

Capítulo 0

Introducción ¹

La teoría de juegos es el estudio de las interacciones estratégicas entre agentes racionales con diferentes objetivos. La principal característica de esta teoría es que reconoce que las decisiones de un jugador pueden afectar de forma directa los objetivos de los demás jugadores. Aquí un agente puede ser un individuo, una firma, un partido político, cualquier entidad que tome decisiones en un ambiente de múltiples agentes. Qué entendemos por racional será algo que se discutirá a lo largo del libro, pero una definición útil para fijar ideas es simplemente agentes que tiene objetivos y buscan maximizarlos.

A modo de introducción, es bueno entender que la teoría de juegos es uno de los dos grandes pilares de la teoría económica. Siendo el otro, la teoría del equilibrio general (Arrow and Debreu (1954)). En la teoría del equilibrio general típicamente existen muchos agentes, consumidores o firmas que interaccionan, que buscan maximizar su utilidad independientemente de las acciones de los demás. En efecto, se supone que las acciones de un agente no afectan directamente a la función objetivo de los demás, pero las consecuencias que cada uno toma es mediada por un precio. Esta es una diferencia importante con la teoría de juegos en donde las acciones sí tienen un efecto directo sobre los demás y típicamente no existe un precio como un mediador de esa interacción. Otro aspecto fundamental es que en la teoría del equilibrio general el precio actúa como una señal eficiente de demanda y oferta de productos. De esta forma, bajo los supuestos, bastante ideales de la teoría, se puede demostrar la existencia de un equilibrio y que este equilibrio es eficiente en el sentido de Pareto; esto último, el famoso primer teorema del bienestar. Estos resultados son dos de los más importantes en la teoría

¹Riascos, A. Juegos Estratégicos y sus Aplicaciones. Versión 9, 2026.

económica. En contraste, en la teoría de juegos usualmente no hay ese mecanismo mediador como el precio y las acciones sí tienen efectos directos sobre los demás (*i.e.* externalidades). Además, los agentes no tienen incentivos a interiorizar esos efectos. Como los agentes actúan de forma independiente y únicamente en función de su propio beneficio, casi siempre la solución del juego va a ser ineficiente socialmente (no se cumple el primer teorema del bienestar). Esto hace que la teoría sea muy rica en ejemplos y esfuerzos por establecer mecanismos compatibles con los incentivos de los agentes y que permitan restablecer la eficiencia de los resultados de una interacción.

Al igual que la teoría del equilibrio general, la teoría de juegos utiliza bastantes supuestos simplificadores de la realidad y muchas veces sus predicciones o descripciones de lo que sucede en una interacción, no se ajustan a la realidad. Sin embargo, en esto también está justamente su valor. Tener un marco de análisis con supuestos claros para entender el mundo como desviaciones de la teoría. Esto nos ilumina sobre los supuestos, restricciones, incentivos que son importantes para entender el mundo y cuándo estos no se cumplen. Este marco conceptual y la idea de entender el mundo como desviaciones del mundo ideal, permite sugerir formas de intervenir o cambiar nuestra forma de pensar para lograr mejores resultados. Es, entendiendo de esta forma estas teorías, ideales y sobre simplificadoras, que podemos apreciar su enorme utilidad para comprender y analizar la realidad económica.²

Ahora un poco de taxonomía. La teoría de juegos la podemos dividir en dos grandes ramas; la teoría de juegos estratégicos (o juegos no cooperativos) y la teoría de juegos cooperativos. La principal diferencia entre estos juegos es que en los juegos estratégicos los jugadores actúan en forma independiente mientras que en los juegos cooperativos se supone que los jugadores pueden hacer coaliciones, o arreglos de obligatorio cumplimiento, mediante algún mecanismo que no es explícito. En este libro nos vamos a concentrar únicamente en la teoría de juegos estratégicos.

Los juegos estratégicos se pueden representar de dos formas.

1. Forma normal, para juegos estáticos.³
2. Forma extensiva, para juegos dinámicos.

²Aspiro a que las aplicaciones que se presentan en este libro sean un testimonio de la utilidad de la teoría que se expone. El psicólogo alemán, Kurt Lewin, hizo esta observación: *There is nothing more practical than a good theory.*

³También se denominan juegos en forma estratégica. Sin embargo, reservamos este nombre para toda la clase de juegos estratégicos (en forma normal y extensiva).

Los juegos en forma normal, a su vez, se pueden clasificar en juegos de información completa e incompleta. Los juegos en forma extensiva, en cambio, se clasifican en juegos de información perfecta y juegos de información imperfecta. La representación de un juego en forma extensiva es la forma más general de representar un juego. En este libro vamos a comenzar estudiando los juegos estratégicos en su representación normal. La Figura 0 muestra esquemáticamente una taxonomía de la teoría de juegos estratégicos y algunos ejemplos o aplicaciones.

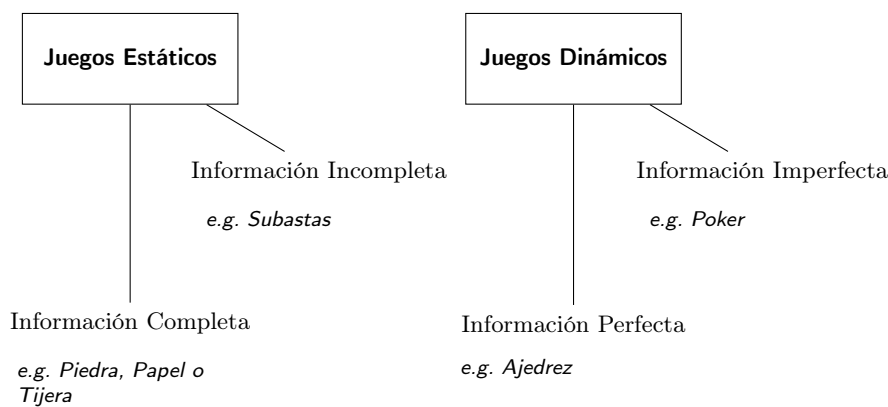


Figura 1: Taxonomía de la teoría de juegos estratégicos (*i.e.*, no cooperativos) con algunos ejemplos o aplicaciones importantes.

Todo juego plantea un problema de decisión para los agentes involucrados por eso la teoría de juegos es un ejemplo del estudio de la toma de decisiones con múltiples agentes. Ahora, queremos enfatizar, por lo menos informalmente, cuáles son las características más sobresalientes que intervienen en este tipo de decisiones. Por el momento, ignoramos las potenciales asimetrías de información entre los jugadores (*i.e.*, este será uno de los temas más relevantes cuando estudiemos juegos de información incompleta). Para el caso de juegos de información completa, los elementos fundamentales son:

1. El conjunto de acciones y/o un conjunto de consecuencias.
2. Una relación de preferencia sobre las consecuencias.⁴

⁴Usualmente, como en todos los ejemplos del libro, las preferencias se expresan sobre las acciones. Sin embargo, si se desea analizar las preferencias de forma detallada, por ejemplo, verificar si cumplen o no ciertos axiomas, lo adecuado es hacerlo sobre el espacio de consecuencias.

3. Una estructura de conocimiento (una estructura de información y el conocimiento que tienen los agentes de esta información).
4. Una hipótesis de comportamiento o racionalidad.

A lo largo del libro vamos a formalizar algunas de estas ideas, especialmente 1,2 y 4, pero también tendremos algunas cosas que decir, aunque sea informalmente, sobre la importancia de la estructura de conocimiento. En efecto el conjunto de acciones es típicamente fácil de describir, al igual que las consecuencias de un perfil de acciones. Como en general las acciones determinan de forma unívoca las consecuencias, casi siempre supondremos que las preferencias dependen directamente de las acciones.

Ahora, el concepto de estructura de conocimiento es el más complejo de formalizar y por eso en este libro usaremos una aproximación intuitiva (abajo el Ejemplo 0.1 a manera de motivación).

Por último, la hipótesis de comportamiento está íntimamente relacionada con la estructura de conocimiento. Básicamente, el supuesto fundamental es que los agentes toman acciones que son las mejores posibles dada sus conjeturas sobre las acciones de los demás. Las conjeturas a su vez dependen de la estructura de conocimiento de cada uno.

La relevancia de la estructura de conocimiento en un problema de decisión se explora informalmente en el próximo capítulo y será una constante a lo largo de este libro. En el siguiente ejemplo se pone de manifiesto la importancia que tiene el conocimiento que suponemos tiene un agente sobre el conocimiento que tienen los demás y viceversa.

Ejemplo 0.1 (Paradoja de los sombreros). Supongamos que hay tres personas sentadas en un salón. Cada una de ellas tiene un sombrero que saben que puede ser blanco o rojo. Cada una de ellas puede observar el sombrero que los demás tienen en su cabeza, pero no el que tiene en su cabeza. Supongamos que los tres son rojos. Lo único que saben las personas es que el sombrero puede ser blanco o rojo y lo único que pueden hacer es ver el color del sombrero de los demás. Si preguntáramos uno a uno cuál es el color del sombrero que ellos tienen en su cabeza, todos responderían “no sé”. Ahora, supongamos que les hacemos el siguiente anuncio: “existe por lo menos un sombrero rojo en la cabeza de ustedes”, y volvemos y les preguntamos uno a uno si saben con certeza cuál es el sombrero que tienen en la cabeza. Cada uno de ellos puede oír lo que responden los anteriores. En esta ocasión los dos primeros a quienes se les pregunta responden que no saben, pero el último debería de responder que su sombrero es con seguridad rojo.

Para ver esto, obsérvese que, antes del anuncio, las tres personas saben que hay por lo menos un sombrero rojo entre los tres (dado que sabemos que los tres sombreros son rojos y cada uno puede ver el sombrero de los demás). Más aún, el tercero sabe que el segundo sabe que hay por lo menos un sombrero rojo (debido a que sabe que el primero tiene un sombrero rojo y el segundo lo puede observar). Sin embargo, antes del anuncio, el tercero no sabe que el segundo sabe que el primero sabe que hay por lo menos un sombrero rojo entre los tres.

Ahora, una vez hecho el anuncio, saber que hay por lo menos un sombrero rojo entre los tres se convierte en conocimiento común: todos saben que los demás saben y todos saben que los demás saben que los otros saben, etc. En particular, el tercero sabe que el segundo sabe que el primero sabe que hay por lo menos un sombrero rojo entre los tres. Por lo tanto, dado que el primero y el segundo afirmaron no saber de qué color era el sombrero, el tercero puede deducir que su sombrero es rojo. De esta forma, si una vez hecho el anuncio los dos primeros responden “no sé”, el tercero puede razonar así. Si mi sombrero no fuera rojo, como el segundo jugador sabe que el primero respondió no sé, entonces es porque el segundo sabe que el primero está viendo un sombrero rojo. Si el mío no es rojo, entonces el segundo jugador debería deducir que el sombrero rojo que está viendo el primer jugador es el del jugador dos, luego el jugador dos debió responder que su sombrero era rojo. Como no lo hizo, eso quiere decir que mi sombrero es rojo. Obsérvese que la clave de este argumento es que después de hecho el anuncio, el jugador tres sabe que el jugador dos sabe que el jugador uno sabe que hay por lo menos un sombrero rojo entre los tres.

En contraste con el concepto de conocimiento común, existen muchas otras formas de relajar este supuesto. Por ejemplo, puede ser que los agentes no sean conscientes de todos los elementos del juego. Es decir, el juego no es conocimiento común (un jugador desconoce que existe un tercero, desconoce todas las acciones de sus oponentes, etc.) o puede tener un conocimiento limitado (sabe algo sobre su oponente, pero no sabe que su oponente sabe que él sabe), etc.

0.1. Ejercicios

1. Considere la siguiente variación de la paradoja de los sombreros. Se les dice a las tres personas: “voy a esperar un minuto para que me digan de qué color es el sombrero que tiene cada uno en su cabeza. Durante ese

minuto pueden levantar la mano y decir rojo o blanco. Si después de un minuto no han dicho nada se entiende que cada uno de ustedes no sabe el color de su sombrero con certeza. Una vez finalizado el juego los que tomaron una decisión ilógica serán penalizados, a los demás se les dará una bonificación. Por ejemplo, haber levantado la mano sin tener certeza, o haber dejado que pase el tiempo cuando hubiera podido levantar la mano porque debería de saber, serán penalizados.”

- a)* Sin ninguna información adicional, ¿Qué pasará después de un minuto?
- b)* Ahora se anuncia con un megáfono que existe por lo menos un sombrero rojo entre ellos tres y las tres personas juegan de nuevo con las mismas reglas de este ejercicio. ¿Cómo actúan las tres personas?