V(y) = \left(1 + \frac{\delta}{2} + \frac{\delta^2}{4} + \ldots\right) \log y + A$$

► Para comparar el óptimo social y los resultados del equilibrio de Nash consideremos dos sociedades distintas, una en la que se gestiona el consumo socialmente y otra en la que cada jugador determina de manera unilateral cuánto consumir.

La función de consumo socialmente óptima c (y) es siempre menor que la función de consumo en equilibrio (dando cada individuo su mejor respuesta) c*(y)

$$c(t) = \frac{1 - \frac{\delta}{2}}{2} y < \frac{1 - \frac{\delta}{2}}{2 - \frac{\delta}{2}} y = c^*(y)$$

por tanto, si ambas sociedades comienzan con el mismo stock, en el siguiente periodo la primera sociedad tendría un stock más grande, ya que invirtió más en el primer período. Este aumento implica a su vez que esta sociedad vuelve a invertir más en el siguiente y así sucesivamente. Así, la primera sociedad invierte una fracción más grande de cualquier stock y siempre cuenta con un stock de recursos más grande disponible. Más aún, el stock sostenible también es más alto:

- En el primer caso el stock sostenible es $y = 100\delta/2$
- En el segundo caso el stock sostenible es $y^* = 100\delta / (4-\delta)$

La lección general es que la extracción unilateral conduce a un exceso de extracción; el consumo en la solución de equilibrio es mayor que en la solución socialmente óptima. En la solución de equilibrio un jugador sólo recoge parte de las consecuencias de sus acciones. Por lo tanto, es más probable que consuma de más, ya que parte del consumo excesivo se soporta con un descenso del consumo futuro del otro jugador. En cada período la sociedad administrada socialmente tendría un mayor stock de recursos, invertiría más, y por tanto seguiría teniendo un mayor stock en el futuro (incluso si empezamos con la misma cantidad de recursos). El stock sostenible en el primer caso es más alto que en el segundo (un stock sostenible es aquel que se mantiene regenerado y está al mismo nivel en todos los períodos). ♣

Una conclusión que podemos extraer de la tragedia de los comunes para los juegos dinámicos en general es que en el juego con horizonte finito cada jugador mira el número de períodos de extracción con el fin de tomar su decisión. Sin embargo, con horizonte infinito hay un equilibrio en el que cada jugador decide cuánto consumir sólo mirando el tamaño del recurso actual (entorno del juego). Esta situación es típica de los juegos markovianos y las estrategias de este tipo tienen un nombre técnico, son las llamadas estrategias de Markov o markovianas. Estas estrategias son atractivas debido a su simplicidad, ya que no requieren que el jugador tenga información de lo que ha hecho su rival en el pasado o de cómo el estado del juego ha evolucionado en períodos anteriores. El primer paso para su definición es definir formalmente un juego markoviano.

Definición 30.37 *Un juego markoviano está formado por:*

- *Un conjunto finito de jugadores* $I = \{1, 2 \ldots, N\}$.
- *Un espacio de estados* S *(finito o infinito).*
- *Un conjunto de acciones* A^i *para cada jugador (finito o infinito).*
- *Una correspondencia que asocia a cada estado las acciones que puede tomar cada jugador,* $\phi_i(s)$.
- *Una corriente de pagos* $\pi_1^{(i)}, \pi_2^{(i)}, \ldots$, *en la que cada pago* $\pi_t^{(i)} = \pi_t^{(i)}(s_t, a_t)$ *depende de:*
  - *el estado del juego en el periodo t,* s_t,
  - *el conjunto de acciones elegidas simultáneamente por los jugadores,* $a_t = (a_t^i)_i$ *(perfil de acción).*
- *Una función de transición que determina el estado del siguiente periodo, dado el estado actual,* s, *y el perfil de acción elegido por los jugadores,* a
  - ♦ *El juego comienza en un estado inicial,* s_0
  - ♦ *En cada etapa t:*

  I.- los jugadores observan el estado del juego, s_t,

  II.- dependiendo del estado, eligen simultáneamente acciones a_t^i

  III.- dependiendo tanto del estado como del perfil de acción, reciben su pago. ♣

Nota En un juego markoviano estocástico la función de transición no es determinista y depende de la probabilidad de que el siguiente estado el juego sea un estado determinado dado el estado del juego en este periodo y el perfil de acción jugado. ♣

Nota Por convención, un juego markoviano comienza en cero, con lo que el último periodo de un modelo con T periodos es $T-1$, ♣

Nota En los juego markovianos en cada estado se pueden considerar **acciones mixtas** como distribuciones de probabilidad sobre el conjunto de acciones disponibles para un jugador, $\phi_i(s)$. El conjunto de las acciones mixtas disponibles para el jugador i se denota por $M_i(s)$ y por $\pi_i(s,\nu)$ el correspondiente pago esperado para un perfil de acciones mixtas ν. ♣

El objetivo de cada jugador es seleccionar una estrategia que maximice su recompensa total (descontada adecuadamente). El equilibrio aparece cuando ningún jugador puede mejorar su recompensa desviándose unilateralmente del vector de estrategias correspondiente al equilibrio. Por tanto, el primer paso es definir las estrategias posibles en un juego marcoviano.

Definición 30.38

► *La historia del juego hasta el periodo t, t-historia, de un juego, h_t, es una descripción completa de la evolución del juego hasta el principio del periodo t y le corresponde un estado del juego al final del periodo $s_t(h_t)$.*

▹ *La descripción de una t-historia, además del estado del juego al principio del periodo, especifica en todos los periodos anteriores el estado del juego y las acciones escogidas por los jugadores.*

► *La estrategia de cada jugador es una función que en cada periodo t especifica para cada t-historia la acción a realizar por el jugador, $\sigma_{i,t}(h_t)$.*

► *Un perfil de estrategias es una estrategia para cada jugador y se denota por σ y en cada periodo, t, determina:*

- *el pago de cada jugador como función del estado inicial (recompensa total descontada):*

$$V(\sigma)(s) = \sum_{t=0}^{\infty} \delta_i^t \pi_t^{(i)}(s_t, a_t)$$

- *un perfil de estrategias para el resto del juego para cada t-historia, $\sigma^*(h_t)$.* ♣

En cada periodo la historia pasada del juego continua influyendo en las posibles recompensas de los jugadores solo mediante sus efectos sobre el valor del estado. Una estrategia markoviana es una estrategia que solo depende del estado del juego.

Definición 30.39

► *Una **estrategia markoviana** es una estrategia que en cada periodo depende de la historia hasta dicho periodo solo a través del estado del juego en el periodo y puede representarse por una sucesión de funciones del espacio de estados en el conjunto de acciones disponibles para el jugador, $\phi_i^t(s)$.*

► *Esta estrategia es **estacionaria** si las funciones son iguales en todos los periodos y, en este caso, la denotamos por $\phi_i(s)$.* ♣

La idea de equilibrio en un juego markoviano es la de equilibrio de Nash y corresponde a un perfil de estrategia, σ^* en el que ningún jugador se beneficia por una desviación unilateral del perfil.

Definición 30.40

► *El equilibrio en un juego markoviano es perfecto en subjuegos si en todo periodo, t, y para cada historia desde t, h_t, el perfil de estrategias inducido $\sigma^*(h_t)$ constituye un equilibrio para el resto del juego.*

► *Un equilibrio perfecto en subjuegos mediante estrategias markovianas recibe el nombre de **equilibrio perfecto de Markov** (MPE).* ♣

Nota Un equilibrio en estrategias Markovianas estacionarias siempre es perfecto en subjuego. Esto no es necesariamente cierto para equilibrios en estrategias markovianas no estacionarias y hace que sea necesario incluir el requerimiento de perfección en subjuegos en la definición de equilibrio. ♣

Nota Si el número de entornos del juego es finito siempre existe un equilibrio perfecto de Markov. Pero cuando este número es infinito, aunque algunos juegos como la tragedia de los comunes tienen un equilibrio, no tenemos resultados generales sobre si siempre hay un equilibrio perfecto de Markov. Cuando existe tampoco podemos garantizar la unicidad, ya que puede haber muchos equilibrios perfectos de Markov y si tenemos en cuenta estrategias del disparador aparecen aún más equilibrios. ♣

En juegos repetidos los equilibrios perfectos de Markov no conducen a ningún patrón de comportamiento interesante, ya que en un juego repetido sólo hay un entorno del juego; por lo tanto, una estrategia de

Markov es lo mismo que la elección de la misma acción una y otra vez y un equilibrio perfecto de Markov es lo mismo que jugar un equilibrio de Nash del juego de etapa repetidamente.

La conclusión final sobre la tragedia de los comunes es que en equilibrio el recurso está siempre por debajo de la solución socialmente óptima. Se extrae demasiado recurso y las utilidades de equilibrio son más bajas que las utilidades socialmente óptimas. A veces se puede remediar este problema con la amenaza de un castigo mediante estrategias del disparador, ya que si los jugadores creen que en el futuro el buen comportamiento será recompensado y el mal comportamiento castigado son propensos a cooperar (esta conclusión es válida en una diversidad de juegos).