

APRENDIZAJE Y EVOLUCION EN TEORIA DEL EQUILIBRIO

Por: ALFREDO GARCIA MORALES*

Entendemos por equilibrio o convención un patrón de comportamiento que es esperado y que se ratifica por sí solo. Es decir, todos los agentes se conforman con el equilibrio, todos esperan que los otros se conformen con ese equilibrio y todos se conforman con el e-

quilibrio, dado que los otros lo hacen. En nuestra vida cotidiana tenemos diversos ejemplos, manejar del lado derecho de la calle cuando otros lo hacen, aceptar papel moneda si otros lo hacen, ser poco corteses en el conducir cuando otros también lo son etc...

nos aclara el por qué de la robusteza o estabilidad de la Convención o situación de Equilibrio frente a pequeñas perturbaciones, es decir, agentes que deciden ir por fuera de la convención, pero no nos explica cómo se puede llegar a esa situación.

El enfoque de aprendizaje asume que los agentes "juegan" a distintas convenciones, cada uno de los agentes posee memoria finita que le permite ir guardando el recuerdo de aquellas convenciones o situaciones en las que salió mejor librado.

En cada una de estas situaciones, existe un patrón de comportamiento esperado y cada uno de nosotros decide cómo actuar teniendo en cuenta cómo actúan los otros.

Por qué estas convenciones persisten?

Simplemente porque si alguno de los agentes decide "desviar", es decir, tomar una decisión diferente a la establecida en la situación de equilibrio, aisladamente, la situación resultante no será la más favorable. Por ejemplo, si uno no es agresivo conduciendo en Bogotá, nunca llegará a su destino. Esta explicación educativa es lo que se conoce en Teoría de Juegos como Equilibrio de Nash (1).

Pero este concepto solo

Para explicar cómo se llega a una Convención o situación de Equilibrio examinaremos dos enfoques: Aprendizaje y Evolución. Ambas teorías son de naturaleza dinámica, es decir, aducen que una Convención o Situación de Equilibrio es fruto de un proceso dinámico en el cual los agentes evalúan y repiensen sus decisiones, hasta llegar a una situación particular con un grado de atracción tal que es estable a pequeñas perturbaciones como las que acabamos de ilustrar.

El enfoque de Aprendizaje asume que los agentes "juegan" a distintas convenciones, cada uno de los agentes posee una memoria finita que le permite ir guardando el recuerdo de aquellas convenciones o

situaciones, en las que salió mejor librado. En la siguiente iteración, la decisión de cada agente no solo tendrá en cuenta la situación actual sino que tendrá a su disposición en memoria la experiencia de las pasadas iteraciones y de acuerdo a estas tomará la que mas le convenga. Este proceso de memorización y discernimiento es nuestro modelo simplificado de aprendizaje. Quizás, el ejemplo mas famoso de esta familia de modelos sea el proceso de "tatonnement" o ajuste con el que Cournot explicaba la situación de equilibrio en cantidades a producir en oligopolio. En este ejemplo, las firmas tienen una memoria muy corta y solamente retienen las acciones tomadas por los otros agentes en el período inmediatamente anterior, lanzan al mercado en cada iteración una cantidad, y de acuerdo a la ley de demanda se establecen beneficios para cada una de ellas. Entonces, cada firma asumiendo que las otras firmas repetirán los niveles de producción del período precedente, evalúa cuál es la cantidad óptima a producir. Cournot aducía que este mecanismo llevaba a la situación de equilibrio del mercado. Lo cierto es que bajo supuestos,

este proceso conduce al equilibrio.

Estos supuestos de índole técnico que no quiero detallar, exigen en cierta forma que cada firma al optimizar de manera miope o con muy poca memoria, y de manera aislada e individual, conduzcan a nivel global hacia la situación de bienestar general.

Aunque para el caso de oligopolio este modelo simplísimo de aprendizaje parece funcionar, en otras instancias es completamente errático e inestable. Otro aspecto que escapa al modelo de Cournot es el llamado supuesto de la Racionalidad "Acotada o Limitada". Los agentes bien sea sociales o económicos toman decisiones que no son necesariamente "óptimas", puesto que el cálculo de estas puede ser engorroso. Sin embargo, sus acciones si llevan, por lo menos ex-post, un intento de optimización implícito.

En nuestro marco de aprendizaje, este supuesto vendría a traducirse en la posibilidad de no tomar a cada iteración la acción "óptima", sino otra relativamente buena.

Y lo que es más importante, la posibilidad de imitar o copiar acciones de otros, que han sido exitosas. Estas dos posibilidades permiten a los agentes cometer pequeños "errores" en la evaluación de sus acciones, y aunque parezca insólito, no permiten que el proceso de aprendizaje se estanque en una convención o situación de equilibrio que bien no sea la mejor globalmente o de "Pareto", puesto que esta es la mas robusta a "errores" como los mencionados.

En el enfoque evolutivo, lo que se trata es de aplicar el marco conceptual de las Teorías Evolutivas en Biología. De manera apresurada, diremos que las especies se reproducen o sobreviven de manera proporcional a su grado de adaptación al medio. Mientras más adaptado más chance de reproducirse y sobrevivir. Adicionalmente, existe la posibilidad de pequeñas mutaciones de naturaleza aleatoria que modifican y renuevan el patrimonio genético de las especies que se reproducen (las mas adaptadas).

La idea entonces es que el efecto combinado de es-

tas dos operaciones, selección y mutación, explican la evolución y diversidad de las especies en nuestro mundo.

En el intento de explicar el origen de las Convenciones o Situaciones de Equilibrio, identificamos las estrategias empleadas por los agentes como las "especies", a estas según su evaluación, es decir, qué tanto beneficio depararon a sus ejecutores, aplicaremos un criterio de "selección y mutación" como sigue: al inicio del proceso iterativo todas las estrategias posibles se encuentran disponibles, los agentes escogen cada uno, una estrategia, y a esta "población" le aplicamos un criterio de "reproducción" proporcional al beneficio deparado por ella, en otras palabras, en la siguiente iteración aquellas "malas", no estarán disponibles para el agente que decide, las "buenas", en cambio permanecen disponibles. Lo que se tiene entonces es que en cada nueva iteración el espacio de estrategias posibles se va reduciendo y homogenizando, permaneciendo en él, aquellas que seguramente conformarán la Convención o Situación de Equilibrio. Es importante

también resaltar el rol cumplido por la mutación, que es análogo al de los pequeños "errores" en el modelo de aprendizaje. La mutación de estrategias, evita una preconvergencia o convergencia prematura a una convención no muy robusta. En breve, una Convención o Situación de Equilibrio se presenta aquí como una situación estable a pequeños cambios o mutaciones en las estrategias previstas y su origen se explica a partir de un proceso evolutivo de selección del más "fuerte" o "adaptado".

Estos dos enfoques han permitido justificar situaciones de Equilibrio social y económico que son aparentemente "irracionales" (como la forma de conducir en Bogotá), que aunque no parezcan satisfacer el bienestar general permanecen inmutables.

Para ilustrar la anterior exposición consideramos el famoso "dilema de los prisioneros". Dos sospechosos son capturados por un cierto delito. Cada prisionero es alojado en una celda completamente incomunicado del otro. El detective, que aún no tiene pruebas contundentes, pe-

ro esto no es de conocimiento de los prisioneros, les ofrece a cada uno por separado una rebaja de pena si delata al otro. Es evidente, que la situación de Pareto o bienestar general será la de no delatarse puesto que no se les podrá condenar y deberá indemnizarseles por la encarcelación arbitraria. Si ambos se delatan, ambos serán condenados pero a una pena menor. En el siguiente tablero trataré de ilustrar estas situaciones:

	Delatar	No Delatar
Delatar:	(-1,-1)	(0,-2)
No Delatar:	(-2,0)	(1,1)

Entonces, si por ejemplo ambos se delatan cada uno "pierde" -1, pues serán condenados, si uno delata y el otro no, el que delata no "pierde" nada o gana cero, y el que no delató es condenado y "pierde" -2, la pena mayor.

Noten que existen dos Convenciones o Situaciones de Equilibrio: "Ambos delatan" y "Ambos no delatan". Consideremos, por ejemplo, la situación "Ambos delatan": si alguno de-

cidiese no delatar, sufrirá una pena mayor (-2 en vez de -1) y por tanto no lo hace. De la misma forma, "Ambos no delatan" es robusta a cambios aislados, si alguno decidiese delatar dejaría de percibir la indemnización (con valor + 1).

El "dilema" está en que amplias y muy serias investigaciones prácticas (1), (tomar dos personas y ponerlas a decidir como si fueran los prisioneros) muestran que la gran mayoría de nosotros, escojemos la situación "Ambos delatan" que paradójicamente, no es la de bienestar general. Para no ir tan lejos, consideren la situación del tráfico en Bogotá, en este caso "delatar" es "ser descortés o agresivo conduciendo" y no delatar es "ser cortés". La situación que ratifican día a día los bogotanos es "ser descorteses y agresivos" en el conducir.

Como explicar esto? El enfoque "aprendizaje" nos diría que nuestra memoria es demasiado corta y por tanto, aunque cada mañana al levantarse los bogotanos "tomen" conciente o inconcientemente la decisión de cómo manejar, al día siguiente "olvidarán" su experiencia de ayer y si ésta, por ejemplo, fue encontrarse con mucha gente cortés en la ruta, será desechada por olvido. No "aprendemos", entonces, que hay otra Convención o Situación de Equilibrio, aún mejor en el sentido de Pareto, es decir, para todos. El enfoque "evolutivo" por su parte nos diría que de cierta forma está "pagando" mas ser descortés o agresivo y que de esta manera esa estrategia es la mas "adaptada" y por lo tanto persiste.

Para concluir, es este punto en donde estas teo-

rías de naturaleza descriptiva pueden servir como herramientas de evaluación de decisiones a tomar por el planificador social, que busca el bienestar general. Por ejemplo, en el caso del tráfico de Bogotá, la modificación de la estructura de recompensas (cuanto por delatar, etc., algo así como punir la descortesía) podría ser una medida para incentivar el mecanismo de aprendizaje de los agentes sociales, por muy corta que tengan la memoria.

Bibliografía:

- (1) FUDENBERG D. Tirole J. "Game Theory" MIT Press 1.993.
- (2) YOUNG, Peyton. "The Evolution of Conventions" Econometrica, Vol 61 Enero 1.993. pp. 57 - 84

* El autor es Ingeniero Eléctrico. Estudiante de Doctorado en Investigación de Operaciones. Universidad de Michigan.