

EVOLUCIÓN DE LA COOPERACIÓN A TRAVÉS DE LA IMITACIÓN (*)

CONFERENCIA DE DAVID LEVINE (UCLA)

Le agradezco su amable presentación, mucho más de lo que yo merezco. Después de sus palabras no tengo mucho para agregar, por lo que voy a comenzar.

Esta ponencia se refiere al trabajo que realicé junto con Wolfgang Pesendorfer y trata sobre la teoría de juegos evolucionista. Creo que lo mejor sería que comenzara por darles alguna información básica, pues quizás no todos los participantes estén familiarizados íntimamente con los elementos que se han desarrollado en la teoría de juegos evolucionista durante los últimos diez o doce años. Supongo, pues, que lo mejor es comenzar con una breve introducción de las preguntas y respuestas que existen en el momento actual.

Voy a comenzar con un ejemplo del tipo de situaciones que fascinan a quienes se dedican a la teoría de juegos, y que ha generado recientemente muchas investigaciones. Para examinar estas situaciones en un marco muy concreto, les propongo analizar la situación de Alemania en la época de la caída del régimen comunista. No sé si ustedes recuerdan cómo se sucedieron los acontecimientos, pero lo básico es que la gente salió a la calle, y cuando todo el mundo, todo el país había salido a la calle, el gobierno cayó.

Pensemos ahora en esa época desde la perspectiva de una persona que vive en un régimen represor como lo era el régimen alemán. Esa persona debe tomar una decisión. Puede salir a la calle a manifestar contra el gobierno o puede quedarse en su casa. Obviamente, si usted es la única persona que sale a la calle a manifestar contra el gobierno, no le irá nada bien, se sentirá muy desgraciado, lo encarcelarán, etc., etc. Pero si todo el

(*) Conferencia corregida por el autor en el lenguaje original. Traducción del editor.

mundo a la vez sale a la calle a manifestar contra el gobierno, el gobierno caerá y todo el mundo estará muy contento. Y la otra cosa que puede pasar es que todos salgan a la calle a manifestar contra el gobierno, menos usted. Y eso tampoco será bueno para usted, porque sus vecinos después le preguntarán dónde estaba cuando todos estaban en las calles manifestando, le preguntarán por qué se quedó escondiéndose en su casa.

Esto es un ejemplo de lo que en la teoría de juegos se conoce como “juego de coordinación”. Se trata de un juego en el que hay más de un equilibrio. Hay dos equilibrios. En un primer equilibrio todo el mundo se queda en su casa, y si todo el mundo se queda en su casa no hay motivos para salir a manifestar a la calle, porque uno no desea ser el único que sale a la calle. Por otra parte, si todos salen a la calle a manifestar, también hay un equilibrio, porque uno no desea ser la única persona que se quede escondiéndose en la casa cuando todos están manifestando en la calle.

Otra característica de este juego es que no solo hay dos equilibrios, sino que ambos equilibrios son diferentes en cuanto a cuán deseable pueda ser alcanzarlos. Todos están de acuerdo en que prefieren el equilibrio en el que todo el mundo sale a manifestar a la calle, porque a nadie le gusta el régimen. Así que existen estos dos equilibrios diferentes, y de hecho en Alemania Oriental vimos los dos equilibrios. Vimos el equilibrio en el que todo el mundo se quedaba en su casa y luego súbitamente un día cambiamos al equilibrio en el que todo el mundo salía a la calle a manifestar.

Una de las cosas que los teóricos desearían tener es una teoría que nos diga, cuando hay más de un equilibrio, por qué motivo en un momento dado podríamos estar en uno de estos equilibrios y no en otro. Hace unos diez o doce años, un grupo de economistas dedicados a la teoría de juegos realizó un trabajo destinado a responder a dicha pregunta. Nombraré, entonces, a quienes podrían ser considerados los padres fundadores del campo de la teoría de juegos evolucionista moderna, que serían Kandori, Mailath, Rob y Young, que escribieron de hecho entre los cuatro dos trabajos diferentes. Ellos introdujeron la idea de que los equilibrios se alcanzan a través de un proceso de evolución y de que el proceso de evolución es un proceso imperfecto que se produce a través de la acumulación de errores aleatorios. Lo que ellos tenían in mente era un proceso similar al de la evolución biológica, donde se producen mutaciones, y la interpretación de las mutaciones es que la gente a veces hace cosas que no favorecen quizás sus mejores intereses inmediatos y que lo hace

quizás por accidente o quizás a propósito. Se trata, pues, de un proceso aleatorio, de forma que es posible que se produzcan muchos accidentes aleatorios y esto podría provocar el pasaje de un equilibrio a otro equilibrio.

En particular, en el tipo de encuadre que describí en Alemania, la predicción es que existe un proceso de ponderación entre el tipo de beneficios que aportaría el buen equilibrio y lo costoso que puede ser cometer un error y ser la única persona que sale a la calle; y la teoría que ellos proponen hace una predicción muy específica. Dice que uno debe hallar la respuesta a la siguiente pregunta: Supongamos que la mitad de la gente ha salido a la calle y la otra mitad de la gente está en su casa. ¿Qué debo hacer? ¿Debo salir a la calle o debo quedarme en mi casa? Y la predicción de la teoría es que el equilibrio que deberíamos observar en el muy largo plazo es el equilibrio de que cuando la mitad de la gente está en la calle y la otra mitad de la gente está en su casa, si uno desea salir a la calle y la mitad de la gente constituye una tentación suficiente para salir a la calle, entonces en el largo plazo veríamos el equilibrio en el que toda la gente está en la calle. Pero si cuando la mitad de la gente está en la calle salir a la calle todavía es demasiado peligroso y es probable que si uno sale a la calle uno se meta en problemas, entonces uno preferirá quedarse en la casa y el equilibrio que deberíamos observar en el largo plazo es el equilibrio en el que toda la gente se queda en su casa.

Esto vendría a ser un breve resumen de los principales resultados y motivaciones de la teoría de juegos evolucionista que existen en la actualidad. Ahora me gustaría hacer un par de comentarios sobre el trabajo que hicimos con Wolfgang Pesendorfer.

Me referiré a mi investigación conjunta con Wolfgang Pesendorfer, que es básicamente un avance en la agenda de la teoría de juegos evolucionista. Hay dos motivaciones para la investigación que realizamos en el contexto de lo que les he explicado acerca de la situación de la teoría de juegos evolucionista. La primera es que estamos interesados en un conjunto de aspectos algo diferentes. Originalmente, la teoría de juegos evolucionista se centró en el tema de los equilibrios por coordinación; pero a nosotros también nos interesan los encuadres en los que hay más de un equilibrio, porque ahí es donde la teoría de juegos evolucionista ha demostrado su valor, el de poder sugerir cuál de los muchos equilibrios será el equilibrio que realmente se alcanzará. Pero nuestro interés en los equilibrios múltiples se despliega de hecho en un encuadre diferente, en el

encuadre de la teoría tradicional de los juegos repetidos, en el encuadre conocido como el teorema de Folk. Les diré rápidamente qué es el Teorema de Folk y trataré de darles un ejemplo alternativo. El Teorema de Folk dice que en un juego repetido hay muchos tipos de equilibrio diferentes. Existe un equilibrio en el que la gente se comporta de manera bastante egoísta; en términos económicos ustedes podrían pensar en cuanto a prestar y pedir prestado. Un equilibrio en cuanto a prestar y pedir prestado es que nadie presta dinero, porque la gente cree que no les van a devolver lo que presten, y nadie devuelve lo que le han prestado. Ese sería un equilibrio. Todo el mundo persigue su propio interés personal, pero no hay intercambio intertemporal; nadie presta; nadie toma prestado; así que estamos ante un colapso del sistema económico.

El otro caso es el equilibrio en el que la gente presta dinero y espera que le devuelvan lo que prestó, y el motivo por el que esperan que le devuelvan lo que prestó es que saben que la gente que no devuelve lo que le prestaron recibe un castigo y será excluida de toda posibilidad de tomar préstamos en el futuro.

Así que éste es un equilibrio para el funcionamiento del mercado del crédito, y el punto que deseo aclarar es que en este tipo de encuadre hay dos tipos de equilibrios, un equilibrio autárquico en el que no hay intercambio y un equilibrio cooperativo en el que hay intercambio, pero el motivo por el que la gente paga las deudas en las que incurre para poder realizar intercambios es la amenaza de alguna venganza en el futuro si no paga sus deudas.

Así pues, por ejemplo, el motivo por el que Argentina no ha declarado una moratoria es evidentemente el temor de sufrir consecuencias negativas en el futuro si no paga su endeudamiento.

Pero hay, naturalmente, otro equilibrio en el que de todas formas no habrá intercambio en el futuro y por tanto no tendría sentido pagar lo que se debe, pues uno sabe que aun cuando pague las deudas en el futuro uno no obtendrá los beneficios del intercambio.

En el encuadre de una teoría de juegos en la que hay un equilibrio de autarquía y un equilibrio de cooperación que se da por medio de amenazas de venganza, desearíamos saber cuál de estos dos diferentes tipos de equilibrio va a ser alcanzado. Esta es, de hecho, una pregunta

muy tradicional en la teoría de juegos, que se plantea casi desde sus orígenes; se trata de una pregunta para la que solo los métodos recientes de la teoría de juegos evolucionista han comenzado a brindar razonables perspectivas de encontrar una respuesta.

Nuestra idea básica consiste, entonces, en adoptar el enfoque evolucionista de este problema; pero también proponemos otra innovación para complementar este procedimiento para el estudio del tema de la cooperación y la aparición de la cooperación en la teoría de juegos evolucionista. Esta modificación se refiere al propio funcionamiento de la teoría de juegos evolucionista. Permítanme que se las explique.

En la teoría de juegos evolucionista original uno debe comprender de qué forma se pasa del equilibrio en el que todo el mundo se queda en su casa al equilibrio en el que todo el mundo sale a la calle. La explicación de dicho fenómeno conforme a la teoría de juegos evolucionista original es que los cambios se dan a través de un proceso de mutaciones. Los sujetos realizan acciones y actúan en pro de sus propios intereses. Luego un día a alguien se le prende la lamparita y dice: Para mí sería bueno salir a la calle a manifestar. Ahora, si ese día los demás no salen a la calle, se trata de una muy mala decisión. Pero si resulta que, digamos, un 75% de la gente esa misma mañana al levantarse tiene la misma idea y todos salen a la calle, entonces la idea funciona y comienza la revolución. Pero para mucha gente no resulta plausible esta idea de que este tipo de cambio entre un equilibrio y otro se da porque simultáneamente y por casualidad a la gente se le ocurre la misma idea la misma mañana del mismo día.

La propuesta que hemos hecho para modificar este modelo es la del mecanismo de propagación, que hace que una cantidad suficiente de gente adopte la idea “mala”, ya que si bien salir a la calle solo es un mala idea, cuando sale a la calle un número suficiente de gente la idea resulta ser buena. El mecanismo de propagación no es un proceso de mutación en el que la gente decide al azar hacer algo sino más bien un proceso de imitación. Así, una persona, que lidera la revolución, pone un pie en la calle esperando que otras personas lo sigan. Las demás personas lo miran y se dicen: bueno ya hay alguien en la calle, quizás sea buena idea que yo también salga a la calle. Entonces una segunda persona sigue al primero. Obviamente tampoco sería una buena idea para la segunda persona si no hay más seguidores, porque entonces esas dos personas la pasarían muy mal. Pero si hay un tercero, un cuarto, un quinto, que los sigan, quizás no

sea imposible, y cuanto más gente salga a la calle mayores serán las posibilidades de que aumenten los seguidores. Así pues, la propuesta que haríamos es la que se refiere al mecanismo de propagación, por su interés para las sociedades humanas; el mecanismo de propagación en el que hay líderes, en el que hay seguidores, en el que la gente probablemente imite lo que hacen los demás sin tener en cuenta si esa acción es una buena idea o una mala idea. O sea que no actúan simplemente por motivos completamente aleatorios.

La posibilidad de que uno se despierte una mañana y aleatoriamente decida salir a la calle es, en mi opinión, muy baja en comparación a la posibilidad de que si uno ve que hay gente en la calle uno se decida a salir también a la calle y unirse a los demás.

Esa es, pues, nuestra hipótesis básica, que adoptamos por dos motivos: en primer lugar, porque nos parece una explicación más plausible de este fenómeno del vuelco de un equilibrio a otro equilibrio, y, en segundo lugar, por un motivo práctico, porque hace que sea posible realizar un análisis. Este supuesto lleva a una teoría matemática mucho más simple y que, en consecuencia, podemos aplicar al problema de la cooperación aplicada por medio de amenazas de venganza, que es un problema de mayor complejidad. Sería muy difícil estudiar este problema haciendo uso de la teoría de las mutaciones. La teoría de la imitación es más simple y, por tanto, puede ser usada como herramienta para examinar problemas que entrañan mayores dificultades.

Con esto termino la introducción básica, y creo que ahora debería resumir las conclusiones que hemos alcanzado en nuestro trabajo. Luego, para aquellos de ustedes que tengan interés, haré rápidamente una descripción más técnica del funcionamiento del modelo. Pero para el grupo en su conjunto, permítanme intentar mostrarles las conclusiones que pudimos sacar aplicando este método en particular.

En primer lugar, el encuadre que hemos estudiado. A nosotros nos interesa un encuadre en el que sea posible tener equilibrios, equilibrios cooperativos cuyo cumplimiento se exija a través de amenazas de venganza y equilibrios no cooperativos en los que las personas no cooperan entre sí. Y una de las cosas que hicimos en este estudio fue tratar de hallar un encuadre simple con ese tipo de equilibrios múltiples en donde se da la venganza. El encuadre que encontramos fue el siguiente. Imaginamos

personas que coinciden y se encuentran unas con otras para encarar algún tipo de interacción, donde hay una amplia variedad de cosas que pueden hacer. Se pueden comportar de forma más cooperativa o menos cooperativa, pueden pagar deudas, pueden hacer préstamos, pueden realizar pagos, pueden aplicar sanciones, pueden destruir cosas. Las personas tienen a su disposición una amplia variedad de opciones, y al desarrollar estas interacciones las personas pueden hacer ciertos compromisos en cuanto al tipo de estrategias o normas que seguirán en dicha interacción. Y la regla que cumple una de las personas puede ser potencialmente conocida por la persona con quien interactúa. Por ejemplo, es posible que una persona sea del tipo de personas que pagan los préstamos. Entonces, nosotros imaginamos que cuando esa persona se encuentra con otra para solicitar un préstamo, la otra persona recibe algún tipo de señal, posiblemente una señal ruidosa¹, que indica si la persona que pide el préstamo es una persona que pagará su deuda. Y lo que imaginamos es que cuando una persona emplea una estrategia en particular, existe la oportunidad de que la otra persona con la que interactúa pueda saber cuál es la estrategia que la primera persona está empleando.

Pasaré, entonces, a describir cómo podemos encontrar dos tipos diferentes de equilibrio en este encuadre en particular. Una posibilidad es una estrategia en la que no se realizan préstamos, no se pagan los préstamos, las personas son autárquicas y no hay intercambio. Cuando las demás personas adoptan esta estrategia y no hay intercambio, uno no obtendrá realmente ninguna ventaja del intercambio, uno nunca recuperará lo que haya prestado, porque nadie paga sus préstamos. Entonces, si uno realiza préstamos pero nadie paga los préstamos, uno no obtiene ningún beneficio. En ese caso, más vale ser como los demás y convertirse en una persona que no hace préstamos, no paga los préstamos, etc.

Pero la otra posibilidad es que uno adopte la estrategia de pagar los préstamos y de darle préstamos a las personas que en el pasado han pagado sus deudas, buenas personas, personas buenas como nosotros, personas a las que les daremos préstamos, pero a las personas que se identifican como personas malas, que no cooperan, no les haremos préstamos.

1 N.T. Las palabras “noise” (traducida como “ruido”) y “noisy” (traducida como “ruidoso”) que se usan en este trabajo corresponden a la jerga de las telecomunicaciones, donde indican una distorsión o deterioro en la recepción de la imagen o los mensajes que se transmiten.

Y como al encontrarse con una persona uno recibe una señal, podremos determinar, quizás con cierto grado de imperfección, si la otra persona es una persona que pagará su préstamo, y uno solamente hará préstamos si descubre que quien solicita el préstamo es una persona que pagará lo que debe.

Ahora bien, en este tipo de mundo, ¿debería uno decidir ser una persona que pagará sus préstamos o debería uno decidir ser el tipo de persona que no paga los préstamos? Sabemos que si uno elige ser el tipo de persona que paga sus préstamos, recibirá préstamos de los demás.

Sé que no estoy explicando una gran parte de este mecanismo que dice que es posible saber que una persona es honesta cuando uno ve una cara honesta, pero de hecho esto es algo que observamos en la práctica, y que es de mucha importancia. Estoy pensando en situaciones en las que las señales pueden ser señales ruidosas. Alguien puede tener la apariencia de ser honesto y de hecho ser un ladrón, y ésa es una posibilidad que también hay que contemplar.

Pasaré entonces a describir nuestras conclusiones. Tenemos una situación que es muy común en la teoría de juegos. Tenemos multiplicidad de equilibrios, los equilibrios buenos, cooperativos, y los equilibrios malos, no tan cooperativos; entonces permítanme que describa las conclusiones que podemos alcanzar aplicando este método de la teoría de juegos evolucionista imitativa.

Estudiamos básicamente dos situaciones diferentes. El encuadre más simple es el encuadre de identificación perfecta, donde las señales no presentan ruido. Si uno se encuentra con otra persona uno puede simplemente decir si esa persona es tramposa o no es tramposa, y no hay ambigüedad en dicha determinación. En ese encuadre lo que ocurre, y quizás no deba sorprendernos, es que el equilibrio que surge en el largo plazo como resultado del proceso evolucionista es que no hay tramposos. Todos cooperan y todos usan la estrategia cooperativa de cooperar solamente con las personas que cooperan. Y obviamente si uno encuentra a una persona que no coopera, uno no coopera con esa persona. Ese es el mecanismo que genera el cumplimiento.

Ahora bien, el rasgo interesante, que al menos no era tan obvio para mí cuando comenzamos este trabajo, es el tipo de tratamiento que uno da a las personas que no son el tipo correcto de persona. Si uno piensa en el

grupo de las personas cooperadoras, el grupo de las personas que tienen éxito, que cooperan entre sí, ¿cómo se comportan estas personas cuando se reúnen e identifican perfectamente a alguien que no es cooperador? Se supone que deben aplicarle una sanción a esta persona, pero ¿qué tipo de sanción deberían usar? La respuesta acerca de la sanción es la siguiente, que lo que ellos deberían hacer es tratar de maximizar la diferencia entre la utilidad que reciben y la utilidad que la otra persona recibe. En cierto sentido es como una contienda bélica, pero la principal diferencia es que uno no debería tratar de destruir al rival sino más bien de maximizar la brecha entre uno y el rival. Lo que uno debería intentar es hacer cosas que sean dañinas para la otra persona pero que para uno tengan un costo muy bajo, porque eso nos dará la mayor diferencia entre los beneficios que recibe la otra persona y los que uno puede recibir. Esta teoría de hecho hace una propuesta muy específica acerca de la forma en que se debe castigar a los tramposos. Se refiere a una forma en particular en la que deben ser castigados y un criterio en particular para castigarlos.

Esto tiene una interesante ampliación en un mundo de señales imperfectas. El segundo encuadre que estudiamos es un mundo en el que uno recibe una señal ruidosa que solo nos permite identificar de manera imperfecta si la persona con la que estamos interactuando está siguiendo una estrategia en particular. Y esto lo hacemos en un contexto relativamente simple; lo hacemos en el contexto de un juego determinado sobre el que simplemente diré que la utilidad que uno recibe del juego consta de dos partes: una parte de la utilidad depende solamente de las acciones que desarrolle el otro jugador y una segunda parte de la utilidad depende solamente de las acciones que uno desarrolle. Así pues, técnicamente, existe un supuesto subyacente acerca de la posibilidad de separar las dos partes de la utilidad, lo que me permitirá describir el resultado.

La idea es que uno recibe una señal ruidosa de la otra persona. Y las estrategias que son exitosas en el muy largo plazo son las estrategias que de hecho responden solamente a una pregunta muy simple, que no intentan identificar de forma complicada exactamente qué estrategia está usando la otra persona. De hecho, las estrategias que surgen como estrategias de equilibrio en el largo plazo solamente intentan determinar si la estrategia que deben enfrentar es igual o diferente de la que el sujeto está aplicando, y si es diferente las diferencias específicas entre una y otra estrategia son irrelevantes. Así pues, la estrategia tratará al rival de forma diferente según la probabilidad de que la otra persona con la que uno está

interactuando esté usando la misma estrategia que uno o una estrategia diferente. Y quizás una interpretación de esto es una teoría de preferencia endógena por el altruismo o por la malevolencia.

Así que yo me encuentro con alguien para desarrollar una interacción y recibo una señal sobre esa persona y sobre el tipo de persona que es. Y en base a dicha señal yo puedo sacar una conclusión y puedo calcular la probabilidad condicional de que esa persona con la que me he encontrado esté usando la misma estrategia que yo, la probabilidad de que dicha persona sea del mismo tipo de persona que soy yo. Entonces calculo esta probabilidad, y lo que dice la teoría es algo muy específico acerca de lo que yo debería hacer. Debo maximizar la suma ponderada de mi utilidad y la utilidad del otro jugador, y esta ponderación dependerá de la probabilidad condicional, de la siguiente manera.

Si yo sé con un 100% de certeza que la persona es del mismo tipo que yo, empleo ponderaciones iguales para mi utilidad y para la utilidad de la otra persona. Soy perfectamente altruista hacia las personas sobre las que percibo que son perfectamente iguales a mí, las trato de la misma forma que me trataría a mí mismo. Pero si recibo pruebas concluyentes de que esta persona es diferente a mí, si estoy seguro de que esa persona es diferente a mí, uso una ponderación positiva para mi propia utilidad y una ponderación negativa para la utilidad de la otra persona; una ponderación de 1 para mi utilidad y una ponderación de -1 para la utilidad de la otra persona, de forma que estoy maximizando la diferencia entre mi utilidad y la de la otra persona.

Si recibo una señal que no me da ninguna información, de tal manera que es igualmente posible que la persona sea igual que yo o diferente a mí, usaré una ponderación de 1 para mi propia utilidad y una ponderación de 0 para la utilidad de la otra persona. Me comportaré de forma totalmente egoísta, sin tomar en cuenta el bienestar de la otra persona, sin darle un valor, ni negativo ni positivo.

Si esta probabilidad condicional se mueve en el entorno de 1, la persona es igual a mí y si se mueve en el entorno de 0 es diferente a mí, y lo que sucede es que yo me comportaré de manera más o menos altruista, en el sentido de que elegiré mi comportamiento de forma de maximizar esta combinación entre mi utilidad y la utilidad de la otra persona, tratando a la persona de manera más favorable si percibo que es más parecida a mí

o tratándola de manera más desfavorable si percibo que es menos parecida a mí.

Ahora bien, quizás esto solamente le pueda resultar sorprendente a un economista. Creo que la mayoría de la gente entiende perfectamente que uno es más amable con los familiares cercanos, por ejemplo, que con los desconocidos. Quienes hemos recibido capacitación en comportamiento social hallaremos que esta idea es transparente, pero los economistas durante mucho tiempo han considerado que la gente es totalmente egoísta, que no se preocupa por los demás, que no le dan a los demás ni un valor positivo ni un valor negativo, y quizás por este motivo este resultado pueda ser algo sorprendente para un economista.

Así quedaría redondeado el tema. Ahora, para quienes tengan mayor interés, les daré algunos detalles, es decir una descripción detallada del modelo matemático. Usaremos un modelo matemático simplificado para establecer estos resultados. Permítanme que les explique esto brevemente, porque si conocen los supuestos comprenderán mejor y más claramente las limitaciones de estos resultados en particular.

Veamos los detalles formales. Estudiamos una situación en la que hay jugadores, parejas de jugadores que se reúnen para jugar un juego de forma normal y simétrica, y estos jugadores se extraen de una única población. El mundo en el que trabajamos es, pues, un mundo con una gran población de personas idénticas, y estas personas se toman en forma aleatoria de a dos y cada par de personas juega un juego determinado. Todas las personas son iguales.

El contexto es el de un juego finito. La razón de plantear esto es simplemente que ustedes sepan que s representa la estrategia en algunos de los supuestos que se harán más adelante. Los jugadores, que son idénticos, reciben cierta utilidad que depende de dos cosas: la estrategia que están usando y σ , que representa las probabilidades de las diferentes estrategias empleadas por el jugador con el que se encuentran.

Esto es simplemente una descripción del juego. Ahora bien, lo que yo realmente quiero describir es el proceso de imitación, porque es el proceso que constituye el meollo de los resultados que tenemos. También describiré cuáles son los supuestos que son generales en la bibliografía sobre teoría de juegos evolucionista y qué es lo que diferencia nuestros

supuestos de los supuestos más tradicionales. Por el momento estoy adoptando solamente supuestos tradicionales; estoy simplificando los supuestos para que resulte más fácil analizar el modelo.

Hay un número fijo de jugadores, por lo que el mundo tiene m jugadores. Cada jugador usa una estrategia determinada. Por tanto, en cualquier momento dado habrá una habitación llena de personas, como ésta, para identificar la estrategia particular de cada persona, y luego extraeremos personas y los jugadores usarán sus estrategias contra sus oponentes. Y cada vez que jugamos, por motivos de simplicidad, lo que sucede es que una persona, solamente una persona, está habilitada para cambiar su estrategia. Y así, la forma en que esta persona cambiará su estrategia será mediante el uso de varias técnicas diferentes.

Hay tres formas diferentes en que una persona puede cambiar una estrategia. La primera forma en que una persona puede cambiar su estrategia es mediante el proceso que la teoría de juegos evolucionista ha denominado mutación. ¿Qué significa? Significa que ustedes tienen un universo de estrategias, 100 estrategias entre las que pueden elegir, y eligen una al azar. En este caso no hay ningún elemento de racionalidad en el proceso. Luego aquí lo que se describe es la innovación. La innovación se da cuando uno intenta algo nuevo que quizás nadie ha intentado. Entonces ustedes ven qué es lo que ocurre. No se trata de un proceso de racionalidad sino de un proceso de irracionalidad, aleatorio.

La segunda forma en que la gente puede elegir una estrategia es a través de un proceso de imitación. En concreto, lo que proponemos es un modelo muy simple de imitación. Imaginemos que ustedes nacen, llegan al mundo, y deben decidir cómo jugar. Una posibilidad es que ustedes no sepan nada, en cuyo caso se verán más o menos obligados a hacer una elección al azar. Pero supongamos que yo les doy algo de información, supongamos que identifico a una de las personas del grupo y les digo cómo está jugando. Ahora ustedes saben que al menos una persona del grupo cree que cooperar es una buena idea, cualquiera sea la interacción en la que esté participando. En tal caso, si ustedes no tienen ninguna otra información y ven que alguien está haciendo algo, lo lógico es que imiten a esa persona. Eso tiene más sentido que elegir al azar entre todas las posibilidades, pues al menos una persona ya creyó que ésa era una buena idea. Entonces la segunda forma es la idea de que uno elige una estrategia por imitación. ¿Cómo funcionaría? Uno elige la estrategia de forma

proporcional con respecto a la forma en que muchas otras personas están jugando dicha estrategia en la población. Así por ejemplo, si un 99% de las personas están haciendo lo mismo, cuando uno observa a una persona cualquiera lo más probable es que uno vea lo que el 99% de las personas están haciendo, y entonces uno jugará del mismo modo que el 99% de las personas. Existe una posibilidad de un 1% de que uno elija a la persona diferente que está haciendo algo diferente, en cuyo caso uno terminará imitando a esa persona.

Finalmente, debo decirles la tercera forma en que una persona podría elegir su estrategia, que es a través de un proceso más racional, y este proceso más racional estará limitado a las estrategias que los demás jugadores están de hecho utilizando. En este caso yo ingreso al mundo y no puedo calcular y determinar cuál es la mejor forma de jugar, pero sí puedo determinar qué resultados están obteniendo los diferentes jugadores con las estrategias que están utilizando. Puedo observar a Bill Gates y saber qué está haciendo y saber que es rico. Entonces, dentro del universo de personas de la habitación, yo básicamente elijo a la persona más rica y la imito. Esto es mucho más racional que elegir una persona al azar, y ahora puedo esperar que si me comporto como Bill Gates puedo llegar a ser rico. No quiero exagerar, pero es obviamente mucho mejor imitar a las personas exitosas que a quienes han fracasado. Y ésta es la tercera forma de elegir una estrategia.

Lo más probable es que uno actúe con un alto grado de racionalidad. Así que el supuesto básico en este caso es que la elección más razonable y racional es la más probable, y lo más probable que ocurra es que uno imite a una persona exitosa. Así que esto podría decirse que es una probabilidad residual. Y luego tenemos un ϵ proporcional con menor probabilidad – siempre consideramos que ϵ es un valor pequeño... Uno está obligado a elegir uno de estos métodos alternativos para elegir una estrategia. Así pues, con el ϵ proporcional de probabilidad uno imita a una persona y con la probabilidad ϵ^n , uno innova, intenta algo nuevo.

Ahora estaría en condiciones de describirles con cierto grado de claridad de qué forma este modelo difiere en forma esencial del modelo evolucionista estándar, que es la siguiente. En el modelo estándar no existe la imitación. Entonces, esta constante C sería igual a 0. No hay imitación, solamente hay innovación.

Ahora bien, ¿a qué nos referimos cuando hablamos de modelo imitativo? Lo que planteamos es que la innovación es menos probable que la imitación, y el punto central es decir que la imitación debe ser suficientemente más probable que la innovación, que la innovación debe ser mucho menos probable que la imitación. El supuesto preciso es el supuesto acerca de la magnitud del coeficiente n en comparación al tamaño de la población, que es m . Y el supuesto específico es que el coeficiente n debería ser mayor que el tamaño de la población m . ¿Y eso qué significa? Significa que cuando una persona hace algo, la probabilidad de que la totalidad de la población imite a esa persona es mayor que la probabilidad de que un integrante de dicha población tenga una idea totalmente nueva.

Ese es el punto central del supuesto. Es un supuesto fuerte, pero justamente porque es un supuesto fuerte lleva a una gran simplificación. Yo no quiero plantear que yo crea que ésta es la magnitud correcta, pero ciertamente creo que la innovación es menos probable que la imitación. No quiero plantear que en todos los casos prácticos la imitación es tanto más probable que la innovación, pero por ser un teórico debo estudiar los casos más extremos, que son más fáciles de analizar, y debo esperar que los resultados aclaren los casos más intermedios.

Y entonces ahora destacaré brevemente, como dije, de qué forma esto nos ayuda a obtener los resultados de un modelo que sea demasiado complicado como para ser estudiado, en un encuadre más simple. El motivo de ello es que básicamente la evolución es una competencia entre estrategias, de alguna forma es una guerra entre diferentes estrategias. En la teoría estándar, puede haber muchas estrategias, todas compitiendo a la vez por sobrevivir. Y puede ser muy complicado cuando muchas estrategias deben competir con muchas estrategias rivales diferentes. Pero en un mundo de imitación, es muy poco probable que una estrategia tenga que competir con más de una estrategia rival. Lo más probable es que todos terminen haciendo lo mismo. Y si todos hacen lo mismo, la imitación lo lleva a uno a hacer lo mismo que los demás.

Es solamente en el caso de la innovación que se da que alguien tenga una idea nueva que lleva a todos los demás a hacer algo diferente. Imaginemos entonces un mundo en que todo el mundo cumple lo establecido, todos juegan de la misma forma, todos hacen lo mismo, y de pronto alguien tiene una idea novedosa y propone algo diferente. ¿Qué va a ocurrir? Básicamente pueden suceder dos cosas: que todos piensen que

es una mala idea y sigan haciendo lo mismo que estaban haciendo o que haya un número suficiente de personas que imiten al innovador y se pongan a hacer lo que él propone. Pero es muy probable que todo este proceso de imitación que llevaría a otro equilibrio se produzca antes de que una persona tenga una tercera idea. Durante todo este proceso de pasaje de la idea 1 a la idea 2 es muy poco probable que la población reciba una tercera idea; de forma tal que la idea 1 y la idea 2 básicamente compiten solamente entre sí. Yo considero que esto es bastante sensato; todos sabemos que no hay millones de ideas compitiendo entre sí a la vez; en un momento dado, lo que hay son ciertas ideas clave que compiten una con la otra. Así pues, ya sea que éste sea un supuesto bueno o malo, lo cierto es que éste es el supuesto que hace que el modelo matemático sea mucho más fácil de analizar, porque podemos estudiar simplemente la oposición entre pares de ideas y no tenemos por qué preocuparnos acerca de la competencia entre un mayor número de ideas.