

EL DILEMA DE LOS PRISIONEROS Y LA MORAL

CARLOS ROSENKRANTZ
SADAF

La pregunta de por qué debo ser moral ha sido siempre una importante cuestión filosófica. El problema de la justificación racional de la restricción de mis acciones para preservar los derechos de otros es uno de los temas filosóficos más difíciles y que tradicionalmente ha dividido a la teoría ética. Los escépticos argumentan que no hay razones para ser moral, porque la moral es sólo una cuestión de preferencias; en consecuencia, la decisión de ser moral no puede ser producto de razones de ningún tipo. Otros filósofos han argumentado que existen razones para ser moral y ellas son, sorprendentemente, razones de autointerés. Éstos dicen que es posible probar que es en nuestro interés actuar en forma moral. Sostienen que el hecho de que algo esté en mi interés, dada la actitud práctica hacia mis intereses, es suficiente para calificar ese hecho como una razón. Por ello hay razones para ser moral. Estos filósofos han usado el dilema de los prisioneros (DP), como una herramienta para probar que su razonamiento es plausible; que desde un punto de vista autointeresado es mejor ser moral que inmoral. En este trabajo intentaré mostrar que basar la moral en razones autointeresadas falla porque las enseñanzas del dilema de los prisioneros no son las que comúnmente se sostienen.

Sin embargo esto no significa que no haya razones para ser moral. Por el contrario, creo fervientemente que hay razones y que son poderosas. En la última parte del trabajo mencionaré cuáles son esas razones.

Algunas elucidaciones son necesarias antes de encarar el nudo de la cuestión.

Lo que impugnaré en este trabajo es el intento de resolver lo que se ha denominado “el problema motivacional de la moral” recurriendo a razones autointeresadas. En otras palabras, el intento de explicar por qué debo ser moral con el hecho de que ser moral me beneficia. No me detendré en el intento de resolver el “problema sustancial de la moral” recurriendo a esa clase de razones. Es decir, el intento de contestar la cuestión del contenido de la moral recurriendo a razones autointeresadas. Sin embargo, quiero señalar que la plausibilidad de este intento depende de la plausibilidad de la respuesta autointeresada al “problema motivacional”. Si fuese verdad que es en mi interés actuar moralmente, sería posible decir que el contenido de la moral está dado, por lo que ésta es mi interés. Pero si esto no fuese verdad, como sostendré, serían necesarias otro tipo de razones para justificar la respuesta autointeresada al problema sustancial.

También quiero destacar que concederé en este trabajo, en beneficio de la tesis que impugno, que las interacciones sociales son como interacciones en DP. Presupondré que si fuese verdad que razones autointeresadas son suficientes para resolver esquemas de DP, estas razones serían suficientes para basar la moral.

El dilema

El DP es una formulación abstracta de situaciones muy extendidas en la realidad en las cuales lo que es me-

jor para cada persona lleva a desventajas para todos, incluidas esas mismas personas, pudiendo todos estar mejor si todos cooperasen absteniéndose de hacer lo que es mejor para cada persona individualmente. Así, cada conductor conduce más rápido si desobedece las señales de tránsito aunque de hecho va más lento si todos los demás también desobedecen esas señales. Cada soldado está mucho más seguro si dispara tan rápido como puede en el frente, pero ciertamente terminará muerto si todos disparan como él. Cada pescador obtiene más pescados si pesca más de lo permitido pero menos si todos hacen lo mismo.

Una típica formulación abstracta de estas situaciones es la siguiente:

Dos hombres sospechosos de haber cometido un crimen juntos son arrestados por la policía y colocados en celdas separadas. Cada sospechoso puede confesar o permanecer en silencio sabiendo cuáles son las consecuencias de cada uno de estos actos. Ellas son: 1) si un sospechoso confiesa y su compañero no lo hace, el que confiesa sale libre y el otro va a la cárcel por 20 años; 2) si ambos confiesan, ambos van a la cárcel por 5 años; 3) si ambos permanecen en silencio los dos van a la cárcel por un año por haber tenido armas escondidas, un cargo menor. Se asume que no hay "honor entre delincuentes" y que el único interés de cada sospechoso es su solo autointerés. Bajo estas condiciones, ¿qué deben hacer estos criminales?¹

Los rasgos centrales de DP son las particulares relaciones entre los resultados de las cuatro posibles situaciones que resultan de las dos opciones disponibles para cada jugador y el hecho de que los prisioneros no pueden salir del dilema explotándose recíprocamente porque esa opción (por definición de explotar y ser explotado) no

¹ Véase Morton Davis, *Theory. A Nontechnical Introduction*, Basic Books, Nueva York, 1983.

es tan buena para cada jugador como la cooperación mutua.²

El argumento básico de la cooperación

Algunos filósofos han sostenido, con el objeto de basar la moral en razones autointeresadas, que en el esquema del DP no es verdad lo que así parece a primera vista. Afirman que la búsqueda racional del autointerés no genera consecuencias nocivas para el puro autointerés porque un prisionero autointeresado permanecería en silencio para obtener la condena más leve.³ Los prisioneros razonarían de la siguiente manera:

- 1) Yo sé que cada uno de nosotros sabe que ambos somos racionales.
- 2) Yo sé que nuestras situaciones son absolutamente simétricas y consecuentemente nuestra conducta debe ser exactamente simétrica también porque somos exactamente racionales y autointeresados.
- 3) Yo sé que hay una razón *prima facie* para confesar, porque hacerlo me concedería el mejor resultado sin importar lo que haga mi compañero delincuente. Sin embargo, veo como consecuencia de (1) y de (2) que la única conducta posible es que ambos confesemos o que ambos no confesemos. Como ambos preferimos las consecuencias de que ninguno confiese sobre las consecuencias de que ambos confesemos, permaneceremos sin confesar.
- 4) Por consiguiente yo no confesaré.

² Véase Amartya Sen, "Reply to Comments", en Stephan Körner (ed.), *Practical Reason*.

³ Véase R. D. Luce y M. Raiffa, *Games and Decisions*, Wiley, Nueva York, 1957.

Esto es lo que ha sido llamado “el argumento básico para la cooperación”, un argumento que recurre sólo a razones de autointerés. Cada prisionero se rehúsa a perseguir la maximización de su utilidad (lo racional), pues como ser racional, sabe que su intento de maximizar su utilidad arrojará graves consecuencias.

Sin embargo, tan pronto cada prisionero decide no confesar, advierte que si lo hace quedará desprotegido frente a la confesión de su compañero. Para protegerse él tiene también que confesar aunque sabe que si lo hacen ambos, ambos estarán peor. Por ello cada uno piensa que como ambos son racionales ninguno confesará. Cada prisionero se tranquiliza nuevamente, pues lo racional es confiar en el “argumento básico” para cooperar; ser moral y no confesar.

El “argumento básico” mostraría, si fuese correcto, que es posible cooperar y actuar moralmente recurriendo a razones autointeresadas. Pero el argumento no es sano. Sobresimplifica el problema porque presupone que sólo son posibles conductas simétricas. Presupone que cada prisionero asume que la acción del otro prisionero será una función de su propia acción y coincidirá con ella.

Esto será posible siempre y cuando exista una alternativa fija “prescrita por motivos de autointerés” que pueda servir de base para predecir qué es lo que el otro jugador hará. Pero la única alternativa impuesta por razones de autointerés, y que por lo tanto puede considerarse fija, es confesar, dado que, mediante ello, cualquiera que sea la acción del otro, uno maximiza su interés (argumento dominante). Consiguientemente, en DP, prisioneros interesados y racionales no podrán decidirse por la solución mutuamente más beneficiosa, que es la decisión de cooperar.

La paradoja

La razón por la cual los prisioneros no pueden obtener los mejores resultados no es su racionalidad sino su egoísmo. La paradoja del DP se produce al pensar en maximizar la utilidad de cada prisionero, lo que es autofrustrante, pues cada uno termina con menos utilidad de la que tendría si actuase sin pensar en la maximización de su utilidad. Cuando no pienso en mi utilidad y sé que mi contraparte tampoco lo hará, puedo no decidir basándome en el "argumento dominante" y actuar de la manera que dé a ambos mejores resultados. Como puede ser fácilmente advertido, este escape es producto de la modificación de uno de los rasgos que califica a los prisioneros del dilema: su autointerés. No pensar en mi utilidad significa actuar altruistamente.

Pero, ¿es verdad que un actor racional y autointeresado no puede obtener mejores o al menos iguales resultados que un actor racional y altruista? En otras palabras, ¿puede un ser racional y autointeresado convertirse en altruista por conveniencia personal? Algunos han argumentado que sí puede. Como el altruismo es lo mejor en términos de autointerés, en casos como el DP se ha dicho que el altruismo puede emerger entre gente autointeresada. Así, personas racionales y autointeresadas podrían obtener los mejores resultados del dilema a través de su conversión a la moral. Mi posición, por el contrario, es que del egoísmo nunca se derivarán el altruismo y la cooperación. El argumento que basa esta afirmación es este: si el egoísta de mi compañero prisionero ha decidido ser altruista basado sólo en razones de autointerés, yo nunca cooperaría con él porque sé que tan pronto como lo hiciera, él se abstendría de cooperar conmigo. Autointeresadamente, lo que él debe querer es que yo coopere para maximizar su utilidad, pues puede lograr las mejores consecuencias cuando yo coopero y

él no. Yo sé todo esto, pues sé que él es autointeresado; por lo tanto, yo nunca cooperaría no importa cuánto trate mi compañero de convencerme de que cooperará conmigo.

Los prisioneros autointeresados que quieren convertirse en altruistas por conveniencia no escapan a la paradoja del DP, pues nunca pueden generar las expectativas necesarias para que la cooperación aparezca. La cooperación necesita que ambos prisioneros tengan expectativas fundadas de que el otro prisionero cooperará, pero ellos no tienen ninguna razón para tener esas expectativas, pues ambos saben que su compañero es un egoísta autointeresado que sólo trata de maximizar su utilidad.⁴ Si la cooperación no puede emerger, los egoístas nunca obtendrán los mejores resultados en interacciones del tipo del DP.

El escape moral

Se ha dicho que un individuo “desde premisas no-morales aceptará los límites de la moral”. Es interesante señalar que una de las bases de la confusión en los autores que sostienen esto, es que se concibe como autointeresado a aquel prisionero que decide cooperar sin saber qué posición ocupará en el dilema. En este caso, los atributos de este prisionero (autointerés e ignorancia de la posición que ocupará) lo llevan a decidir con imparcialidad, es decir, a decidir moralmente. Esta clase de prisioneros pueden escapar al DP porque no están interesados sólo en ellos sino en todos; adoptan una perspectiva general que no es adoptada por personas realmente autointeresadas sino sólo por gente altruista. Como estos prisioneros son en realidad individuos al-

⁴ D. M. Hodgson, *Consequences of Utilitarianism; A Study in Normative Ethics and Legal Theory*, Clarendon Press, Oxford, 1967.

truistas, el hecho de que ellos aceptarían las restricciones que la moral impone no ayuda a la tesis que estoy tratando de impugnar.

La razón que estos “especiales” autointeresados tienen para aceptar las limitaciones de la moral es que todos estarán mejor en un mundo moral. Esto explica por qué la moral es una institución social recurrente pero no implica que ser moral sea la mejor forma de actuar desde un punto de vista autointeresado. No lo implica porque de la proposición “todos estamos mejor si todos somos morales” no se sigue la proposición “cada uno está mejor si es moral”. Por el contrario, cada uno está mejor si es inmoral y todos los demás son morales. La confusión radica en la ambigüedad de “todos”. “Todos” en “todos estamos mejor si todos somos morales” no significa la suma de cada uno; “todos” se relaciona con la perspectiva de todos, la cual no es la perspectiva de nadie en particular ni la suma de las perspectivas de cada uno. Por ello, de que desde el punto de vista autointeresado adoptado desde la perspectiva de todos sea conveniente cooperar no puede derivarse que es también conveniente cooperar desde el punto de vista de cada uno cuando ellos son autointeresados.

He sostenido que en interacciones del tipo del DP individuos autointeresados y racionales no pueden convencerse de que la mejor solución es cooperar. Si son autointeresados, no serán capaces de escapar del argumento dominante. No podrán decidir autointeresadamente ser altruistas, pues no pueden generar las expectativas que individuos altruistas generan y que hacen posible la cooperación.

Las alternativas

El argumento básico para la cooperación no ha sido el

único dado en apoyo de la emergencia de la cooperación en un mundo no altruista. Últimamente se han hecho dos muy sofisticados intentos, recurriendo a distintas clases de argumentos.

En *Morals by Agreement*, Gauthier reconoce que el argumento básico para la cooperación es insuficiente para basar la cooperación entre individuos racionales y egoístas en situaciones estratégicas (situación de comportamientos mutuamente dependientes).⁵ Gauthier desarrolla una rica argumentación para mostrar que “lo racional y, más estrictamente, lo racional en términos de maximización de utilidad es remplazar la respuesta maximizadora de utilidad (estrategia egoísta), por la respuesta optimizadora de utilidad (estrategia cooperativa)”. Para mostrar esto, comienza con ciertos comentarios sobre la racionalidad práctica (racionalidad que lleva a la acción). Dice que un actor racional debe determinar sus elecciones en términos de expectativas recíprocas de una manera que satisfaga las siguientes condiciones de elección racional estratégica:

- A La elección de cada persona debe ser una respuesta racional a la elección que ella espera que otras personas hagan.
- B Cada persona debe esperar que toda decisión de otra persona satisfaga la condición A.
- C Cada persona debe esperar que sus decisiones y expectativas sean reflejadas en las expectativas de toda otra persona.

Contrariamente a lo que sucede en situaciones de certeza, donde se decide entre acciones que están relacionadas con determinadas consecuencias, en situaciones de riesgo o incertidumbre en que cada acción está correla-

⁵ David Gauthier, *Moral by Agreement*, Clarendon Press, Oxford, 1986.

cionada con un grupo de posibles consecuencias, cada una con una probabilidad, no puede elegirse directamente y sin más la consecuencia o el resultado preferido. De este modo, cada acción debe considerarse como una lotería en la que sus probables consecuencias cumplen la función de premios. Las preferencias de cada actor por cada acción (o lotería) deben depender exclusivamente de la utilidad de las consecuencias de cada acción por su probabilidad (utilidad esperada). La situación es más compleja aún en situaciones estratégicas (decisiones mutuamente dependientes), porque en este contexto la utilidad esperada de una acción depende de las acciones de otros y las acciones de otros dependen a su vez de la utilidad esperada de esas acciones (condiciones A, B, C).

Para superar el problema que presenta la mutua dependencia, Gauthier desarrolla un mecanismo de toma de decisiones caracterizado por ofrecer la posibilidad de maximizar el autointerés. Sostiene que la elección que maximiza el autointerés no debe ser una elección entre distintas acciones, es decir, no debe elegirse la realización de la acción A o de la acción B, sino una lotería que asigne a A y B una determinada probabilidad de ser realizada. Gauthier sostiene que la asignación de probabilidades a la acción A o B puede ser una asignación racional (satisface A, B y C), y por lo tanto puede ser la manera de maximizar racionalmente el autointerés. Dada la elección de una lotería y su distribución de probabilidades entre las distintas acciones, la acción que en definitiva se realice se determina por el uso de esa lotería. Una lotería sobre posibles acciones es denominada por Gauthier "una estrategia". Cada grupo de estrategias, una por cada persona interactuante, determina una lotería sobre las consecuencias de las acciones que en definitiva se realizarán. Esta lotería sobre conse-

cuencias está en equilibrio si y sólo si es el producto de estrategias, cada una de las cuales maximiza la utilidad esperada (utilidad por probabilidad) de cada una de las personas interactuantes dadas las estrategias elegidas por otras personas. Una lotería sobre consecuencias en equilibrio es a lo que los seres autointeresados deben tender, pues este equilibrio es lo único que garantiza que todos maximizan la utilidad esperada. La estrategia que maximiza mi utilidad esperada depende de las estrategias de los demás; de este modo y para evitar el círculo de estrategias mutuamente dependientes, es necesario un proceso de ajustes mutuos de utilidades esperadas.

Para usar el ejemplo de Gauthier: supongamos que María tiene dos posibilidades, ir a la fiesta de Ana o quedarse en casa. María tiene muchas ganas de ir a la fiesta, pero tiene aún más ganas de evitar encontrarse con Juan, quien puede estar allí. Si María supiese que Juan va a la fiesta ella no iría. Juan, por su parte, quiere evitar ir a la fiesta de Ana pero quiere aún más encontrarse con María, pero si Juan supiese que María va a la fiesta él iría a esa fiesta.

¿Cómo hace cada uno para maximizar su utilidad? Ésta es una situación estratégica, por lo cual, según Gauthier, la única forma de maximizar la utilidad es decidirse por una lotería que adjudique alguna probabilidad a las acciones de ir a la fiesta y de no ir. Dada la utilidad (supongamos que para María es cero si ambos van a la fiesta, un tercio si ninguno va, dos tercios si María no va y Juan va, y uno si sólo María va, y para Juan es cero si ninguno va a la fiesta, un tercio si Juan va y María no, dos tercios si ninguno va, y uno si ambos van), María debe elegir una lotería x con probabilidades (r) de ir y $(1-r)$ de no ir, tal que Juan tenga igual utilidad esperada vaya o no a la fiesta. Juan, por su lado,

debe elegir una lotería y que deje a María con igual utilidad esperada entre ir o no ir si ella adopta la lotería x . x debe ser tal que si adjudicase más probabilidades a la acción de ir, Juan no elegiría y sino alguna otra lotería que otorgue más probabilidades a su acción de ir, con lo cual María no quedaría, si adopta x , en situación de indiferencia respecto de su utilidad esperada entre ir y no ir.

No obstante la sofisticación de la presentación de Gauthier, su análisis, como él mismo lo reconoce, no puede resolver situaciones como el DP, porque el caso de Juan y María es un caso que presenta problemas de coordinación pero no de cooperación entre seres egoístas. Todo el problema radica en descubrir cuál es la estrategia maximizadora de utilidad para cada participante.

En esta situación uno puede diseñar una estrategia maximizadora de su propio bienestar dadas iguales estrategias por parte de otros, pues uno sabe que todos tienden a maximizar sus intereses. Las expectativas necesarias para adoptar una estrategia maximizadora siempre emergerán, pues lo único que todo jugador espera es que todo otro jugador actúe tratando de maximizar su interés y sabe que todos los demás jugadores actuarán en igual sentido. Obviamente este no es el caso de DP, pues allí la cuestión radica en la posibilidad de cooperación entre seres autointeresados que para cooperar deben esperar que los demás no estén dispuestos a maximizar su autointerés pues en este caso adoptarían la "alternativa dominante" (confesar) explotando toda disposición a cooperar.

La metasolución

Advirtiendo que su solución no es suficiente para resolver problemas cooperativos, pero sí importante pues nos

motiva a decidir basados en cálculos de probabilidades, Gauthier sostiene que es posible obtener estrategias cooperativas ya que "personas racionales y autointeresadas elegirán, para maximizar la utilidad, no hacer elecciones basadas en este criterio". Esta decisión será producto de la necesidad de cooperar con otros para obtener resultados mutuamente beneficiosos. Agentes racionales y autointeresados se dispondrán así a no ser maximizadores de utilidad porque ello los hará "elegibles como partes en cooperaciones mutuamente beneficiosas". Gauthier sostiene que los agentes se convertirán en "maximizadores-restringidos", lo que implica que ellos "tendrán una disposición a basar sus acciones en una estrategia conjunta sin considerar si alguna otra estrategia individual les arrojará una utilidad mayor". Seres racionales y autointeresados elegirán (metadecisión) estar dispuestos a ser "maximizadores-restringidos" considerando no la utilidad que se obtendría si todos cooperasen, sino la utilidad que obtendrían si cooperasen dada la estimación del grado en que los demás cooperan. Si esta estimación excede lo que obtendrían con una no-cooperación universal, la disposición a cooperar es una decisión racional. Gauthier no se equivoca con lo que afirma, aunque presupone la cuestión.

Si tenemos en cuenta que las expectativas acerca de la conducta de otros dependen estrictamente de nuestras decisiones acerca de cómo actuar y otros saben que lo que nos motiva es la maximización del interés, es difícil explicar por qué en una sociedad autointeresada habría individuos con disposición a cooperar (maximizadores-restringidos) en lugar de individuos que simulan esa disposición para motivar a otros a convertirse en maximizadores-restringidos, y así explotar su cooperación.

Ante la imposibilidad de determinar con certeza quiénes son maximizadores-restringidos y quiénes directos,

sólo sería racional disponerse como un maximizador-restringido si la probabilidad de que una interacción termine en cooperación fuese mayor que la probabilidad de que una interacción termine en explotación. Esta probabilidad está en relación con el número de maximizadores-restringidos, de tal modo que a medida que éstos aumentan disminuye la probabilidad de que una interacción termine en explotación. Podría ser racional disponerse a cooperar cuando en una sociedad hubiese suficientes maximizadores-restringidos.

Si bien esto parece ser el caso, no soluciona situaciones del tipo de DP, porque sigue sin tener explicación por qué personas autointeresadas se convertirán en maximizadores-restringidos cuando las expectativas de que otros se limiten a restringir su utilidad sean irracionales. La presuposición de la existencia de maximizadores-restringidos en una sociedad simplifica el trabajo de Gauthier, pero lo que está justamente en duda es la posibilidad de esta existencia en sociedades de individuos autointeresados. Por lo tanto toda expectativa que implique que otros tendrán la disposición de restringir la maximización de su utilidad basados en razones de autointerés parece ser injustificada. Consecuentemente la fundación egoísta de la moral requerirá otra clase de argumento.

La solución continua

Robert Axelrod ha desarrollado en *The Evolution of Cooperation*⁶ otra línea de argumentación, que asume que dos egoístas jugando una sola vez el juego (DP) elegirían la "alternativa dominante".

⁶ Robert Axelrod, *The Evolution of Cooperation*, Basic Books, Nueva York, 1984.

El análisis de Axelrod reconoce la imposibilidad de una metasolución, pues reconoce la imposibilidad de formular racionalmente expectativas acerca de cómo actuarán los otros jugadores. De este modo, las expectativas no ocupan un lugar central en Axelrod.

Axelrod dice que lo que hace emerger la cooperación es que los jugadores saben que se encontrarán nuevamente en situaciones análogas en el futuro y que la decisión tomada hoy influirá la decisión que se tomará mañana. Axelrod sostiene que la cooperación emergerá si los individuos tienen suficientes oportunidades de volver a encontrarse, de tal modo que tengan un interés en el futuro. Para demostrar esto, recurre a un DP continuo en lugar de un único DP como el que hemos estado tratando. Para fundar su argumento recurre a una experiencia real.

En un torneo de DP continuo en el cual tomaron parte distinguidos profesores de varios países del mundo, Axelrod encontró que los que adoptaron una táctica consistente en cooperar en la primera interacción y luego decidir tal como decidió el otro jugador en la interacción previa, obtuvieron mejores resultados que cualquier otro. Axelrod sostiene que todos los jugadores racionales autointeresados deberían aceptar esta estrategia, que causaría la emergencia de la cooperación aun en un mundo de egoístas. La táctica fue llamada TIT FOR TAT y las propiedades que hicieron a TIT FOR TAT obtener un mejor resultado fueron:

- 1) La propiedad de ser buena: nunca es la primera en defezionar confesando.
- 2) La propiedad de ser respondiente: responder actuando de la misma manera en que el otro ha actuado.

- 3) La propiedad de poder perdonar: propensión a cooperar después de que el otro ha defecionado.
- 4) La propiedad de ser vengativa: defecionar inmediatamente después de una defeción de la otra persona.

Sin embargo, si nos acercamos al argumento de Axelrod, difícilmente encontraremos bases para la fundación autointeresada de la moral. Es verdad que en el torneo, TIT FOR TAT fue la táctica más exitosa, pero lo fue porque no todo participante fue racional en la búsqueda de su interés. En el torneo no todos eligieron la misma táctica, lo que implica que no todos fueron igualmente racionales. Por el contrario, en situaciones de DP todos los participantes deben decidir en forma idéntica por el hecho que de ellos se predicen propiedades idénticas. Con esta sola afirmación no se puede impugnar la sustancia argumentativa que se derivaría de los resultados del torneo, pues lo único que la necesidad de idénticas respuestas puede probar es que todos jugarían la táctica TIT FOR TAT. Pero tal como Axelrod reconoce, existen otras tácticas que, de haberse jugado, habrían arrojado mejores resultados que TIT FOR TAT. TIT FOR TAT triunfó en el torneo, pero de haberse jugado un torneo más amplio, se habrían jugado tácticas que vencen a TIT FOR TAT y los resultados habrían sido distintos.

Si los participantes fuesen plenamente racionales, TIT FOR TAT nunca habría sido jugada, pues no obstante vencer muchas tácticas, es también vencida por muchas otras. Una persona racional y autointeresada habría intentado otra táctica; una capaz de vencer a TIT FOR TAT. De este modo, la cooperación no habría emergido. Sosteniendo que en el torneo se habría jugado TIT FOR TAT, Axelrod presupone que quien lo juega no es autointeresado y racional. Quien está dispuesto a

tomar un riesgo acerca del carácter y de las disposiciones de otros jugadores, sin medir las consecuencias para su autointerés, difícilmente puede ser definido como un ser egoísta y menos aún como un racional egoísta. Tomar este riesgo implica aumentar las probabilidades de ser explotado y nadie absolutamente autointeresado lo haría a menos que este riesgo estuviese justificado por ganancias futuras. Estas ganancias pueden existir cuando ya hay una cantidad de individuos dispuestos a cooperar por otras razones que razones de autointerés. En ese ambiente es probable que individuos autointeresados que participen en varias interacciones para maximizar su bienestar deban cooperar, pero el problema consiste en ver si en una sociedad autointeresada la cooperación emergerá en medida alguna.

Hay también otro argumento en contra de la elección de una táctica cooperativa como TIT FOR TAT. En un número finito y determinado de interacciones existe la certeza acerca del número de jugadas. Esa certidumbre nos hará optar por una estrategia no cooperativa porque sabemos, como la contraparte sabe, que explotar a otros será la mejor decisión en la última jugada. Como sabemos que ambos debemos defeccionar en la última jugada para maximizar nuestro interés, debemos defeccionar también en la penúltima; de otro modo seremos explotados por la defección de los demás, que saben que en la próxima jugada no se cooperará, por lo que no es redituable cooperar en ésta. Así hasta llegar a la primera jugada, en la cual lo mejor será también defeccionar. Ésta es una razón, reconocida por Axelrod, que hace que la cooperación no pueda emerger si el número de situaciones de DP que enfrentamos está previamente determinado. Por ello Axelrod sostiene que el juego debe tener indeterminado número de jugadas. No obstante, la probabilidad de que la próxima jugada

sea la última aumenta a medida que el juego avanza, con lo que aumenta la probabilidad de ser explotado. Esto es así incluso en aquellos casos en que el número de jugadas se extienda considerablemente después de haber pasado cierto umbral. Es decir, el caso en el que después de haber atravesado cierto umbral las probabilidades de que la próxima sea la última jugada se reducen considerablemente. Aun en este caso la situación impedirá que la cooperación emerja, pues la expansión del número de jugadas no es conocida por los jugadores. Con base en el único dato que ellos poseen, la finitud del juego, lo racional será computar únicamente el aumento de la probabilidad de que la próxima jugada sea la última. Los jugadores estarán dispuestos a interactuar un número de veces determinado pasado el cual la probabilidad de que la próxima sea la última jugada será demasiado grande como para justificar el riesgo que implica cooperar. Pero fijado ese número de veces (x), en ese número menos uno ($x-1$) tampoco será racional cooperar, pues ambos saben que dado que la próxima vez (x) no cooperarán, cooperar esta vez ($x-1$) no es racional por la razón arriba citada. Ese número menos dos presentará idéntico problema y así hasta la primera jugada, en la cual lo más redituable, razonando sobre bases autointeresadas, será defeccionar no cooperando. De este modo, incluso en la primera jugada los jugadores intentarán explotar a los demás, con lo que la cooperación se convertirá en imposible.

De alguna manera estos problemas son sugeridos por Axelrod cuando propone cuatro condiciones para ser exitoso en un DP continuo. Ellas son: 1) no ser envidioso; 2) no ser el primero en defeccionar; 3) devolver la cooperación y la defección, y 4) no ser demasiado inteligente. Sin embargo, estas condiciones frustran la empresa de mostrar cómo la cooperación emergerá en-

tre personas racionales y autointeresadas, pues implican claramente que los individuos se desenvuelven mejor en el DP cuando no son autointeresados.

Intenté mostrar en este trabajo que ni el argumento basado en expectativas, ni el basado en la metasolución, ni el basado en el dilema continuo, son argumentos plausibles para fundar las bases egoístas de la moral. Ni aun en los casos más formalizados los filósofos han encontrado argumentos que fundamenten esta posición. La moral es una institución que requiere pensar en términos de todos, tomando la utilidad o el bienestar de otros tan seriamente como si fuesen los nuestros. Sería realmente sorprendente si esta institución permitiese la adhesión basándose en razones que por definición son las más opuestas a ella. Por esto, fundar la moral en el autointerés parece ser una empresa fútil.

El argumento racional para la cooperación

Sin embargo, la falta de justificación de la moral por razones de autointerés no significa que la decisión de ser moral sea una decisión irracional. Como lo mencioné al comienzo del trabajo, existen fuertes razones que fundamentan esta decisión. Ciertamente es que puede decidirse no ser moral, pero esta decisión difícilmente será calificada como racional si no se ofrecen argumentos que la justifiquen. Si lo hago, es decir, si ofrezco razones, es porque valoro la adhesión de otros a mis razones. Valoro su libre adhesión a los principios que explican y justifican mi acción. No siendo este el caso, nunca daría razones de mis acciones sino que forzaría a mis semejantes a aceptar mis principios y las consecuencias de mis actos. El ofrecer razones se inserta en una práctica que aspira a lograr la convergencia con las razones ofrecidas. Esto es innegable diga lo que diga, pues si ofrezco ra-

zones es porque me encuentro inserto en la práctica de ofrecer razones, aceptando y valorando implícitamente la adhesión de otros a las razones ofrecidas. Si intento negarlo caigo en lo que Carlos Nino llamó la "inconsistencia pragmática". Esto describe el intento de negar, mediante razones, lo que es aceptado implícitamente por las acciones que uno realiza. En otras palabras, lo que está presupuesto en la actividad de dar razones es el valor de la decisión autónoma de otros de converger con nuestras razones y justificaciones. De este modo, en toda discusión, más allá de afirmaciones en contrario, se valúa la autonomía de otros. La fundamentación autointeresada de la moral pretende basarse en argumentos que desconocen el valor de esa autonomía, pues considera como razones para ser moral únicamente aquellas que maximizan nuestra autonomía independientemente de la autonomía de los demás. Quien decida no ser moral intentando justificar esta decisión mediante razones autointeresadas incurre indefectiblemente en la inconsistencia pragmática, pues intenta negar aquello que con sus argumentos presupone. Sus argumentos, por el hecho de estar insertos en la práctica de ofrecer razones, presuponen el valor de la autonomía, aunque por ser autointeresados intentan negar esa presuposición. De este modo, la única forma de ser consistente (lo que es un requisito de la racionalidad) con lo que uno implica cuando argumenta es adoptar las restricciones impuestas por la necesidad de salvaguardar la autonomía de otros. En otras palabras, ser moral.

Por supuesto puedo decidir no ser racional y no hay razón alguna que pueda convencerme de serlo, ya que la decisión de no ser racional implica no tomar en consideración ninguna razón. Pero si para negar la moral es necesario negar la racionalidad, también el intento

de fundar racionalmente esta afirmación será autofrustrante.

SUMMARY

In this paper I try to highlight the relationship between self-centered reasons and morality. My basic aim is to show that not even the most sophisticated attempts of current philosophical literature can bridge the wide gap that detaches from one another. I focus the 'prisoner's dilemma' showing that the best pay-off for everybody in that situation can only be obtained if everybody refuses to think in terms of the maximization of his or her utility and decides to cooperate keeping silent. I argue that this decision cannot be based in self-interested reasons because self-interested people can never generate the expectations needed for cooperation to arise. This impossibility stands as a wall against which Gauthier's and Axelrod's attempts bounce. Which I call in the text the metha-solution (Gauthier) and the iterated solution (Axelrod) cannot overcome the fact that knowing what the rest of as self-interested I would never cooperate because they will try to exploit my cooperation and were they altruistic I would never cooperate either because I will try to exploit their cooperation. Nevertheless there is a limited plausibility in the approaches I challenge which derives from the assumption that some people have already decide to cooperate not on bases of self-interested reasons but on reasons of different sort. Obviously this cannot serve as a proof of the existence of the bridge I referred above. Finally I sketch why there are reasons to be moral which make the decision of restraining my actions to preserve the interests of others a rational one.