

William Poundstone

# El dilema del prisionero

John von Neumann,  
la teoría de juegos y la bomba



**Alianza** editorial  
El libro de bolsillo

Título original: *Prisoner's Dilemma. John von Neumann,  
Game Theory and the Puzzle of the Bomb*  
Traductor: Daniel Manzanares Fourcade

Primera edición: 1995  
Tercera edición: 2015  
Primera reimpresión: 2018

Diseño de colección: Estudio de Manuel Estrada con la colaboración de Roberto Turégano y Lynda Bozarth  
Diseño de cubierta: Manuel Estrada  
Fotografía de Amador Toril

Reservados todos los derechos. El contenido de esta obra está protegido por la Ley, que establece penas de prisión y/o multas, además de las correspondientes indemnizaciones por daños y perjuicios, para quienes reprodujeren, plagiaren, distribuyeren o comunicaren públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier medio, sin la preceptiva autorización.

© 1992 Copyright by William Poundstone  
© Alianza Editorial, S. A., Madrid, 1995, 2018  
Calle Juan Ignacio Luca de Tena, 15  
28027 Madrid  
[www.alianzaeditorial.es](http://www.alianzaeditorial.es)

ISBN: 978-84-206-9396-5  
Depósito legal: M. 26.645-2014  
Printed in Spain

Si quiere recibir información periódica sobre las novedades de Alianza Editorial, envíe un correo electrónico a la dirección: [alianzaeditorial@anaya.es](mailto:alianzaeditorial@anaya.es)

# Índice

- 15 Agradecimientos
  
- 17 1. Dilemas
  - 22 El dilema nuclear
  - 24 John von Neumann
  - 30 El dilema del prisionero
  
- 33 2. John von Neumann
  - 35 El niño prodigio
  - 38 Hungría en la época de Béla Kun
  - 41 Los comienzos de su carrera
  - 44 El Instituto
  - 47 Klara
  - 50 Su carácter
  - 62 La fase *Sturm und Drang*
  - 69 El mejor cerebro del mundo
  
- 75 3. Teoría de juegos
  - 76 Kriegspiel
  - 81 ¿Quién fue el primero?
  - 83 Teoría de juegos y comportamiento económico
  - 85 Repartirse un pastel
  - 86 Jugadores racionales
  - 88 Los juegos son como árboles

93	Los juegos como tablas
98	Juegos de suma cero
100	Minimax y el pastel
105	Estrategias mixtas
112	Bolas con efecto y genes asesinos
116	El teorema minimax
118	Juegos de $n$ personas
122	4. La bomba
123	Von Neumann en Los Álamos
126	La teoría de juegos en la guerra
129	Bertrand Russell
132	Un gobierno mundial
136	La operación Crossroads
140	El ordenador
144	La guerra preventiva
150	5. La organización RAND
153	Historia
163	Pensar sobre lo impensable
168	El surf, la semántica y la fonología finlandesa
172	Von Neumann en RAND
175	John Nash
177	Comentando el partido el lunes por la mañana
181	6. El dilema del prisionero
182	La venta de un Buick
185	El honor de los ladrones
190	El experimento de Flood y Drescher
202	La anécdota de Tucker
210	Sentido común

- 213 Los dilemas del prisionero en la literatura  
218 El viajero gratis  
224 La carrera de armamentos nucleares
- 228 7. 1950  
229 La bomba soviética  
233 El hombre que vino de Marte  
234 El discurso de Urey  
237 El caso Fuchs  
242 La guerra de Corea  
243 Las implicaciones de los avances tecnológicos  
en armamento  
249 Agresores por la paz  
253 Francis Matthews  
257 Las consecuencias finales  
261 La reacción de la opinión pública  
267 ¿Acaso fue un sondeo de opinión?  
268 El discurso de MacArthur  
270 Orvil Anderson  
273 La reacción de la prensa  
276 ¿Cuántas bombas hacen falta?  
284 Epílogo
- 288 8. La teoría de juegos y sus detractores  
289 Críticas a la teoría de juegos  
293 Maquiavelo y el utilitarismo  
296 ¿Es racional la gente?  
300 Los análisis del Estado de Ohio
- 308 9. Los últimos años de Von Neumann  
309 La bomba de hidrógeno

311	Un tigre muy hermoso
314	Comisionado
320	El momento de la esperanza
325	La enfermedad
335	Muerte
337	10. El juego del gallina y la crisis de los misiles de Cuba
340	Gallina
348	El dilema del voluntario
351	Experimentos con el dilema del voluntario
353	La crisis de los misiles de Cuba
367	La táctica del loco
370	11. Otros dilemas de ámbito social
374	El atolladero
375	La caza del venado
380	Juegos asimétricos
382	La justificación de la cooperación
388	El superjuego de Howard
392	La paradoja de la inducción retroactiva
397	12. La supervivencia de los más aptos
398	Estrategias estabilizadas
402	¿Es genético el interés propio?
406	Robert Axelrod
413	«Donde las dan, las toman»
419	Las dificultades de «donde las dan, las toman»
425	Selección artificial
429	El pez que se mira en el espejo
434	Cooperación y civilización
437	«Donde las dan, las toman» en la vida real

442	13. En puja por el dólar
444	La carrera armamentística
448	La subasta del dólar de Shubik
452	En puja por el dólar en la vida real
459	Estrategias
463	Pujar racionalmente
467	En qué momento falla la teoría de juegos
471	El juego del número más grande
479	La pluma en el vacío
483	Bibliografía
489	Índice analítico





*A Mikolas Crews*



# Agradecimientos

Muchos de los miembros pertenecientes al grupo original de expertos en teoría de juegos de la organización RAND continúan trabajando.

Merril Flood y Melvin Dresher aportaron datos importantes para este libro concernientes a su labor allí y al ambiente que existía.

Gran parte del material biográfico de John von Neumann, incluyendo las citas de su correspondencia, proviene del conjunto de papeles de Von Neumann existentes en el Departamento de Manuscritos de la Biblioteca del Congreso, en Washington, D.F.

Parte de los datos históricos sobre la Administración Truman fueron consultados en la Biblioteca Presidencial Harry S. Truman, en Independence, Missouri.

Mis agradecimientos también deben ser para Paul Armer, Robert Axelrod, Sally Bedrow, Raoul Bott, George B. Dantzig, Paul Halmos, Jeane Holiday, Cuthbert

Hurd, Martin Shubik, John Tchalenko, Edward Teller y Nicholas A. Von Neumann, por las narraciones de sus recuerdos, por su ayuda y por sus consejos.

# 1. Dilemas

Un hombre iba a cruzar un río con su mujer y su madre. Apareció una jirafa en la orilla opuesta. El hombre sacó su rifle y apuntó, mas la jirafa le dijo: «Si disparas, morirá tu madre. Si no disparas, morirá tu mujer». ¿Cuál debería ser el comportamiento del hombre?

Así se narra un cuento tradicional que contiene un dilema, tal y como lo relata el jeque de Dahomey. En el folclore africano abundan disyuntivas excéntricas y difíciles de este tipo. Muchas han sido recogidas por filósofos y escritores occidentales para su uso propio. En la fábula del jeque hay que suponer que todo cuanto diga la jirafa parlante es verdadero. Se puede replantear el dilema en términos más apropiados al entorno y la lógica occidentales: unos científicos locos le han raptado a usted junto con su madre y su mujer, y les han encerrado en una habitación con una máquina de aspecto extraño. Los tres están atados a sillas. Al alcance de usted hay un pulsador.

Una ametralladora se yergue amenazadora delante de su mujer y su madre, y hay un reloj en la pared que marca los segundos. Uno de los científicos le explica que si oprime el pulsador, la máquina apuntará a su madre y la matará. Por el contrario, si no pulsa el botón en un lapso de sesenta segundos, el artefacto apuntará a su mujer y disparará. Tras una inspección detenida, se ha dado cuenta de que la máquina funcionará irremisiblemente como se le ha explicado. ¿Qué debería hacer?

Tales dilemas se discuten a veces en las clases de ética de los institutos. Evidentemente, no hay una solución adecuada. Sería evadir su responsabilidad si arguye que no tomaría ninguna acción (es decir, no pulsar el botón, con lo que la máquina mataría a su cónyuge), basándose en que no podría ser culpable por no hacer nada. Sólo puede decidir cuál de las dos personas aprecia más, y salvarla.

Las decisiones son aún más difíciles cuando otro también está decidiendo, y el resultado depende del conjunto de todas las decisiones tomadas. En el libro *The Book of Questions* (1987), de Gregory Stock, hay un dilema parecido pero que incita a una mayor reflexión: «Usted y una persona amada son situados en habitaciones separadas provistas de un pulsador. Saben que matarán a ambos a no ser que uno pulse el botón antes de una hora; además, la primera persona que accione el pulsador salvará a la otra, pero morirá inmediatamente. ¿Qué decisión tomaría usted?».

En este caso hay dos personas deliberando sobre su situación que realizan decisiones independientes. Es necesario que haya alguien que pulse el botón. Lo difícil es

decidir cuándo deberá sacrificarse. El dilema le obliga a tomar una «decisión de tipo bote salvavidas» que afecta a la persona a quien usted quiere y a sí mismo. ¿Quién debería ser el que sobreviviese?

Existen muchas situaciones en las que uno decide salvar a otro a expensas de perder la propia vida. Un padre o madre podría salvar a un niño basándose en que el niño, dada su juventud, tiene mayores expectativas de vida. Sea el que fuere el criterio usado, y sabiendo que ambas personas no tendrían el mismo, hay tres desenlaces posibles al tomar una decisión de tipo «bote salvavidas».

El caso menos angustiante es cuando ambos coinciden en quién debe sacrificarse y quién salvarse. Entonces aquél debería pulsar el botón para así salvar al otro.

Una segunda posibilidad es que ambos decidan salvarse el uno al otro. Una madre decide salvar a su hija, a la que le quedan más años de vida, y la hija decide asimismo salvar a su madre, pues ésta le dio la vida. En este caso se compite por ser el primero en pulsar el botón.

La opción más conflictiva surge cuando *ambas* personas deciden que ella misma es la que debe salvarse. Entonces nadie pulsa el botón, y el reloj marca el tiempo que pasa inexorablemente...

Revisemos este caso algo más a fondo. Según el reloj, han pasado cincuenta y nueve minutos. Usted no ha oprimido el pulsador con la esperanza de que su ser querido lo hiciera, pero no ha sido así. (Presupongamos que la noticia de la muerte de uno es comunicada inmediatamente al superviviente en cuanto éste pulsa el botón.) Ha tenido usted tiempo sobrado para considerar todas

las posibilidades. Ciertas personas pueden tardar la hora entera para decidir quién se salvará, o para reunir el valor suficiente para pulsar el botón. Aun teniendo en cuenta estas posibilidades, parece que la persona que usted adora ha decidido que sea *usted* el que se sacrifique.

No tiene sentido comprometerse a no pulsar *jamás* el botón, ni siquiera en la última fracción de segundo. Por muy egoísta que usted sea, no puede salvarse a sí mismo. Alguien ha de morir; así es como se ha planteado el dilema. Si la persona a la que ama no se presta a sacrificarse, por lo menos sálvela usted. Pues no olvide que ama de verdad a esa persona.

La situación ideal sería que pulsara usted el botón en el último instante. Es posible que la otra persona piense lo mismo. Precisamente esto le da fuerzas para resistir hasta el último momento, y además da a esa persona la oportunidad de pulsar el botón en el penúltimo instante; si no fuera así, entonces, y sólo entonces, pulsaría usted su botón. Por supuesto, es posible que el ser amado planee esta *misma* estrategia.

Al complicarse la situación con este planteamiento mutuo de resistir hasta el último momento posible, el desenlace dependerá de los tiempos de reacción y de la precisión del reloj. La malvada máquina no tiene en cuenta la *intención* de haber pulsado el botón dentro del margen de tiempo estipulado. O bien usted o bien la otra persona han de cerrar mediante su actuación el circuito, antes de que la máquina se ponga en marcha y les mate a los dos. Además, no se sabe a ciencia cierta si el reloj de la pared está perfectamente sincronizado con la máquina. Eso dijo el científico loco, mas quizá no se re-



fería al segundero, y recuerde que después de todo es un chiflado. Para asegurarse de poder pulsar el botón dentro de un margen de seguridad, debería en realidad pulsarlo un instante o dos antes. ¡Sin embargo, este requerimiento es contrario a su intención de resistir hasta el último momento!

Tanto la persona a la que quiere, y que parece aplazar su decisión, como usted, están en el mismo barco. Si ambos decidieran pulsar el botón en el último instante, el desenlace sería resultado del azar. Uno de los dos pulsaría una fracción de segundo antes que el otro, o bien ambos se pasarían del tiempo límite y por tanto morirían. El desenlace se determinaría por casualidad. La suerte primaría frente a la racionalidad.

En el ámbito de la filosofía, las situaciones en las que se han de tomar decisiones límite encerrado en una habitación desconocida han despertado el suficiente interés como para que se las bautice con el nombre de «cajas de dilemas». ¿Por qué tales dilemas llaman tanto la atención? En parte se debe a la curiosidad que tenemos ante cualquier situación nueva. Pero no despertarían tanto interés si sólo fueran rompecabezas que no influyeran en nuestra existencia personal.

Los dilemas de la vida real, que no son planteados por científicos dementes, surgen gracias a las diversas maneras con las que nuestros intereses individuales se debaten con los de otros y con los de la sociedad en general. Diariamente debemos tomar decisiones difíciles, a veces con resultados distintos de los que habíamos esperado. La cuestión esencial que se plantea es simple y apremiante: ¿existe un comportamiento racional para cada situación?

## El dilema nuclear

La URSS hizo explotar su primera bomba atómica en Siberia, en agosto de 1949. De esta manera terminaba el monopolio americano sobre armas nucleares. Mucho antes de lo esperado por los observadores occidentales, había ya dos potencias nucleares.

La bomba soviética provocó el comienzo de una carrera armamentística nuclear. Algunas consecuencias eran fáciles de prever. Cada nación se armaría hasta el punto en que pudiera lanzar una ofensiva atómica rápida y devastadora sobre la otra. Muchos pensaron que se planteaba así un dilema inabordable. Por vez primera en la historia del mundo se contemplaba la posibilidad de un ataque sorpresa que borraría a toda una nación de la faz de la Tierra. Si surgía un conflicto, la tentación de pulsar el botón nuclear sería irresistible. Era igual de importante darse cuenta de que cada nación temería ser *la víctima* del ataque sorpresa de la otra potencia.

Hacia 1950, muchas personas de los Estados Unidos y de Europa occidental pensaban que América debería considerar la posibilidad de llevar a cabo un ataque sorpresa, no provocado, sobre la Unión Soviética. La idea, llamada eufemísticamente «guerra preventiva», afirmaba que, aprovechando la oportunidad, los Estados Unidos deberían establecer un gobierno mundial, bien mediante la amenaza de un ataque –es decir, hacer chantaje nuclear– o bien por ataques sorpresa. Podría pensarse que sólo una minoría compuesta por chiflados apoyaría tal propuesta. De hecho, encontró apoyo entre gentes de innegable inteligencia, entre ellos dos de los mejores ma-

temáticos de todos los tiempos: Bertrand Russell y John von Neumann. Los matemáticos no suelen dar a conocer sus opiniones políticas, o sobre asuntos mundanos de otro tipo. Bertrand Russell y John von Neumann tenían muchos rasgos y actitudes bastante diferentes, pero coincidían en la opinión de que no había lugar suficiente en el mundo para dos potencias nucleares.

Russell, que fue uno de los apologistas más importantes del movimiento a favor de la guerra preventiva, apoyaba la idea de un ultimátum dirigido a la Unión Soviética, en el que se amenazaría con un ataque nuclear a no ser que se sometiera a ser gobernada por un gobierno mundial dominado por los Estados Unidos. En un discurso realizado en 1947, Russell dijo: «Me inclino a pensar que Rusia aceptaría; si no es así, y si actuamos con prontitud, el mundo podría sobrevivir a la consiguiente guerra y surgiría un gobierno mundial, tal y como lo necesitamos».

Von Neumann llegó aún más lejos, pues propugnaba un ataque nuclear relámpago por sorpresa. La revista *Life* citaba a Von Neumann: «Si me propone usted bombardearles mañana, yo le contesto ¿por qué no hoy? Si dice usted que hoy a las cinco de la tarde, yo le digo ¿por qué no a la una?».

Ninguno de ellos tenían aprecio por la Unión Soviética. Sin embargo, pensaban que la guerra preventiva era, sobre todo, una cuestión de planteamiento lógico, la única solución racional al mortal dilema de la proliferación de armas nucleares. Así lo manifestaba Russell en un artículo que abogaba por la guerra preventiva, en el número de enero de 1948 de la revista *New Commonwealth*:

«El argumento que he desarrollado es tan simple e inexorable como una demostración matemática». Pero la misma lógica puede enmarañarse. No hay palabras que mejor hayan recogido el enrarecido ambiente de todos los acontecimientos de la guerra preventiva que la impremeditada y orwelliana frase del secretario de la Marina de los Estados Unidos, Francis P. Matthews, que en 1950 instaba a la nación a convertirse en «agresores por la paz».

Hoy, cuando la distensión Este-Oeste es bien patente, la guerra preventiva parece una aberración producida por la Guerra Fría. Sin embargo, actualmente existen las mismas disyuntivas. ¿Qué debería hacer una nación al chocar su seguridad con el bien de la humanidad? ¿Qué debería hacer una persona cuando sus intereses se oponen a los del bien común?

## John von Neumann

Quizá el mejor ejemplo del planteamiento del angustioso dilema de la bomba lo personifica John von Neumann (1903-1957). El nombre nada dice a la mayoría de las personas. El matemático célebre es una especie casi inexistente. Los pocos profanos a los que les sea familiar, le ubicarán como el pionero de la calculadora digital electrónica, o como uno más del grupo de mentes brillantes que trabajaron en el proyecto Manhattan. Unos pocos le reconocerán como uno de los posibles modelos en que se inspira el personaje del profesor Strangelove, de la película de Stanley Kubrick (1963).

Desde luego, fue Von Neumann el que asistió a las reuniones de la AEC (Comisión de Energía Atómica) en silla de ruedas\*.

El cuerpo principal de la obra de Von Neumann, aquel por el que adquirió al principio de su carrera la reputación de un genio, se halla en los dominios, inaccesibles al profano, de la matemática pura y la física matemática. Pudiera suponerse que, a lo largo de su vida, su trabajo estuvo alejado de los asuntos mundanos. Sin embargo, a Von Neumann le apasionaba la matemática aplicada. Tanto la calculadora electrónica como la bomba fueron para Von Neumann proyectos no académicos, pero son buenos ejemplos de su interés en las aplicaciones de las matemáticas.

Von Neumann jugaba al póquer, sin ser muy competente. Su mente ágil se fijó en determinados aspectos del juego. Le interesaba sobre todo el engaño, el faroleo y las segundas intenciones, es decir, la manera en que los jugadores trataban de dar pistas falsas, usando las reglas aceptadas. Si se emplea la jerga propia de los matemáticos, existía en ello algo que era «no trivial».

Desde mediados de los años veinte hasta los años cuarenta, Von Neumann se entretuvo en investigar la estructura matemática del póquer y otros juegos. A medida que sus estudios tomaban forma, se dio cuenta de que los teoremas podían aplicarse a la economía, la política,

\* La película mencionada es *Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb?*, dirigida por Stanley Kubrick, 1963 [*Teléfono rojo: ¿volamos hacia Moscú?*]. En ella el personaje del profesor Strangelove –cuyo nombre se traduce literalmente como ‘amor aberrante’ (... a la bomba)– aparece siempre en silla de ruedas.