

## LAS REGLAS DE LA COOPERACIÓN

### 1. Introducción

**H**ay algunas palabras que parecen encerrar un sentido mágico. Es como si ellas mismas concretaran un valor fundamental indiscutible, y en consecuencia su negociación es, también por sí misma, la negación de ese valor. Podrá objetarse que no se trata estrictamente de palabras sino de ciertos usos de las mismas; pero en rigor no se puede negar que determinadas palabras se prestan más que otras a tales usos que las colocan en posición magnificente o desdichada, según el caso. Algo de esto sucede con la palabra cooperación. En la actualidad, en el terreno de la Filosofía del Derecho, de la Filosofía Política y de la Filosofía Moral, se destacan marcadamente las bondades de la cooperación, y se asocia el comportamiento cooperativo de los actores sociales, con las llamadas conductas socialmente deseables, como contrarias a las indeseables conductas egoístas. Se ha llegado al extremo de afirmar que el comportamiento cooperativo es portador de un vigor tal, que le permite emerger, desarrollarse y hasta obtener el triunfo final sin auxilio alguno, aun en un mundo extremadamente egoísta. Con esto se llega a admitir (como sucede siempre que se apela a lo mágico) que el valor, la fuerza y el poder que inicialmente se radica en la palabra, es portado de una manera directa a las acciones que la misma nombra. En este sentido, algunas tesis, como la sostenida por R. Axelrod-84, hacen pensar en una ley universal de evolución que conduce de manera inexorable al surgimiento y triunfo de la cooperación, aun en un mundo de extremo egoísmo. El peligro de teorías de este tipo es que, utilizando simulaciones que parten de un cuerpo de restricciones formales no siempre plausibles o diseñables en la realidad, tienden a elaborar un argumento que engañosamente posee base empírica, y luego es convertido de un modo insensible, y generalmente acrítico, en normativo.

Estas ideas han llegado a impregnar no sólo el panorama intelectual, sino también el discurso político de los últimos años, y obtuvieron un importante refuerzo logístico al ser vinculadas a los conceptos y metodología propios de la teoría de los juegos. Sin entrar por el momento a discutir el significado que tales ideas puedan haber alcanzado, tanto en el plano de la teoría como de la práctica social, lo cierto es que hay muchos de sus rasgos que invitan a una consideración crítica, ya que producen la desagradable impresión de ofrecer soluciones de dudosa justificación, a problemas ambiguamente caracterizados. Bueno es señalar que no sólo la palabra cooperación exhibe este sentido mágico. En el escenario de la teoría política, la palabra estabilidad suele ser tratada del mismo modo, es decir, como si encarnara un valor indiscutible. Sin embargo, sólo cuando el concepto se somete a un análisis metódico en su contenido y en sus connotaciones, se descubren rasgos diferenciadores que permiten caracterizar de manera adecuada el núcleo del problema implicado, y al mismo tiempo evaluar las inevitables dificultades teóricas y metodológicas que deben sortearse para dejarlo libre de ambigüedades. (E. Garzón Valdés-87.)

Tal como lo hacen quienes creen en esa fuerza mágica de la palabra cooperación y de las acciones cooperativas, vamos también nosotros a hacer uso del llamado Dilema del Prisionero<sup>1</sup>, ya que el mismo ofrece con una bella frugalidad, un rico panorama para analizar cuánto de explicativo y cuánto de normativo se encierra en ese confuso asunto de la cooperación tan mencionado últimamente.

## 2. Algo acerca de los juegos

Los planteamientos éticos que consideran que las conductas cooperativas pueden exhibir una cierta superioridad respecto de las no cooperativas, han buscado desde hace algún tiempo un núcleo de fundamentación en la llamada teoría de los juegos. Se trata de una teoría formalizada de exquisita elegancia, e indiscutible

---

<sup>1</sup> Originalmente formulado por Merrill Flood y Melvin Dresher y formalizado por A. W. Tucker, el Dilema del Prisionero ha merecido una gran atención por parte de los estudiosos de las Ciencias Políticas, la Teoría de los Juegos, la Psicología Social, y la Filosofía Moral, quienes han ido descubriendo vetas de interés en el planteo general. Desdichadamente, en algunos casos se ha tergiversado la naturaleza del modelo, considerándolo como un reflejo de situaciones reales; o lo que es peor aún, como un paradigma moral que destaca la superioridad absoluta del comportamiento cooperativo.

poder como herramienta analítica<sup>2</sup>. Desde este punto de vista, las aportaciones que la teoría de los juegos puede ofrecer no son pocas; y en efecto, la misma ha contribuido a clarificar muchos aspectos que suelen permanecer oscuros cuando los planteamientos se efectúan desde un lenguaje de mayor vaguedad. Asimismo, la apelación a conceptos y estructuras de la teoría de los juegos ha permitido una fecunda relación entre la Filosofía Moral y la Teoría Matemática, que ha contribuido a enriquecer el panorama de ambas esferas<sup>3</sup>. Sin embargo, para que un modelo formal como el de la teoría de los juegos pueda ofrecer frutos verdaderamente satisfactorios, y no se constituya en un mero elemento justificatorio de conclusiones, o en modo elegante de presentación de sinsentidos, es necesario reconocer los principios básicos del mismo, que siempre constituyen las primeras restricciones del modelo, y cuya consideración es fundamental en el análisis del isomorfismo (J. Barragán-83).

Es razonable la perspectiva que considera que el concepto primitivo en la teoría de los juegos, es el de **interés**. De una manera general, el contenido de este concepto no difiere fundamentalmente de aquello que en el plano intuitivo entendemos por interés. Cuando un sujeto tiene preferencias personales respecto de los varios posibles resultados en una determinada circunstancia, se dice que posee interés en ese resultado. Esto nos coloca claramente en un terreno en el que queda excluida la llamada generosidad altruista como acto de desprendimiento de las propias preferencias: nuestros actores en la teoría de los juegos tienen preferencias, no son desinteresados y no pueden serlo si queremos aceptar el concepto primitivo de la teoría.

Del concepto de **interés**, se deriva el de **conflicto de intereses**. En el marco de la teoría de los juegos cuando dos o más individuos tienen preferencias acerca de posibles resultados que les conciernen recíprocamente, y no logran ponerse de acuerdo acerca de esas preferencias, nos encontramos frente a un conflicto entre los intereses. Tampoco este concepto está demasiado lejos de la intuición, y nos refuerza el anterior concepto de interés,

---

<sup>2</sup> Una referencia sistemática que ilustra de modo completo acerca de los antecedentes de la Teoría de los Juegos se encuentra en R. D. Luce y H. Raiffa-57, sección 1.2.

<sup>3</sup> Entre los autores que han recogido excelentes frutos en esta relación entre Filosofía Moral y Teoría de los Juegos, se encuentran J. Harsanyi y D. Gauthier, quienes, manteniendo una gran coherencia con los conceptos de ambas disciplinas, han logrado construir a lo largo de los años importantes teorías éticas, expresadas en términos formalmente elegantes y de gran plenitud semántica.

ya que pone en evidencia que nuestros actores no sólo tienen preferencias, sino que se mueven en un ámbito en el que la existencia de las preferencias de otros actores (las cuales no siempre pueden armonizarse con las propias) genera conflictos entre esas preferencias.

También del concepto de **interés** es posible derivar el de **utilidad**. Esta palabra arrastra un destino desdichado. Por lo general, cuando se la emplea origina una gran resistencia y muchos malentendidos. Por eso es particularmente importante establecer con claridad cuál es su uso en la teoría de los juegos, ya que si se apela a este aparato analítico es preciso aceptar tal uso, aunque contradiga algunas ideas previas que se puedan tener acerca del término.

Cuando un individuo está en condiciones de ordenar coherentemente sus preferencias acerca de los posibles resultados, podemos hablar de **utilidad** en el sentido de la teoría de los juegos (J. von Neumann and O. Morgenstern-47). Esta consistencia en el ordenamiento de las preferencias permite representarlas numéricamente mediante las así llamadas funciones de utilidad, y también hace que la utilidad exhiba la muy importante propiedad que puede expresarse del siguiente modo: un individuo determinado prefiere un resultado a otro, si y sólo si la utilidad esperada del primero supera a la utilidad esperada del segundo. Dicho con mayor elegancia, cada individuo procura maximizar su utilidad esperada. El empleo de la utilidad como medida de la preferencia, intenta resolver el problema de las comparaciones interpersonales. En efecto, cuando se emplea una unidad monetaria en el análisis de un conflicto de intereses, es posible caer en equívocos, ya que la misma cifra no tiene el mismo significado para un sujeto que nada posee que para alguien que lo tiene todo. Al expresar la preferencia en unidades de utilidad, y bajo el supuesto de que los patrones de preferencia son similares en cuanto a coherencia entre los actores, el marco de referencia naturalmente crece en previsibilidad.

En este concepto von neumanniano de utilidad, ya no podemos apelar a la intuición de una manera tan natural como lo hemos hecho con los anteriores. Sin embargo, para el caso de recurrir al auxilio de la teoría de los juegos como herramienta analítica, parece imprescindible usar la utilidad como medida de preferencia, ya que aunque la teoría de la utilidad tiene diferentes contextos de aplicación, no es posible ignorar que fue construida (y sigue siéndolo) como el pilar fundamental de la teoría de los juegos.

Ante un conflicto de intereses producido como consecuencia

de la no armonización de las preferencias, cada individuo tiene que elegir un determinado curso de acción con el fin de maximizar su utilidad esperada. Esta elección efectuada con tal finalidad define el concepto de **toma de decisión** que contribuye, junto a los anteriormente desarrollados a clarificar el panorama de la teoría de los juegos. La toma de decisión individual<sup>4</sup>, que es la que nos interesa, puede ser efectuada bajo una situación de total certidumbre, bajo una condición de riesgo o una de incertidumbre. La toma de decisión bajo certidumbre se da cuando se conoce perfectamente el resultado específico al que nos conduce una determinada acción; la toma de decisión bajo riesgo, se da cuando la acción determinada se asocia no a un resultado, sino a un conjunto de resultados posibles, cuya probabilidad (objetivamente basada) es conocida por quien toma la decisión; finalmente, en la toma de decisión bajo incertidumbre, la acción también conduce a un conjunto de resultados posibles cuya probabilidad (subjetiva) no es completamente conocida.

Tratando de mostrar como se insertan los conceptos descriptos dentro del escenario de la teoría de los juegos, puede decirse que para cada individuo implicado, el problema del conflicto de intereses se convierte en un problema de toma de decisión bajo condiciones de riesgo e incertidumbre; derivada esta última del hecho de que cada jugador ignora qué curso de acción elegirá su contrincante. Así puestas las cosas, puede afirmarse que la teoría de los juegos procura dotar de un marco general teórico a los problemas implicados en la toma de decisión interactiva bajo condiciones de riesgo; en razón del carácter interactivo, se agrega de modo natural la condición de incertidumbre.

En este contexto, si una persona puede expresar sus preferencias de un modo coherente, esto es si tal sujeto exhibe consistencia en sus elecciones<sup>5</sup>, es posible elaborar una función de utilidad.

---

<sup>4</sup> Cuando en la Teoría de los Juegos se habla de **decisión individual** no se alude necesariamente a la decisión tomada por un solo individuo. Una decisión se considera individual cuando es una voluntad unitaria quien la toma. Tal voluntad puede coincidir con la de una persona, o también ser la expresión de un colectivo que ha logrado establecer por algún medio, tal unidad volitiva.

<sup>5</sup> Las preferencias deben satisfacer los llamados axiomas de consistencia a fin de posibilitar resultados inequívocos que permitan comparaciones interpersonales de utilidad. Tales axiomas son: a) Las alternativas deben ser comparables. Dados A y B, el sujeto preferirá A o B, o será indiferente entre A y B. b) Tanto la relación de preferencia como la indiferencia son transitivas. Dados A, B y C, si el sujeto prefiere A a B y B a C, entonces prefiere A a C; y si el sujeto es indiferente entre A y B y entre B y C, entonces es indiferente entre A y C. c) Si una lotería tiene como una de sus alternativas otra lotería, la primera

Esta función de utilidad no refleja las preferencias del actor en un sentido abstracto, sino que se refiere a las alternativas en una situación perfectamente definida: la del juego, y en ese marco específico procura maximizar la utilidad esperada por el jugador. La teoría de los juegos intenta elaborar patrones generales de toma de decisión interactiva, bajo las condiciones que ya hemos señalado.

Considerando los patrones de preferencia de los jugadores, es posible clasificar los juegos en estrictamente competitivos y no estrictamente competitivos. Como sus nombres lo indican, mientras los primeros exhiben patrones de preferencias estrictamente opuestos, en los segundos es posible encontrar por lo menos una combinación A y A' entre las soluciones del juego, tal que mientras un jugador prefiere A a A', el otro prefiere A' a A. Los primeros son juegos-suma-cero y los segundos son juegos-no-suma-cero.

A su vez, entre los juegos no-suma-cero, se encuentran los llamados juegos cooperativos, que permiten a los jugadores hacer todo tipo de acuerdo previo al juego, y reforzar tales acuerdos mediante sistemas de premios y sanciones; en el mismo grupo no-suma-cero también encontramos los llamados juegos-no-cooperativos, en los que tanto la comunicación previa al juego entre los jugadores, como los acuerdos de cualquier tipo, están expresamente excluidos.

### 3. El Dilema del Prisionero

Del Dilema del Prisionero se ubica entre los así llamados juegos-no-suma-cero-no-cooperativos. En consecuencia, es un juego no estrictamente competitivo, y que exhibe una restricción inicial relativa a la prohibición de que los jugadores se comuniquen entre sí o efectúen acuerdos previos al juego. Como en todo juego, cada uno de los prisioneros es capaz de ordenar coherentemente sus preferencias y siempre procurará maximizar su propia utilidad esperada.

---

puede descomponerse en las loterías básicas mediante el cálculo de probabilidades. *d)* Si dos loterías son indiferentes al sujeto, también son intercambiables como alternativas en cualquier lotería compuesta. *e)* Si dos loterías involucran las dos mismas alternativas, entonces aquella en la que la alternativa preferida tiene una probabilidad más alta de ocurrencia, es también la preferida. *f)* Si A es preferida a B y B a C, entonces existe una lotería que involucra a A y C (con probabilidades adecuadas) que es indiferente a B.

En su versión más simple, el Dilema del Prisionero nos presenta el panorama de dos prisioneros sometidos a juicio y privados de su libertad. Estos prisioneros son entrevistados por separado por el fiscal, quien les efectúa algunas propuestas con los respectivos ofrecimientos de utilidad. Si el fiscal logra que uno de los prisioneros delate al otro, tendrá en sus manos todos los elementos de juicio necesarios para aplicar a este último el máximo de la pena prevista para el delito; para «agradecer» este favor del delator, pedirá su libertad. Como la oferta del fiscal es simétrica, puede darse el caso de que ambos prisioneros la acepten, lo que equivaldría a una mutua delación, situación en la que ambos prisioneros recibirían una pena cercana a la máxima. Para el caso que ambos decidan no delatarse, el fiscal quedará sin elementos suficientes y sólo podrá aplicar a ambos un mínimo de pena. Estas ofertas producen a los jugadores determinadas utilidades según sea la estrategia que sigan (condicional a la seguida por el otro jugador). Estas utilidades esperadas se expresan en la llamada matriz de ganancias (utilidades), que para el caso en consideración es la siguiente:

		B	
		D	-D
A	D	2,2	10,0
	-D	0,10	8,8

Los números de la matriz expresan la utilidad (expresada en años) que cada prisionero obtendría en cada una de las cuatro posibles intersecciones de estrategias.

Como se trata de un juego-no-suma-cero-no-cooperativo, las restricciones establecidas en el planteamiento de este Dilema, restricciones que como es natural cumplen un importante papel en las soluciones posibles, serán: 1) no pueden existir acuerdos de cumplimiento forzoso entre los prisioneros; y a los efectos de garantizar esa primera restricción, 2) los prisioneros no pueden comunicarse entre sí, ni directamente, ni a través de terceros.

A partir del esquema planteado, y considerando que ambos prisioneros son actores racionales, resulta coherente que el prisionero A opte por la estrategia D, ya que la misma le garantiza una ganancia mínima de 2,2 y una máxima de 10; mientras que la estrategia —D le ofrece una posible ganancia que va entre 0,10 y 8,8.

También es coherente con el carácter racional del prisionero B que elija la estrategia D por idénticas razones que A. Esto hace que ambos prisioneros se encuentren en la casilla (2,2), que es la de la mínima suma de ganancias. La salida de esta casilla en la que las ganancias colectivas son las mínimas, se llama **no cooperativa**, mientras que la salida de la casilla (8,8) donde las ganancias colectivas son las máximas, se llama **cooperativa**.

¿Qué es lo que ha impedido que los prisioneros se encuentren en la salida cooperativa, y en su lugar converjan a la salida no-cooperativa? Desde luego que estando vedada por la restricción inicial la posibilidad de un acuerdo forzoso (y de cumplimiento forzoso), cada prisionero teme ser traicionado por el otro, y quedar en consecuencia con 0 de ganancia. Las casillas (0,10) y (10,0), representan para los prisioneros un terreno minado de riesgos que no se deciden a afrontar para llegar a la casilla (8,8). Este temor recíproco a ser traicionados es el factor que favorece la selección de la intersección (D,D) para la mayor satisfacción del fiscal.

Como puede verse, la combinación de estrategias (D,D) se autorrefuerza al interior del juego, ya que contiene la mejor respuesta de cada jugador frente a cualquiera de las estrategias posibles del otro jugador. Este acuerdo autorreforzado dentro de los límites de las restricciones del juego es el punto de equilibrio del mismo; y bajo el supuesto de racionalidad de los jugadores, es decir, utilizando éstos su mejor estrategia posible, tenderán hacia esa intersección, obteniendo sólo un pago de (2,2) en lugar de (8,8).

En consecuencia, tanto para el prisionero A como para el B las estrategias D son estrictamente dominantes, y el par (D,D) es el único par en equilibrio del juego. Esto significa que cualquier jugador racional decidirá utilizar la estrategia D, ya que la misma maximiza su utilidad esperada.

Como sucede siempre con los modelos conceptuales, el Dilema del Prisionero es susceptible de dos interpretaciones diferentes, las cuales definen también dos diferentes usos. Por una parte, un modelo conceptual puede ser primordialmente interpretado como descriptivo de la estructura de determinadas relaciones, cuyo análisis permitiría elaborar categorías explicativas, y predecir el futuro comportamiento de tales relaciones; pero por otra parte también el modelo puede ser interpretado como definidor de conductas socialmente deseables, lo que haría posible producir normas de comportamiento para regir las relaciones sociales. En el marco del uso descriptivo, el Dilema del Prisionero nos refiere cómo se comportarían dos jugadores racionales,



considerando las restricciones establecidas; mientras que dentro del uso normativo, el Dilema puede ser utilizado para poner en evidencia las ventajas de la salida cooperativa sobre la no cooperativa, lo que a su vez puede llevar a afirmar la superioridad ética de las conductas cooperativas en términos de utilidad social<sup>6</sup>.

Ambas interpretaciones son perfectamente legítimas, y cumplen una función no necesariamente antagónica; pero es importante puntualizar con cuál de ellas se trabaja en un determinado momento, ya que la mayor parte de las herramientas conceptuales y las estrategias metodológicas que pueden provechosamente emplearse en un caso, son perfectamente impropias para el otro. Así, en la interpretación descriptiva, los conceptos y métodos lógico-matemáticos son centrales para construir un aparato explicativo consistente, que permita predicciones acertadas de los comportamientos; mientras que en una interpretación normativa, los conceptos y métodos propios de la crítica filosófica desempeñan el papel protagónico para la elaboración de un marco coherente que haga posible la comparación de los sistemas de valores implícitos en todo enfoque normativo.

Considerado desde un punto de vista analítico, el hecho de que en el Dilema del Prisionero la solución en equilibrio sea la no-cooperativa, es un resultado necesario si se aceptan las restricciones que el modelo exhibe, y sólo puede ser objeto, como en efecto lo ha sido, de un tratamiento analítico formal (por ejemplo, Nash-50). Tal solución no puede de ninguna manera ser objeto de valoración moral, ni considerada particularmente perversa, ya que la perspectiva normativa rebasa la capacidad analítica de las herramientas lógico-matemáticas, y requeriría de conceptos y métodos en los que la crítica de los principios juega el rol fundamental. No obstante, la perspectiva analítica puede contribuir en buena medida con la normativa, ya que torna particularmente evidentes las consecuencias que conlleva la aceptación de determinados principios.

A este respecto, con el Dilema del Prisionero sucede algo interesante; muchos autores, en el deseo de encontrar una justificación racional para la superioridad ética de la cooperación, hacen aparecer como relaciones de hecho lo que en verdad son

---

<sup>6</sup> El uso de los modelos en las Ciencias Sociales conlleva siempre importantes riesgos. Por lo general, es difícil separar claramente los diferentes usos, y sus respectivos métodos. Esta circunstancia lleva insensiblemente por caminos confusos que no conducen a ninguna parte, disminuyendo el potencial del propio modelo.

expresiones de valor, confundiendo la interpretación descriptiva con la normativa. El resultado de esta confusión es que el análisis de ambos planos se ve afectado, desde el momento que no se emplean los elementos conceptuales y metodológicos específicos para cada uno de ellos. No es novedoso que los juicios de valor admiten una crítica y una justificación racional; pero nada se hace a favor de la misma si se encubre a los valores bajo la forma de hechos de la realidad, sustrayéndolos a los métodos propios de su esfera específica. Este encubrimiento representa una variante más, aunque esté revestida de elegante sutileza, de los enfoques irracionales acerca de los juicios de valor. Por otra parte, el hacer uso indebido de la fuerza y la plenitud semántica de los métodos lógico-matemáticos, sólo puede contribuir a un inmerecido desprestigio de los mismos en el terreno de las ciencias humanas<sup>7</sup>. La diferenciación apropiada entre ambas interpretaciones del Dilema del Prisionero no sólo ofrece un indiscutible interés analítico, sino que genera saludables repercusiones en la esfera de la política práctica.

#### 4. El Dilema del Prisionero como modelo analítico

Como modelo analítico, el Dilema del Prisionero puede considerarse, por una parte, en el plano de su estructura formal, y por la otra, en el de su plausibilidad como herramienta para visualizar las relaciones de hecho a las cuales se aplica.

Como estructura formal, el Dilema del Prisionero consiste en una red de relaciones que poseen determinadas propiedades, enmarcadas en un grupo de restricciones, las que definen las inferencias que legítimamente pueden efectuarse dentro del modelo. En virtud de estas restricciones es que el Dilema del Prisionero es un dilema; ya que la mejor estrategia conjunta no conduce al mejor pago posible. Esto luce como el resultado «irracional» de una selección de estrategias racionalmente efectuada.

Visto así, el Dilema posee la estructura de una toma de decisión individual interactiva, efectuada bajo condiciones de riesgo e incertidumbre, con un decisor que es capaz de ordenar coherentemente sus preferencias, y procura maximizar la propia utilidad esperada mediante su elección. Para analizar el Dilema en

---

<sup>7</sup> Para una discusión más amplia acerca de la posibilidad del análisis racional de los juicios valorativos, en el marco de la construcción de políticas públicas, puede verse J. Barragán-85.

este plano, son necesarias y suficientes las herramientas conceptuales lógico-matemáticas.

Pero he aquí que tales estructuras son utilizadas normalmente con el propósito de analizar relaciones que se desenvuelven con una dinámica propia en el terreno de los hechos; bajo estas condiciones en las que se plantea una aplicación del modelo analítico al plano fáctico, las reglas lógico-matemáticas resultan necesarias pero insuficientes para evaluar el modelo. En estos casos es necesario considerar cuidadosamente si las restricciones, los supuestos y las reglas de inferencia se vinculan amigablemente con las ocurrencias de la realidad, y resultan, en consecuencia, buenas herramientas de análisis. El plano lógico aquí debe ser necesariamente completado por el metodológico que toma en consideración no sólo las relaciones formales y sus propiedades, sino también la eficacia de las mismas en su función analítica de un determinado problema.

Ya sea que se siga el camino de construir primero el cuerpo de conceptos, para luego ajustar al mismo las restricciones de la estructura formal (R. D. Luce and H. Raiffa-57), o el procedimiento opuesto de ajustar los cuerpos conceptuales al modelo formal que se desea utilizar, siempre se hace necesario el análisis metodológico de la plausibilidad de las restricciones en el específico terreno de aplicación que se pretende. Consideradas de este modo, puede verse que las condiciones establecidas en el Dilema del Prisionero, intentan dar cuenta de algunos rasgos constantes de las relaciones sociales empíricas (en situación de conflicto de intereses), con la finalidad de ayudar a su mejor interpretación y eventualmente a su predicción.

Desde esta perspectiva metodológica hay una primera condición del modelo que debe ser considerada: es la referida a la calidad de los jugadores. Los mismos se suponen actores racionales, supuesto bastante plausible en situaciones de un juego en el que se disputan ganancias. La idea de racionalidad en este contexto se asocia con la de búsqueda por cada jugador de la máxima utilidad esperada, y se orienta por la llamada matriz de pagos. Para poder caracterizar cuál es la «máxima utilidad», el actor tiene la capacidad de ordenar sus preferencias de modo consistente. Aunque pueda considerarse que esta versión de la racionalidad es demasiado simplificada e ingenua, sin embargo, ofrece una regla limpia de ambigüedades y a la vez refleja satisfactoriamente los rasgos mínimos del comportamiento de un jugador en disputa de un juego, y también del de los actores sociales en disputa de un bien determinado. Estos rasgos mínimos corresponden a una capacidad de los actores para efectuar comparaciones,

y producir decisiones independientes que se orientan hacia el fin de maximizar las propias utilidades. Sin tales rasgos mínimos, bajo ninguna circunstancia se podría hablar de «elección» de una estrategia determinada. Esta condición de los jugadores contribuye a que el Dilema del Prisionero sea un dilema, ya que les permite expresar su racionalidad individual sin restricciones, lo que los lleva a una solución «irracional» en términos de utilidades.

Una segunda condición establecida en el Dilema del Prisionero, cuya plausibilidad es necesario considerar, es la referida a los pagos. Si a la solución de la casilla (D,D) se le llama No-Cooperativa, a la (—D,—D) Cooperativa, a la (D,—D) de Traición, y a la (—D,D) de Explotación para cada jugador, los pagos deben guardar la siguiente relación:  $T > C > NC > E$ . Esta condición establecida a los pagos también luce muy razonable, ya que refleja el hecho de que normalmente cuando alguien traiciona la confianza de su contrincante recibe la mayor ganancia posible; aún mayor que si coopera con su oponente y logra que éste a su vez coopere. Paralelamente, la peor utilidad posible es la de quien es explotado por su contrincante a través de una traición. Estas circunstancias son precisamente las que golpean la estabilidad de la casilla C; ya que aunque la misma ofrece un pago superior a NC, para converger a ella cada jugador debe vencer no sólo la desconfianza a ser explotado, sino también la propia tentación de traicionar.

Finalmente hay una condición relativa a que no pueden producirse acuerdos forzosos o reforzados entre los jugadores; y para garantizar esto, se establece la incomunicación absoluta de los mismos. Esta condición impuesta en el Dilema del Prisionero pretende dar cuenta segura en primer lugar de la autonomía intelectual de los jugadores para conocer, evaluar y seleccionar la mejor estrategia; y en segundo lugar, de la autonomía de sus voluntades para dirigir efectivamente su acción en este sentido. También este supuesto parece razonable en un mundo en el que no interviene ninguna autoridad superior a los jugadores para inducir alguno de los resultados en el juego<sup>8</sup>.

---

<sup>8</sup> Respecto de la incomunicación queda un punto interesante de dilucidar en el planteamiento, ya que el fiscal, además de la influencia que ejerce sobre las conductas a través de la constitución de la matriz de pagos (lo cual en términos intuitivos hace pensar en una autoridad que crea el juego con sus reglas, y luego se niega a otra intervención correctora) también debe ser pensado como un jugador (con sus propias preferencias) que se mueve entre uno y otro prisionero, produciendo una cierta forma de comunicación, y ¿por qué

Acerca de esta condición en el Dilema del Prisionero, Harsanyi sostiene que con sólo mantener la prohibición de los acuerdos forzosos el Dilema persistía, aun cuando se permitiera a los jugadores comunicarse entre sí e incluso lograr acuerdos convencionales, ya que nada (excepto el acuerdo de cumplimiento forzoso) puede garantizar la elección efectiva de la salida C (Harsanyi-76). Desde luego que la permisión de un acuerdo forzoso previo al juego disolvería de manera expresa el carácter no-cooperativo del Dilema, y lo convertiría en un juego no estrictamente competitivo y cooperativo. Pero no se puede negar que la permisión de la comunicación previa al juego entre los jugadores, aun a través de terceros, abriría la posibilidad (aunque no puede, desde luego, dar garantía plena) para la disolución del Dilema. Esta es la razón por la cual todos los teóricos de los juegos han mantenido, junto a la prohibición de los acuerdos de cumplimiento forzoso, la prohibición de comunicación entre los jugadores. Por otra parte, la propia dinámica del juego hace que ni la exhibición de la estrategia que el jugador piensa utilizar, ni la planificación de estrategias conjuntas produzcan beneficios a cada jugador, con lo que la incomunicación tiende a autorreforzarse.

Como puede verse, tanto la condición de racionalidad exigida a los jugadores, como la relación que deben guardar entre sí los pagos de la matriz, y la prohibición de acuerdos de cumplimiento forzoso o de comunicaciones previas al juego entre los jugadores, no contradicen lo que puede considerarse razonable en una relación de conflicto de intereses entre actores sociales coherentes en sus preferencias. Esta plausibilidad del modelo formal del Dilema del Prisionero, además de su coherencia lógica, es sin duda la razón por la cual su análisis y empleo en el terreno de la ciencia política, la ética normativa y la economía es cada vez más frecuente y exitoso.

#### **4.1. Nuevos modelos analíticos**

Dentro del marco de los pagos definidos y de las condiciones señaladas, el Dilema del Prisionero es un juego no estrictamente competitivo no cooperativo, y su único punto de equilibrio es la

---

no? de ruido. Tal como queda planteada la matriz es lícito suponer que el fiscal desea y busca que los prisioneros se delaten mutuamente, de modo que aunque de hecho no figura en el propio escenario del juego, su presencia es inevitable y en la sola presentación de la matriz hay un claro mensaje.

salida (D,D), que representa la intersección de las mejores estrategias posibles de cada jugador. Sin embargo, y de allí el dilema, la salida (—D,—D) ofrece un mejor pago, aunque es carente de estabilidad.

Procurando escapar a ese dilema, diferentes autores han planteado modelos analíticos alternativos a través de la variación de alguna de las condiciones establecidas, aunque sin llegar a transformarlo de manera directa en un juego cooperativo. Es necesario admitir que en cada caso se construye un juego diferente del Dilema original, con otros supuestos, y que intenta siempre constituir a la salida cooperativa (la de máxima ganancia colectiva) en una solución estable. Las alternativas ensayadas pasan, de una manera explícita o implícita, bien por el camino de modificar la definición de la racionalidad de los jugadores, bien por el de modificar la matriz de pagos.

Si se observa la estructura del Dilema del Prisionero puede verse que consta de una Premisa-Conjunción, una Premisa-Disyunción, y una Conclusión; y como todo dilema tiene la característica de aprisionar el razonamiento en una contradicción sin aparente salida.

La Premisa-Conjunción en el Dilema del Prisionero se puede expresar así: *Si el jugador elige la mejor alternativa obtiene como pago irracional, y si no elige la mejor alternativa, actúa irracionalmente.* La Premisa-Disyunción se expresa del siguiente modo: *La elección será de la mejor alternativa o no lo será.* La conclusión reza: *Por lo tanto, la elección o es irracional o produce efectos irracionales.* Como la Premisa-Disyunción hasta el momento ha sido considerada inatacable, porque las dos estrategias señaladas en la misma son mutuamente excluyentes y agotarían las posibilidades de elección, la literatura se ha inclinado por atacar el dilema mediante el recurso llamado de «coger el dilema por los cuernos», que consiste en negar la Premisa-Conjunción, negando alguna de sus partes; lo cual implica en el caso del Dilema del Prisionero, o revisar el concepto de racionalidad, o revisar el sistema de pagos de la matriz.

Los que procuran la negación de la Premisa-Conjunción por vía de la modificación del concepto de racionalidad, trabajan con la idea general de que el principio de maximización de la utilidad esperada puede ser «racionalmente» limitado, para lo cual necesitan apelar a un nuevo concepto de racionalidad; mientras que quienes buscan negar la premisa a partir del cambio del sistema de pagos, procuran armonizar la intersección de las mejores estrategias (punto de equilibrio), con la salida cooperativa, apelando a la modificación de la matriz de ganancias.

Los primeros se apoyan en la creencia de que la definición de racionalidad podría modificarse, si se consideran los niveles de evolución moral de nuestro tiempo, en los que la autolimitación en la búsqueda de una maximización de la utilidad esperada, con frecuencia se considera perfectamente «racional». Los segundos encuentran su base argumental en el hecho de que sólo mediante una intervención similar a la del fiscal cuando establece los pagos, se podría lograr que la casilla (—D,—D) cobrara estabilidad. La posición de estos últimos luce más clara y directa; porque en realidad los primeros, aunque de una manera sutil, también finalmente modifican los pagos; pero como supone una intervención autoritaria externa al juego mismo, muchos temen aceptarla.

Es importante poner de relieve que mientras en la solución que procura modificar el concepto de racionalidad se percibe la prevalencia de elementos propios del uso normativo del modelo, en la que considera que sin modificación de los pagos no habrá cambios en el punto de equilibrio, se evidencia un enfoque básicamente analítico.

Ubicada en el primer tipo de vía de acceso al problema se puede considerar la teoría de los metajuegos (Howard-71) que intenta encontrar un punto de equilibrio cooperativo ampliando el concepto de estrategia hasta incluir en el mismo las posibles respuestas de éste a las elecciones estratégicas condicionales de aquél. Se trata entonces de la estrategia para seleccionar una estrategia, lo que equivale a producir una regla de elección a nivel metaestratégico. Puede decirse que mientras a nivel de juego se hace uso de la estrategia pero con la vista focalizada en la utilidad esperada, en el plano del metajuego la focalización se dirige hacia las estrategias como objeto, con miras a seleccionar una de ellas.

El desplazamiento de plano desde el nivel del juego al del metajuego, pone en clara evidencia los elementos normativos presentes en este enfoque. El jugador se plantea la estrategia no sólo como un medio para maximizar sus utilidades esperadas, sino que le asigna un valor diferente según sea cooperativa o no. Este hecho no debe provocar intranquilidad alguna; pero obligaría a plantear una discusión crítica acerca de la fundamentabilidad racional del valor seleccionado.

Si a nivel metaestratégico se considera que —D equivale a Cooperar, en el caso del Dilema del Prisionero con la matriz de pagos anteriormente planteada, para el jugador A caben las siguientes reglas o metaestrategias:

1. Elección incondicional de C.
2. Elección incondicional de D.
3. Elección de C si el otro elige C, y elección de D si el otro elige D.
4. Elección de C si el otro elige D, y elección de D si el otro elige C.

La matriz de pagos para estas alternativas, tendrá ocho soluciones:

		Meta Estrategia del jugador A			
		1	2	3	4
Estrategia del jugador B	C	8,8	0,10	8,8	0,10
	D	10,0	<b>2,2</b>	2,2	10,0

En esta primera expansión el punto de equilibrio (2,2) es no cooperativo.

<sup>9</sup> La matriz de pagos correspondiente a dos expansiones es la siguiente:

		RESPUESTA DEL JUGADOR B			
		1	2	3	4
METAESTRATEGIA DEL JUGADOR A	C/C/C/C	(8,8)	(0,10)	(8,8)	(0,10)
	C/C/C/D	(8,8)	(0,10)	(8,8)	(10,0)
	C/C/D/C	(8,8)	(0,10)	(2,2)	(0,10)
	C/D/C/C	(8,8)	(2,2)	(8,8)	(0,10)
	D/C/C/C	(10,8)	(0,10)	(8,8)	(0,10)
	C/C/D/D	(8,8)	(0,10)	(2,2)	(10,0)
	C/D/C/D	(8,8)	(2,2)	<b>(8,8)</b>	(10,0)
	D/C/C/D	(10,0)	(0,10)	(8,8)	(10,0)
	C/D/D/C	(8,8)	(2,2)	(2,2)	(0,10)
	D/C/D/C	(10,0)	(0,10)	(2,2)	(0,10)
	D/D/C/C	(10,0)	(2,2)	(8,8)	(0,10)
	C/D/D/D	(8,8)	(2,2)	(2,2)	(10,0)
	D/C/D/D	(10,0)	(0,10)	(2,2)	(10,0)
	D/D/C/D	(10,0)	(2,2)	<b>(8,8)</b>	(10,0)
	D/D/D/C	(10,0)	(2,2)	(2,2)	(0,10)
	D/D/D/D	(10,0)	<b>(2,2)</b>	(2,2)	(10,0)

Como puede observarse hay tres puntos de equilibrio, dos de los cuales son cooperativos y uno es no cooperativo.



Si el jugador B puede a su vez predecir esta selección condicional de las estrategias efectuada por A, está en condiciones de elaborar sus propias metaestrategias, en un nivel aún superior. Las alternativas en este caso serán 16, y la matriz de pagos tendrá 64 soluciones<sup>9</sup>. Mientras que en la primera expansión Howard consigue 1 sólo punto de equilibrio y es N. C., en la segunda expansión logra 3 puntos de equilibrio, de los cuales 2 corresponden a C. En un juego de 2 personas no son necesarias ulteriores expansiones, ya que no se evidenciarían nuevos puntos de equilibrio, puesto que como regla general para un juego de  $n$  personas,  $n$  expansiones ponen de manifiesto todos los puntos de equilibrio.

Desde el punto de vista de su lógica interna, este enfoque resulta coherente, y lo es precisamente porque ha desplazado el objetivo desde la maximización de la utilidad esperada hacia la selección de una estrategia, buscando con ello escapar a los efectos del dilema. Se trata estrictamente de un juego acerca de las posibles estrategias a emplear por los prisioneros, y acerca de las ganancias correspondientes a cada intersección de estrategias condicionales. Sin embargo, como el propósito del enfoque no es el de construir un nuevo juego vacío, sino el de determinar bajo qué condiciones de nueva racionalidad se puede superar el carácter no cooperativo del Dilema del Prisionero, el juicio no puede ser solamente interno, sino que tiene que superar también la crítica metodológica.

En este aspecto, la teoría de los metajuegos propone un escenario fuera del juego, en el que la comunicación entre los jugadores no está prohibida, y en consecuencia puede suponerse que hay un intercambio de información que contribuirá a hacer más previsibles las elecciones. De este modo, la selección condicional de estrategias puede apoyarse en predicciones que disminuyen los riesgos de traición. Se puede ir más lejos aún, y suponer que en esa instancia del metajuego se construyen acuerdos tendientes a garantizar las selecciones cooperativas; pero una vez construidos esos acuerdos, se abren dos alternativas: o se los torna obligatorios (o se los refuerza) modificando la matriz de pagos, o se los deja librados al cumplimiento de las partes, con lo que el equilibrio que se exhibe a nivel de metajuego no tiene por qué reflejarse a nivel del juego. Por eso esta teoría de los metajuegos a pesar de que parece estar centrada en el tránsito de una racionalidad utilitarista a una cooperativa, y efectivamente al colocarse en el metaplano está proponiendo otra definición de racionalidad (capaz de mirar el juego desde un nivel superior) si quiere hacerse efectiva en el juego no tiene otra alternativa que

la de reflejar ese cambio en los pagos, o renunciar al mismo<sup>10</sup>. Si la matriz de pagos se mantiene sin modificaciones esta nueva racionalidad que no sólo es capaz de ordenar coherentemente sus preferencias individuales, sino también de ordenar valorativamente el tipo de estrategia a emplear, no parece suficiente para resolver los problemas de convergencia estable a la casilla cooperativa sin renunciar a nivel de juego a la maximización de la utilidad esperada.

Entre las soluciones que centran su atención en una transformación del concepto de racionalidad, se encuentra la llamada paradoja de Newcomb<sup>11</sup> analizada por primera vez por Robert Nozick. En su versión original se refiere a una decisión individual bajo riesgo, y plantea la siguiente situación: hay dos cajas, B1 y B2. La caja B1 contiene 1.000 dólares, y la caja B2 contiene 1.000.000 de dólares o nada, y el decisor ignora cuál de ambas circunstancias está presente efectivamente en B2, y debe decidir entre:

1. Toma lo que está en ambas cajas.
2. Toma lo que está en B2.

Pero lo que efectivamente haya en B2 depende de un Ser Superior que posee la facultad, suficientemente comprobada, de predecir con altísima probabilidad de acierto la selección que efectuará el decisor. Si el Ser Superior prevé que el decisor tomará sólo B2, en la caja habrá 1.000.000 y si prevé que tomará B1 y B2 no habrá nada.

La matriz de pagos será la siguiente:

		SER SUPERIOR	
		Predice que el decisor tomará sólo B2	Predice que el decisor tomará sólo B1 y B2
Toma B2 DECISOR	Toma B2	1.000.000	0
	Toma B1 y B2	1.001.000	1.000

<sup>10</sup> A este respecto, Howard y Harsanyi han sostenido una aguda y fecunda polémica, cuyos puntos de mayor interés pueden encontrarse en «Communications», *American Political Science Review*, 68.

<sup>11</sup> Esta paradoja fue formulada en 1960 por el físico William A. Nescomb y tuvo un brillante desarrollo y análisis de parte del filósofo de Harvard Robert Nozick-69. Más tarde ha recibido mucha atención por parte de matemáticos y

El conflicto en esta paradoja reside en el hecho de que bajo la condición de alta capacidad predictiva que exhibe el Ser, mientras el principio de utilidad esperada indica tomar sólo B2, el principio dominante indica tomar B1 y B2. Tal como ha sido señalado en los análisis de la paradoja efectuados por los matemáticos<sup>12</sup>, hay buenas razones para justificar cualquiera de las selecciones: mientras que quien elige sólo B2 piensa que el Ser predice casi con certeza la elección, y que en consecuencia, si elige B2 tiene alta probabilidad de ganar un millón, quien se inclina por tomar B1 y B2, argumenta que de cualquier modo las cantidades ya están colocadas en las cajas y que al tomar ambas, si en B2 hay 1.000 por lo menos ganará eso, y si en B2 hay 1.000.000 su ganancia aumentará a 1.001.000.

El planteamiento de Newcomb (simetrizado para hacerlo aplicable a la teoría de los juegos), representa un enfoque interesante inscrito en el intento de elaborar un nuevo concepto de racionalidad que no sacrifique el principio de maximización de la utilidad, y al mismo tiempo estimule el comportamiento cooperativo. Esta nueva racionalidad posee un buen sistema de información, y a la vez métodos y técnicas apropiadas que le permiten prever con buen nivel de acierto los cursos de acción que tomará el jugador opuesto, y conforme a ello orientar su propio curso de acción.

Aplicado el enfoque (en su versión simetrizada) al Dilema del Prisionero, la nueva racionalidad dota a uno de los jugadores (líder) de la capacidad para tomar la iniciativa de «proponer» al otro jugador (seguidor) una cooperación condicional. Por otra parte, como trabaja con la condición de una fuerte previsibilidad, cada jugador sabe que su estrategia condicional será conocida por el otro con alta probabilidad, sirviéndole esto de incentivo para escoger a su vez la estrategia cooperativa. Esta condición de mutua predicibilidad pretende ofrecer dentro de un marco probabilístico, un nuevo escenario en el cual los jugadores, al par que corren menos riesgos de ser traicionados, desarrollan también una propensión a no defraudar.

Si se analiza la solución propuesta por este enfoque al Dilema del Prisionero, se percibe que efectivamente a nivel de metajuego es posible someter a prueba (con toda la precisión que se

---

expertos en Teoría de los Juegos. En la actualidad es frecuentemente discutida y utilizada en el terreno de la Ética, especialmente en el tratamiento analítico de temas de construcción de políticas públicas.

<sup>12</sup> Especialmente interesantes son los análisis efectuados por M. Gardner-3, G. Schlesinger-74 e I. Levi-75.

deseo) la previsibilidad de las selecciones de los jugadores. En este sentido es muy claro que el gran desarrollo de las técnicas de seguimiento empírico, y el no menos amplio de los sistemas de información, contribuyen positivamente en esta empresa, fortaleciendo la plausibilidad del supuesto; pero luego de efectuado este paso, el único modo de dotar de equilibrio la solución cooperativa consistirá en hacerla forzosa, o reforzarla a través de la modificación de la matriz de pagos, sometiendo a sanciones las violaciones. El modelo de Newcomb (simetrizado) pretende escapar a esta dificultad estableciendo un líder con capacidad para proponer una cooperación condicional; pero este solo hecho demuestra que la nueva racionalidad, además de ser explícitamente definida como una racionalidad bien informada y buena predictora (lo que es una cuestión de hecho perfectamente aceptable), es también **implícitamente supuesta** como una racionalidad cooperativa; y esto no resulta aceptable como principio, sin efectuar antes una importante discusión crítica. Pareciera que esa nueva racionalidad a pesar de su buena información y su fuerza predictora, carece de la capacidad necesaria para autosostenerse de manera estable, frente a la racionalidad maximizadora de la utilidad esperada.

Hay que destacar que este enfoque posee el mérito indiscutible de haber hecho notar la importancia de la información en la construcción de los patrones psicológicos de decisión, y del poder que la misma exhibe como elemento persuasivo de las conductas. Por otra parte, también se ha preocupado de desarrollar los cálculos necesarios para efectuar las mutuas provisiones de las selecciones; pero se percibe que todos estos hechos no resultan suficientes para garantizar la estabilidad de la solución cooperativa, ya que ha sido necesario apelar a la preexistencia de un valor moral que la garantice. Este enfoque, a su vez, va a influir de manera considerable en muchos de los conceptos básicos que maneja la teoría que se analizará en seguida.

Uno de los intentos más importantes efectuados para atacar la Premisa-Conjunción del dilema mediante la apelación a un nuevo concepto de racionalidad, que no sea maximizante de la utilidad esperada, es el que ha realizado Gauthier en el marco de su propuesta para la construcción de una moral a través del acuerdo (D. Gauthier-86). Esta nueva racionalidad consiste, según Gauthier, en una disposición moral de los jugadores para actuar cooperativamente bajo ciertas y determinadas circunstancias.

La nueva racionalidad propuesta por Gauthier se caracteriza por tres rasgos: primero, es una disposición a utilizar estrategias

conjuntas y no individuales; segundo, esa disposición se encuentra condicionada por una expectativa de beneficio o utilidad, y tercero, persigue una maximización limitada que se apoya en el principio de la concesión relativa mini-max.

Según puede verse, Gauthier coloca al jugador en la posición de decidir entre dos disposiciones: una que representa la racionalidad que puede llamarse clásica en el Dilema del Prisionero, y la nueva racionalidad (la otra disposición) que ya no estaría dirigida a una maximización pura de la utilidad esperada, sino a una maximización limitada de la misma. La selección de esta segunda disposición perseguiría la obtención de resultados muy próximos al óptimo. El decisor portador de esta nueva racionalidad, no sólo puede ordenar coherentemente sus preferencias, sino que también está en condiciones de ordenar jerárquicamente los beneficios derivados de utilizar una estrategia individual o una de conjunto, y puede actuar en consecuencia. Además, su función de utilidad no persigue la maximización pura de las utilidades esperadas, sino que, definido un entorno alrededor de ella, acepta un resultado que se acerque a lo óptimo, definiendo una maximización limitada.

El cambio en la función de utilidad, producida en virtud de la concesión relativa, y del principio de maximización limitada, llevaría al jugador a seleccionar no el resultado que le produce la mayor expectativa de ganancia, sino aquel que le permite utilizar una estrategia conjunta, aunque los beneficios sólo sean próximos al óptimo.

Gauthier coloca al decisor en la situación de seleccionar entre dos disposiciones: una que llama de maximización pura, y la otra de maximización limitada. Para tomar esa decisión el actor necesita previamente determinar el nivel de racionalidad de la maximización limitada. Gauthier desarrolla los argumentos que podrían esgrimirse a ese respecto en uno y otro sentido (Gauthier-86, pág. 170); pero antes establece dos condiciones: 1) La situación concreta en la que se tiene que elegir entre una disposición y la otra, debe ofrecer la posibilidad de la cooperación mutuamente beneficiosa y equitativa. 2) También debe ofrecer la posibilidad de beneficios para el individuo que traiciona.

Con la primera condición, Gauthier quiere dotar de sentido moral a la autolimitación, y con la segunda, eliminar de la consideración los casos en que no es necesaria. Conforme a estas condiciones, el sujeto que va a decidirse entre una disposición y la otra, visualiza las siguientes alternativas de utilidad esperada:

u: Si cada uno actuara conforme a una estrategia individual.

$u'$ : Si todos actuaran conforme a una estrategia cooperativa conjunta.

$u''$ : Si él actuara conforme a una estrategia individual y los demás conforme a una estrategia cooperativa conjunta.

En virtud de la primera condición establecida  $u < u'$  (carácter beneficioso de la cooperación) y  $u' < u''$  (carácter beneficioso de la traición). Con esto, Gauthier establece con gran claridad que su interés es disolver el dilema sin modificar expresamente los pagos de la matriz, y para ello utiliza con gran ingenio los recursos de la estadística bayesiana (L. J. Savage-51 y 54). Entrando en los argumentos del decisor, para inclinarse por la maximización pura o por la limitada, tendríamos:

### Argumento 1

*Supongamos que adopto la maximización pura.* Entonces, si espero que los demás fundamenten sus acciones en una estrategia conjunta, abandono mi mejor estrategia individual, y espero una utilidad  $u''$ . Si los otros adoptan estrategias individuales yo hago lo mismo, y espero una utilidad  $u$ . Si la probabilidad de que los demás basen sus acciones en una estrategia conjunta es  $p$ , mi utilidad esperada global es  $[pu'' + (1-p)u]$ .

*Supongamos que adopto la maximización limitada.* Entonces si espero que los demás funden sus acciones en una estrategia conjunta hago lo mismo y espero una utilidad  $u'$ . Si espero que los demás actúen con estrategias individuales hago lo mismo y espero una utilidad  $u$ . Con lo cual mi utilidad esperada global es  $[pu' + (1-p)u]$ .

Como  $u'' > u$ , entonces  $[pu'' + (1-p)u] > [pu' + (1-p)u]$ , para cualquier valor de  $p$  diferente de 0; siendo igual para  $p = 0$ .

En consecuencia, para maximizar mi expectativa global de utilidad tendría que seleccionar la maximización pura.

### Argumento 2

*Supongamos que adopto la maximización pura.* Debo esperar que los demás utilicen estrategias individuales de maximización pura y hago lo mismo; luego espero una utilidad  $u$ .

*Supongamos que adopto la maximización limitada.* Entonces si los demás fundamentan sus acciones en una estrategia cooperativa conjunta, hago lo mismo y espero una utilidad  $u'$ . Pero si los demás no tienen esa disposición, utilizo una estrategia maximizadora,

y espero  $u$ . Si la probabilidad que los demás estén dispuestos a una maximización limitada es  $p$ , mi expectativa global de utilidad es  $[pu' + (1-p)u]$ .

Como  $u' > u$ ,  $[pu' + (1-p)u] > u$  para cualquier valor de  $p$  distinto de 0 e igual para  $p = 0$ .

Luego para maximizar mi utilidad esperada global debo adoptar la maximización limitada.

Según se observa, ambos argumentos conducen a conclusiones contradictorias y en consecuencia no pueden ambos ser válidos bajo las mismas condiciones. Sin embargo, considerando cuidadosamente el segundo argumento, puede notarse que las condiciones han variado, ya que los maximizadores autolimitados interactúan con quienes tienen su misma disposición de una manera, y de forma muy diferente con los maximizadores puros. De modo que los que exhiben una tendencia pertinaz a la racionalidad maximizadora pura, serán excluidos del trato y de los beneficios de la cooperación.

Como puede verse, a pesar de que el objetivo de Gauthier es el de atacar el dilema mediante la construcción de una nueva racionalidad, a la hora de plantearse el problema de la estabilidad de la solución, necesita apelar a una modificación de la matriz de pagos; pero aun así, al no aceptar tal cambio de una manera expresa sino como un proceso, no hay posibilidades de garantizar la estabilidad, si no se utiliza un recurso punitivo (reforzamiento) contra quienes procuran maximizar la utilidad esperada, es decir, contra quienes persisten en la racionalidad utilitarista. Aunque Gauthier sostiene que cuando el jugador se inclina por la disposición a la estrategia conjunta, induce un cambio en las utilidades e induce a los otros jugadores a efectuar una selección análoga, el crucial problema de las eventuales traiciones/explotaciones subsiste, y es por eso que se ve obligado a apelar al concepto de «translucencia», cualidad que permite descubrir la disposición de los jugadores, a fin de eliminar del juego a los que tienen una propensión pertinaz a la antigua racionalidad. Los jugadores, según Gauthier, no son ni totalmente transparentes ni completamente opacos, sino que son traslucidos, es decir, que sus decisiones tienen un cierto nivel de previsibilidad. Esto, unido a la capacidad de los jugadores para leer esa translucencia permite elaborar estrategias conjuntas que van generando un clima donde se impone la nueva racionalidad.

Sin embargo, a pesar de todos estos supuestos, la salida cooperativa sigue siendo inestable, ya que aun cuando se logre construir un acuerdo de estrategias conjuntas, y aun cuando se logren algunos convenios, es imposible garantizar el cumplimiento

de los mismos sin apelar a un refuerzo externo a los jugadores. En el momento en que la traición produce la expectativa de una gran ganancia, no existiría ninguna garantía de que uno de los jugadores no abandone su propensión a la estrategia conjunta.

En respuesta a esto, y aplicando probabilidades bayesianas, Gauthier llega a establecer que en un mundo con un número suficiente de maximizadores limitados, la solución cooperativa podría llegar a ser estable. Esto en razón de que cuando en el juego uno de los actores selecciona la estrategia cooperativa, no sólo produce un efecto directo sobre el juego, sino que también modifica en un sentido positivo la probabilidad de que el otro jugador seleccione a su vez una estrategia cooperativa. De esta manera, tras sucesivos regateos, la matriz de pagos habrá cambiado (sin intervención de ninguna autoridad externa al juego), al asociarse los pagos brutos a una probabilidad específica. Esto equivale a decir que el comportamiento cooperativo se habrá autorreforzado.

El intento de Gauthier tiene algunos méritos metodológicos indiscutibles: entre ellos se puede destacar en primer lugar, el del tratamiento riguroso aplicado al ámbito de la Ética de la Teoría de los Juegos asociada a la Teoría de la Probabilidad, lo que abre vastas perspectivas analíticas; y en segundo lugar el de la incorporación del concepto de traslucencia, vinculado a la llamada probabilidad subjetiva, que permite pensar razonablemente en un ulterior tratamiento formal. Sin embargo, en la definición analítica de la nueva racionalidad, no ha obtenido el mismo éxito. La racionalidad maximizadora autolimitada no es suficiente para sostenerse y dotar de estabilidad cooperativa al juego: una muestra de ello está claramente señalada en la condición 1—, establecida para que tenga sentido la cooperación, acerca de que la cooperación debe ser beneficiosa y equitativa. Aunque el beneficio es una cuestión que puede expresarse en términos de utilidades, la equidad representa un incentivo exterior al juego mismo, y aparece como imprescindible para el desarrollo de esa propensión a la racionalidad maximizante limitada. Por otra parte, la necesidad de punir a los jugadores que no se adhieren a la nueva racionalidad, pone en evidencia que ambas racionalidades están en pugna y que el único modo de modificar la relación de fuerza entre ellas es dotando a la nueva racionalidad de un poder de previsión y de sanción (semejante al del Ser Superior de Newcomb) que escapa a los límites analíticos del juego. Es evidente que Gauthier cree firmemente en la superioridad ética del nuevo *proviso*, pero esto no es suficiente para convertir en



estable la solución cooperativa del dilema, en un mundo donde el famoso fiscal ha dejado tan complicadas las cosas para los maximizadores limitados.

Los modelos analíticos considerados hasta el momento se caracterizan por intentar la disolución del Dilema del Prisionero a partir de la definición de una nueva racionalidad. El otro grupo considera que tal disolución sólo puede lograrse mediante una modificación expresa en la matriz de pagos, que refuerce la salida cooperativa.

Los autores que sostienen la postura de que sólo un cambio en la matriz de pagos puede dotar de estabilidad a la salida cooperativa, consideran que para que tal resultado se produzca es necesario alterar la relación:  $T > C > NC > E$ , señalada como una condición por el dilema. Sin embargo, como la modificación de la matriz de pagos aparece como un acto de alguien externo al juego que estaría induciendo u obligando a elegir una salida cooperativa, no todos los autores se deciden a aceptar de una manera clara e irrestricta este enfoque, seguramente en razón del temor a que esa aceptación pueda tener consecuencias normativas indeseables en el terreno de la política práctica. Este es un asunto que preocupa fundamentalmente a los planificadores de políticas públicas en los estados democráticos, que se ven enfrentados a la necesidad de estimular comportamientos cooperativos, sin vulnerar la libertad de elección individual de los actores sociales. Por esta razón, aunque ni la promoción de acuerdos, ni la traslucencia, ni la previsibilidad de los comportamientos han podido por sí mismos evitar que el punto de equilibrio sea analíticamente no cooperativo, no resulta demasiado fácil aceptar de manera directa, que la modificación de la matriz sea la única solución al dilema, y menos aún que tal solución deba venir desde fuera del juego. Este hecho ha determinado que entre los autores que intentan disolver el Dilema del Prisionero mediante una modificación de la matriz de pagos, encontramos algunos que procuran alcanzar su objetivo a través de intrincados rodeos conceptuales que tienden a desdibujar la presencia de quien efectúa la modificación de los pagos; y otros, muy pocos por cierto, que deciden asumir un camino más directo. Consideraremos dos planteamientos (uno de cada clase) que abordan el problema desde el ángulo de la modificación de la matriz.

A la vista de que los acuerdos no bastan para impedir los resultados negativos de una convergencia a NC (Hoerster-75), procura que se modifique de una manera artificial la jerarquía de las ventajas. Para ello comienza presentando un modelo que mantiene la jerarquía de ventajas que estimula la no cooperación

y que en su versión simplificada tiene la siguiente forma<sup>13</sup>:

1. A actúa egoístamente — los demás actúan altruístamente
2. A actúa altruístamente — los demás actúan altruístamente
3. A actúa egoístamente — los demás actúan egoístamente
4. A actúa altruístamente — los demás actúan egoístamente

Las ventajas personales de A disminuyen desde 1 a 4; esto es que la jerarquía de utilidades es la siguiente:  $1 > 2 > 3 > 4$ . Por lo tanto, actuar egoístamente beneficia a A. Pero como los otros jugadores piensan de igual manera, el equilibrio se presenta en la alternativa 3. El dilema consiste en que la alternativa 2 ofrece mayores ganancias, pero parece inalcanzable sin la presencia de un elemento que la fortalezca. Hoerster considera que un acuerdo para actuar altruístamente entre los jugadores no sería suficiente, ya que siempre un jugador podría violarlo, con lo que demostraría su astucia al beneficiarse con el pago correspondiente a la alternativa 1—. Por eso propone modificar la matriz de pagos, mediante lo que él llama un desplazamiento artificial de la jerarquía de las ventajas. La modificación de los pagos tendría que hacerse en el sentido de que las alternativas 2 y 4, en las que A actúa altruístamente, le sean más favorables que las 1 y 3, en las que actúa egoístamente.

Desde luego esto equivale a reforzar el acuerdo cooperativo, a través de una intervención ajena a los jugadores, situación que como se dijo no todos están dispuestos a asumir. Para tratar de sortear esta dificultad, Hoerster propone un ejemplo que a la postre termina siendo la mejor prueba de la inutilidad de su intento. Se trata de un rico anciano que desea que sus dos hijos se reconcilien tras años de distanciamiento, para lo cual los invita a la celebración de su cumpleaños, y les comunica que después del mismo redactará su testamento. Si ambos vienen a la fiesta la

---

<sup>13</sup> Hoerster incorpora a su exposición las palabras egoísta y altruista para calificar las salidas de NC y C, respectivamente, lo cual coloca a su ejemplo una importante carga no sólo valorativa, sino también emocional. Esto se evidencia en la gran dificultad que tiene para aceptar que si se modifica la matriz de pagos (por acuerdo o a través de una autoridad) surge un nuevo punto de equilibrio. Para él lo que surge es una nueva conducta egoísta, ya que aunque se trate de la casilla C, las razones que llevan allí a los jugadores es el propio interés y no el altruismo. Envuelto en el lenguaje valorativo, no descubre que es perfectamente coherente con la tesis utilitarista el hecho de converger a la casilla C por razones de expectativa de mayor ganancia esperada, la que esta vez coincide con la mayor ganancia colectiva. En otras palabras, que la cooperación no tiene por qué ser ejercida en virtud de sentimientos altruistas.

herencia se repartirá por partes iguales. Si no viene ninguno de los dos, cada uno heredará sólo  $1/5$  (mientras los otros  $3/5$  los heredará el ama de llaves). Si viene sólo uno, el que venga será desheredado, y el que no vino heredará  $3/5$  (los  $2/5$  restantes irán al ama de llaves); la relación que guardan los pagos puede expresarse en la siguiente matriz de ganancias, en la que las cifras corresponden a la parte de la herencia que les corresponderá a cada uno según la estrategia (concurrir o no concurrir) que elijan.

		B	
		C	-C
A	C	$1/2, 1/2$	$0, 3/5$
	-C	$3/5, 0$	$1, 1/5$

Hoerster, deseando que el reforzamiento externo de la salida cooperativa no sea demasiado evidente, ha mantenido la relación entre las ganancias, ya que  $3/5 > 1/2 > 1/5 > 0$ ; en una palabra, sigue pagando mejor el comportamiento egoísta que el altruista, y el punto de equilibrio del juego es el no cooperativo. Lo cual autoriza a pensar que si ambos son actores racionales, que quieren maximizar su utilidad esperada, ninguno de los dos concurrirá a la reunión.

Es entonces cuando Hoerster empieza a debatirse en una complicada red de suposiciones que lo llevan a pensar que los hijos se pondrán en contacto. Asimismo, también supone que de ese solo contacto surgirá **necesariamente** un acuerdo razonable acerca de la estrategia conjunta que más le conviene para garantizarse la mitad de la herencia cada uno; ya que ambos reconocerán que si los dos concurren a la reunión, las ventajas serán mayores que si no concurren. Sin embargo, cuando todo parece definitivamente bien encaminado (si se aceptan las pocas plausibles suposiciones), Hoerster señala: «Y en verdad, para evitar que el otro pueda romper la promesa y no asistir, acordarán hacer la vida *juntos*.» (*op. cit.*, pág. 136). Con una enorme elocuencia queda expresada la dificultad que se viene señalando, ya que «acordar hacer la visita juntos» significa establecer un reforzamiento de las conductas, a todas luces externo al juego. Por una parte el acuerdo de asistir (que ya es externo) no parece ser suficiente, sino que también hay que dotar a ese acuerdo de un extraordinario poder coactivo exterior al juego: «hacer la visita

juntos», que procura eliminar la posibilidad de una traición/explotación, y que impone un nuevo escenario al juego.

Por la otra parte, aunque Hoerster sostiene que las conductas altruistas de los hijos «están impuestas por el propio interés... sin que se necesite establecer un poder coactivo por encima de los dos...» (pág. 136), en realidad la maximización de su propio interés los llevaría a promover que el otro concurriera a la reunión, y al último momento dejar de concurrir, con lo que su ganancia sería la máxima de  $3/5$ . Precisamente para evitar esa «traición» es necesario establecer una coacción que no se refleja en la matriz de pagos (que incluso va más allá del acuerdo) y que prácticamente representa una imposición física de la obligación de concurrir.

El análisis de Hoerster tiene el mérito de poner de relieve la necesidad de modificar la jerarquía de las ventajas, estableciendo una relación que haga más fáciles las cosas para quienes desean actuar cooperativamente. Sin embargo, probablemente en virtud de un manejo analítico rudimentario del modelo básico, ha entrado en numerosas contradicciones. Por otra parte, la confusión demasiado acentuada entre el uso analítico y el normativo, representa un factor que ha contribuido mucho al fracaso del intento por resolver un problema, que bueno es señalarlo, no es sencillo, y representa el máximo desafío para los estudiosos de los juegos.

Entre los autores que sostienen que sólo una modificación directa de la matriz de pagos puede hacer que el dilema se convierta en un juego en equilibrio cooperativo, analizaremos el compacto punto de vista de J. Harsanyi-76 y 78, quien parte de dos aceptaciones meridianas: primera, la solución de un juego no cooperativo debe ser un punto de equilibrio; y segunda, en el Dilema del Prisionero el punto de equilibrio es no-cooperativo. Estas dos aceptaciones ahorran mucho esfuerzo intelectual, no pocas complicaciones, y evitan grandes confusiones de esferas. De estos puntos de partida es posible deducir que para disolver el dilema es necesario modificar la matriz de pagos, ya sea a través de la punición legal de la violación de los acuerdos, o a través de una minusvaloración (firmemente aceptada tanto moral como socialmente) de tales conductas. El ejemplo considerado que tiene la estructura del Dilema del Prisionero enfrenta a dos jugadores: 1 y 2, según las estrategias y pagos de la siguiente matriz:

		Jugador 1	
		B1	B2
Jugador 2	A1	2,2	0,3
	A2	3,0	1,1

En este juego el par de estrategias (A2, B2) es el punto de equilibrio, ya que representa la mejor respuesta condicional de cada jugador; es, asimismo, el único punto de equilibrio. El dilema se deriva de que utilizando esas estrategias los jugadores obtendrían una ganancia de (1,1), que es claramente inferior a la que obtendrían empleando las estrategias (A1, B1). La solución planteada es bien aceptada por los teóricos de los juegos; pero como se ha visto, no se la considera deseable desde el punto de vista ético, ya que pareciera negar la posibilidad de la cooperación sin coacción.

Considerando el problema, Harsanyi se plantea la posibilidad de un mundo en el que los jugadores sean personas extremadamente decentes, a los cuales actuar no-cooperativamente (utilizar las estrategias A2 o B2 cuando los otros usan las estrategias cooperativas A1 o B1) les produce un desagrado tal que llega a modificar su función de utilidades. Vistas las cosas en el contexto de este mundo, se produce un cambio en la matriz del juego. Por ejemplo, si ambos jugadores le asignan un valor  $-2$  (desagrado) a las estrategias no cooperativas, automáticamente se produce una variación en las intersecciones (A1, B2) y (A2, B1), ya que a la utilidad esperada de 3 unidades, habrá que sumarle el valor  $-2$  del desagrado, con lo que se tendría  $3-2 = 1$ , y la matriz quedaría establecida con los siguientes valores<sup>14</sup>:

---

<sup>14</sup> A este respecto, se podría analizar en el campo internacional lo sucedido con las Comunidades Europeas, que han acordado renunciar a una parte de su soberanía nacional de manera irreversible, a favor del surgimiento de la llamada Comunidad Europea. Esto ha producido un acatamiento generalizado de determinadas creencias morales que han modificado las expectativas de utilidad de cada comunidad; por esta razón, aunque un determinado curso de acción conduzca a mayores ganancias, puede no ser elegido porque no produce las mayores utilidades.

		Jugador 1	
		B1	B2
Jugador 2	A1	2,2	0,1
	A2	1,0	1,1

Como puede verse en el juego, que no es más un dilema, se observan dos puntos de equilibrio, de los cuales el (2, 2) es superior al (1, 1); y los jugadores no dudarían razonablemente en converger al mismo. Desde un punto de vista de la búsqueda de la mayor utilidad, este punto no necesita de ningún reforzamiento externo, ya que en la propia matriz ha quedado expresado tal reforzamiento, en la medida que la ganancia atribuida a la traición se ve disminuida por razones de un comportamiento moral generalizado. La plausibilidad de un mundo tal no preocupa a Harsanyi; él se limita a demostrar que bajo esas condiciones (de minusvaloración generalizada de los comportamientos no-cooperativos) el dilema quedaría disuelto al quedar modificada la matriz de pagos.

Pero si la condición señalada no se cumple, la otra vía para disolver el dilema consistiría en modificar la matriz de pagos a través de un cambio originado fuera del juego, que tienda a estimular las estrategias cooperativas, y desestimular las no cooperativas. Esta solución equivaldría a una nueva distribución de ganancias que sustituiría a la hecha inicialmente por el fiscal. Es importante señalar que tanto la solución del mundo hipotético como la del cambio efectuado por una autoridad, son externos al juego y suponen opciones valorativas.

Evaluando la importancia metodológica de las soluciones analíticas al Dilema del Prisionero que fueron consideradas, se observa que aquellas que abordan el problema a través de un cambio en el concepto de racionalidad tropiezan, en primer lugar, con la dificultad insalvable de la falta de plausibilidad de la nueva definición en el estadio actual de evolución moral del hombre. En efecto, para atacar esa parte de la premisa, sería necesario suponer un hombre que privilegia el logro de un acuerdo a la maximización de sus propias ganancias, lo que resulta difícil de aceptar como condición. Pero además, aunque de un complejo análisis resultara que en efecto nuestros actores tienen un grado tal de evolución ética que les permite asignar un valor especial al comportamiento cooperativo por sí mismo, este

solo hecho conduciría a la modificación de la matriz de ganancias, generando un nuevo punto de equilibrio. Sin embargo, esta vía de acceso al problema ha producido un importante rédito analítico: es el de la incorporación de la teoría de la probabilidad para el cálculo de previsibilidad de un comportamiento, lo que a su vez puso de manifiesto la importancia de la información en los procesos estratégicos.

En la otra vía de acceso, es decir, la de quienes sostienen que el dilema sólo puede resolverse mediante una modificación de la matriz de pagos, el panorama se presenta más simple, ya que el tipo de planteamiento facilita un uso estrictamente analítico del modelo. En este grupo la solución de Harsanyi, siendo la más directa y limpia, es también metodológicamente la más fecunda; ya que las condiciones que la misma establece a la vez que resultan plausibles, no entran en ningún momento en contradicción con las consecuencias que generan. La primera alternativa de solución se plantea en un mundo en el que el respeto por los acuerdos, la búsqueda de soluciones conjuntas y la limitación racional, son valores generalizados; y si esa condición se presenta, cada vez que un jugador se enfrente a otro en un esquema de juego como el del Dilema del Prisionero, él mismo modificará la matriz de pagos disminuyendo la utilidad asignada a una eventual traición. Pero si ese mundo no se da, o bien hay que admitir que el Dilema del Prisionero no tiene solución y que en consecuencia (aceptable o no moralmente) el equilibrio es no-cooperativo cuando quiere respetarse el principio de maximización individual, o bien hay que aceptar que del mismo modo que desde fuera del juego se establecieron las reglas iniciales que tornan tan difícil la cooperación, es también posible modificarlas. La mayor contribución que este enfoque del dilema ofrece, es la de permitir una caracterización analítica clara de las condiciones bajo las cuales se puede hablar de cooperación, y aquellas bajo las cuales es imposible alcanzarla.

## **5. El Dilema del Prisionero como modelo normativo**

Según fue anteriormente señalado, la diferenciación entre el uso analítico y el normativo en el Dilema del Prisionero obedece a una importante razón metodológica: las herramientas que resultan fecundas en una esfera, son inaplicables a la otra; y en consecuencia, la confusión de ambos planos sólo puede contribuir a un oscurecimiento de los objetos que en cada momento se abordan.

Así como el uso llamado analítico se dirige a describir una red de relaciones formales, con miras a elaborar categorías explicativas, y de ese modo predecir el futuro comportamiento de tales relaciones; el uso normativo procura inducir conductas consideradas socialmente deseables porque son valoradas como éticamente superiores. Este considerar el Dilema del Prisionero desde un punto de vista normativo, nos coloca automática y expresamente en el terreno de los valores; ya que cuando se habla de la conveniencia o no de preferir una salida a otra, se está privilegiando (por alguna determinada razón), una conducta sobre la otra, y ese *por alguna razón* expresa la presencia de un valor al que se adhiere.

En el caso del Dilema en su uso normativo, cuando se destaca la superioridad de la salida cooperativa porque hace máxima la ganancia conjunta, hay una adhesión a una racionalidad de segundo orden, es decir, una racionalidad que no es maximizadora de las ganancias esperadas para cada jugador, sino que busca la maximización de las ganancias sociales, aunque eso signifique sacrificar la expectativa de una mejor ganancia individual. Esta adhesión implica necesariamente un juicio de valor que debe ser analizado y justificado con herramientas conceptuales específicas.

Aunque en muchas ocasiones los juicios de valor referidos a un problema social, ético o político suelen aparecer como grupos confusos de opiniones y preferencias, si se los considera analíticamente se descubre con frecuencia que constituyen un orden de los llamados blandos, que a pesar de su falta de rigidez preservan auténticos sistemas jerárquicos de prescripciones; este hecho permite suponer legítimamente que toda vez que nos encontramos frente a una conclusión prescriptiva es que la misma se deriva de alguna premisa, también prescriptiva, que da lugar a esa conclusión. Este principio de coherencia en las derivaciones valorativas es la condición racional mínima que puede exigirse a un modelo en su uso normativo. Admitido este principio, queda automáticamente abierta la necesidad de discutir racionalmente no sólo los métodos de derivación utilizados, sino también las propias premisas, antes de aceptarlas como base de la conclusión. Tal discusión de las premisas no tiene la forma de un ejercicio vacío; ya que en el terreno normativo, que produce consecuencias prácticas tan importantes, analizar la validez de lo que se encuentra en la base de un argumento valorativo, es una obligación inevitable de racionalidad.

Por lo señalado, cuando se hace un uso normativo del Dilema del Prisionero, es necesario someter a discusión racional



(desde luego que con las herramientas metodológicas propias de la esfera) algunos aspectos sustantivos de lo que se ha llamado cooperación, formulando preguntas que contribuyan a la crítica de las premisas prescriptivas que se encuentran en la base de justificación de tal preferencia por la salida cooperativa. Desde luego que la enorme fuerza emocional que posee la palabra cooperación, y su aparentemente inequívoco sentido, no constituyen argumentos suficientes para evitar una discusión de este tipo.

Visto el Dilema del Prisionero en su uso normativo, se percibe que a través del mismo se procuran inducir conductas cooperativas, en razón de considerarlas socialmente deseables. Este juicio obedece al hecho de que dentro del marco establecido por el modelo, la estrategia conjunta de cooperación hace máximas las utilidades colectivas. En el modelo, las estrategias individuales quedan siempre encerradas en la contradicción que se produce cuando el jugador, en su deseo de maximizar la función individual de utilidad, obtiene una ganancia menor que la de la casilla cooperativa. Sin embargo, tratándose de una toma de decisión interactiva, el hecho de que uno de los jugadores decida cooperar, no garantiza por sí solo un resultado cooperativo, ya que si el otro jugador no coopera, se producirá el muy indeseable resultado de la traición-explotación. Por esta razón, desde el punto de vista práctico, aunque sea conveniente cooperar, los jugadores se inclinan a favor de la no cooperación como modo de protegerse de eventuales explotaciones. Ante esta tendencia a seguir una estrategia individual, los autores que hacen uso normativo del modelo han procurado exhibir argumentos que prueben la superioridad de las estrategias conjuntas, e incluso las ventajas que se derivan de utilizar la cooperación como regla de vida.

Como se demostró anteriormente, las restricciones establecidas al Dilema del Prisionero resultan muy plausibles en un escenario de conflicto de intereses. Esa plausibilidad ha hecho que el Dilema del Prisionero logre un amplio desarrollo como modelo analítico en muy diversas esferas. El mismo, bueno es recordarlo, se encuadra claramente entre los juegos no-estrictamente-competitivos-no-cooperativos. Sin embargo, como el uso normativo se dirige a prescribir con carácter general determinadas normas de conducta, no parece ni razonable ni práctico que el cumplimiento de esas normas implique correr sistemáticamente el riesgo de graves pérdidas, ya que esta circunstancia golpearía la eficacia y estabilidad de la prescripción. Por esta razón, quienes han utilizado el dilema en su uso normativo, se han enfrentado a la necesidad de intentar correcciones que pudieran favorecer

la estabilidad de la solución cooperativa. Este deseo de asignar a un modelo analíticamente no-cooperativo un uso normativo cooperativo es el que ha movido a ciertos autores a modificar levemente algunas de las condiciones establecidas. Tal relajamiento de condiciones tiene por finalidad no sólo justificar ese uso normativo, sino también tornarlo eficaz y razonable; en consecuencia, para que tales propósitos se cumplan es imprescindible someter las nuevas condiciones a una prudente revisión crítica.

La primera restricción que ha sido relajada es la de incomunicación entre los jugadores. Como el dilema es no-cooperativo en virtud de que están prohibidos los acuerdos pre-juego, el relajamiento procura establecer una forma de comunicación dentro del juego mismo, que minimice los riesgos que corre un jugador que desea cooperar. De esta manera se busca que la conducta cooperativa sea racionalmente aconsejable.

En general, hay acuerdo en que si el Dilema se juega en una única oportunidad, la barrera de la incomunicación no puede romperse, y los riesgos de la estrategia cooperativa son tan altos que no resulta aconsejable correrlos. En este caso el discurso normativo podrá apoyarse en ideales de sacrificio a favor de los demás, o en principios altruistas, pero no resistirá el juicio de las restricciones impuestas por la teoría de la utilidad. Sin embargo, se ha comprobado que por regla general en el mundo de las relaciones sociales o políticas, tales juegos no se juegan por una única vez, sino que tienden a repetirse a lo largo del tiempo entre los mismos actores. Esas repeticiones tienen la característica de producir una cierta forma de comunicación entre los jugadores, los cuales van exhibiendo insensiblemente sus patrones de decisión a lo largo de las sucesivas repeticiones del juego. Si se considera que analíticamente la prohibición de los acuerdos previos al juego es el elemento determinante del carácter no cooperativo del mismo, resulta lícito suponer que relajando la incomunicación entre los jugadores, será posible hacer menos riesgosa la estrategia cooperativa, y en consecuencia darle al uso normativo del modelo una base coherente con los principios inicialmente aceptados. No se trataría de una comunicación previa el juego, sino de una comunicación mediante el juego. Desde el trabajo del M. M. Flood-52, hasta el de R. Axelrod-84, en muchas ocasiones se ha procurado demostrar operacionalmente que es posible hacer coincidir las estrategias cooperativas con los resultados maximizadores, si el dilema es jugado un gran número de veces entre los mismos actores.

Axelrod considera que en tales casos el hecho de que los jugadores

deban encontrarse reiteradamente, permite que la cooperación vaya desarrollándose entre ellos. Efectuando una serie de simulaciones en las cuales se enfrentan diferentes estrategias, determina que la llamada estrategia TIT FOR TAT (comenzar cooperando y luego hacer que lo haga el oponente en la jugada anterior) tiene condiciones de estrategia ganadora, es decir, que quien la aplique acumulará una ganancia mayor que el contrincante. Para que el TIT FOR TAT sea la estrategia más conveniente (sin importar qué estrategia use el contrincante), se establecen dos condiciones: una es que el número de repeticiones del juego sea suficientemente grande; y la otra es que a las ganancias actuales se les asigne un mayor peso que a las futuras.

Para garantizar la primera condición se establece que los jugadores no pueden retirarse del juego, ni naturalmente, retirar sus ganancias: el número de veces que se repite el juego es independiente de la voluntad de los jugadores, como también lo es el momento preciso en que el juego concluye. Y para garantizar la segunda condición, se establece el llamado parámetro de descuento, que asigna un peso decreciente a las ganancias a lo largo del tiempo. Bajo esas dos condiciones los jugadores que apliquen TIT FOR TAT, estrategia inicialmente cooperativa, serán también los más favorecidos. Pero aún más que esto, su comportamiento amigable influirá sobre los otros jugadores, contribuyendo a un desarrollo de la cooperación como valor social.

Aparentemente, bajo estas restricciones, la prescripción de las conductas cooperativas (socialmente deseables), resultaría coherente con los principios básicos del modelo. Pero sucede que se han incorporado dos condiciones cuya plausibilidad no ha sido discutida. Respecto de la segunda, no podrían formularse objeciones importantes, ya que parece razonable que una ganancia obtenida actualmente tenga mayor utilidad que una ganancia futura. Pero respecto de la primera condición, es decir, que el juego deba repetirse un gran número de veces (generaciones) para que la acumulación de las ganancias favorezca al jugador cooperativo; y que por otra parte, los jugadores tengan que estar aprisionados en el juego por un tiempo cuya duración desconocen, no parecen condiciones muy fácilmente admisibles. Para que la cooperación se desarrolle en este escenario, los jugadores deben tener dos creencias básicas: una, en el largo plazo, y la otra en que las reglas del juego permanecerán constantes. En realidad, tanto la una como la otra, no parecen resistir un análisis histórico ni político. En consecuencia, nuestros actores podrán voluntariamente, y bajo su propio riesgo, cooperar si así lo

desean; pero la consecuencia que prescribe una conducta cooperativa se deriva en este caso a partir de premisas de muy dudosa plausibilidad. Para considerar la cooperación como un valor racionalmente sostenible, bajo las condiciones planteadas, habría que garantizar que tales condiciones, el valor del parámetro y las propias reglas del juego se mantendrán constantes en el largo plazo establecido; obviamente esto supone una importante intervención de una autoridad ajena al juego. Por otra parte, la prohibición de abandonar el juego excepto cuando el mismo finalice impone una obligatoriedad a todas luces inadmisibles como marco valorativo.

La otra restricción que ha sido relajada, con el fin de lograr que el Dilema del Prisionero pueda servir como modelo para inducir razonablemente conductas cooperativas, es la relativa a la maximización de la utilidad esperada. Cuando Gauthier opone el principio de maximización limitada al de maximización pura, efectúa simultáneamente la incorporación de dos condiciones relativas a las situaciones en que puede razonablemente desarrollarse la primera: una es que debe existir la posibilidad de una cooperación mutuamente beneficiosa y equitativa (*fair*), y la otra es que debe existir la posibilidad de que quien cometa una traición obtenga ventajas de la misma. Según Gauthier, si no se da la primera condición, la cooperación no tiene sentido, y si no se da la segunda, no es en absoluto necesaria.

Varias son las condiciones que en este planteamiento han sido modificadas desde el dilema original, y que en consecuencia merecen un análisis de plausibilidad. En primer lugar, el concepto de preferencia y su derivado, el de utilidad, como preferencias coherentemente ordenadas, son ampliados. Ya no tenemos solamente a un actor capaz de ordenar coherentemente sus preferencias en el marco del juego mismo, sino a alguien capaz de ordenar preferencias acerca de dos disposiciones valorativas. Una decisión que seleccione una de estas dos disposiciones implica evaluar previamente las premisas prescriptivas que se encuentran en la base de cada una de ellas. Cabría, en primer lugar, preguntarse, **¿con quién se coopera?** En el juego del Dilema del Prisionero parece estar muy claro que cooperar significa no delatar al otro prisionero; y esta claridad se deriva del supuesto (no suficientemente expresado), de que ambos prisioneros constituyen una cierta forma de sociedad, con comunidad de intereses frente al fiscal que es su común enemigo. Es evidente que sólo para quienes se sientan asociados puede tener importancia la maximización de las ganancias comunes, aun a costa de riesgos y pérdidas personales. Pero sucede que en la jerga policial, y

aun en la judicial, suele decirse que un prisionero **coopera** con la justicia cuando delata al otro, ya que proporciona a los órganos judiciales elementos necesarios para hacerla efectiva; y no es infrecuente que la sociología y la psicología de la delincuencia consideren una acción de este tipo (delación), como la manifestación de una potencial capacidad de adaptación del delincuente a las reglas sociales, o como una actitud positiva hacia las mismas.

Planteado en términos de seleccionar una disposición es bastante razonable que el prisionero se pregunte con quién le conviene o desea cooperar... si con el fiscal o con el otro prisionero..., actitud que nos permite percibir que ese **con quién** no es inequívoco. No podría, desde luego, alegarse que se coopera cuando se colocan los intereses sociales por encima de los intereses individuales; ya que, tanto cuando el prisionero delata, como cuando no delata, persigue una maximización de sus propias ganancias en relación con sus riesgos. Está claro que la determinación de quién es el verdadero oponente en un juego, y quién es el socio, es en sí mismo un problema de aprendizaje, y simultáneamente un problema de información que se escenifica fuera del ámbito mismo del juego, y que justifica un análisis de las premisas normativas correspondientes.

El otro concepto que ha sufrido modificaciones es el de función de utilidad, ya que no se trata sólo de maximizar la utilidad esperada bajo las condiciones del juego, sino de maximizarla bajo la condición del empleo de una estrategia conjunta. Para evaluar este nuevo escenario, que naturalmente rebasa al juego mismo, cabría preguntarse, **¿para qué se coopera?** Se supone que cada uno de los prisioneros quiere hacer máximas sus ganancias esperadas dentro de un marco que le ha sido previamente fijado, y cuyas reglas no ha contribuido a crear. Sabe, sin embargo, que esas reglas no le ofrecen garantía alguna acerca del tránsito por el terreno minado que representan las casillas (0,10) y (10,0), y que en consecuencia los riesgos que ese tránsito implica debe asumirlos sin ayuda alguna. Es lícito suponer que una recompensa tan grande como la ofrecida por el fiscal ha de tener una muy poderosa razón, la que naturalmente no está expresada en la limitada información que los prisioneros poseen. Esta poderosa razón es que el fiscal también es parte del juego, y ha fijado unas reglas que tienden a desalentar la cooperación entre los prisioneros, a saber: limitación de la información, aislamiento recíproco, y un sistema de ganancias que estimula fuertemente la traición recíproca, colocando de ese modo en manos del fiscal la más poderosa herramienta para la acusación. En este marco, el cooperar

(en este caso no delatar al otro prisionero) no es algo moralmente aconsejable, ni tampoco desaconsejable, ya que faltan los elementos imprescindibles para elaborar un juicio moral. Parece razonable que la discusión del marco del juego, del sistema de ganancias, y de las reglas de actuación constituyen una condición básica para saber hacia dónde conduce la selección de una determinada acción. Un prisionero que decide delatar, puede hacerlo porque los fines del fiscal le parecen plausibles, y dirige su cooperación hacia ese destino. En este aspecto parece difícil hablar de cooperación en un sentido normativo en un escenario que no exprese un marco de acuerdo acerca del valor moral de las ganancias asignadas, y de las reglas establecidas. Esta es la razón por la que Gauthier habla de una situación en la que la cooperación sea **beneficiosa y equitativa**, como una garantía de que la cooperación tenga sentido; pero esta afirmación rebasa claramente los límites del juego, y del modelo, y requeriría de una fundamentación no sólo muy amplia, sino también poliargumental.

Según puede verse, ni el relajamiento de las condiciones existentes, ni la incorporación de algunas nuevas permiten razonablemente utilizar el Dilema del Prisionero como demostración de la superioridad (ni en términos de utilidad, ni en términos éticos) de los comportamientos cooperativos. Al interior del juego esta empresa luce como inalcanzable. En todos los casos, el elemento valorativo que puede minimizar los riesgos de la cooperación se origina, y encuentra su justificación, fuera del marco del juego. A la vista de estas circunstancias, la norma de conducta tanto en Axelrod como en Gauthier no prescribe una cooperación incondicional, sino una cooperación condicionada a la actuación del contrincante. Cooperar con los cooperativos, y no hacerlo con los no cooperativos parece ser la regla a la que puede razonablemente apelarse si el juego se repite entre los mismos actores.

La consideración analítica del dilema nos mostró claramente que con los valores que la matriz de ganancias tiene asignados, y las reglas establecidas, el punto de equilibrio es no cooperativo. Esto significa que dentro de ese marco el par no-cooperativo es el más razonable. Esta afirmación produce en el plano normativo una sensación de desaliento; pero sólo aceptándola se avanza en el camino. Hay que admitir que sólo a través de un cambio producido en aquel marco original se puede llegar a una solución cooperativa en equilibrio. El problema entonces consistiría por una parte en saber a cargo de quién debería estar la constitución del nuevo marco que valore la cooperación; y por la

otra, cuáles son las condiciones más apropiadas para alcanzar esa solución.

Acerca de quién puede hacerse cargo de los cambios, un camino sería el de que una entidad superior (por ejemplo un árbitro, o el Estado), interviniera de manera tal que con un sistema apropiado de distribución de las ganancias estimulara, en lugar de desalentar, la salida cooperativa. Esta misma entidad podrá modificar las reglas mejorando la comunicación entre los jugadores, propiciando acuerdos y penalizando las traiciones. Esta solución evidencia claramente dos deficiencias fundamentales: una es que cuanto más en equilibrio se encuentre la salida cooperativa por esta vía, más se alejará la solución del carácter básicamente interactivo del juego. Los *jugadores* dejarán de serlo, ya que sus cursos de acción tendrán un trazado lógicamente inevitable, y en consecuencia sus *selecciones* no serán tampoco tales. Esta deficiencia podría caracterizarse como de incoherencia con los principios del modelo básico. La segunda deficiencia, estrechamente relacionada con la primera, se vincula al carácter potencialmente absolutista que se atribuiría a una sola entidad, la que a pesar de que las reglas fijadas y los pagos otorgados fuesen inobjectables, siempre adolecería de una peligrosa falta de control, situación ésta racionalmente indeseable. Esta deficiencia contradice la base ética de racionalidad del modelo.

La otra solución sería la de que los propios jugadores realizaran un prejuego para establecer, tanto las reglas mismas como los pagos más convenientes que regirán en el escenario del juego. A simple vista, esta solución luce extraordinariamente atractiva: *los jugadores jugando un juego para establecer los pagos y reglas del nuevo juego...* Sin embargo, no estamos frente a una solución improblemática. A primera vista parecería razonable que la estructura de este prejuego fuera la de **suma cero**, ya que si no, se retornaría infinitamente sobre el dilema; pero el hecho de que en la propia definición de un modelo normativo de cooperación se instale un juego suma cero, luce extraordinariamente intranquilizante, y por otra parte, no se exhibe como muy plausible. Sin embargo, si se piensa en este prejuego como regateo acerca de la distribución más apropiada y de las reglas más estimulantes de la conducta cooperativa, es factible el logro de una solución en equilibrio que además satisfaga condiciones plausibles. En un uso normativo, tal solución cooperativa se constituiría en una desiderata a alcanzar a través de sucesivos intercambios. Mediante este procedimiento no se construiría una **norma** en sentido estricto, respaldada por su correspondiente sanción, sino más bien una regla de juego. Es cierto que por esta vía queda sin

resolver la situación de la falta de coercitividad de tales reglas, y en consecuencia el carácter relativamente voluntario de su acatamiento; pero quizás convenga señalar que eso es lo que hace que un juego sea un juego y no un campo de trabajos forzados. Tal vez haya que aceptar que la solución acerca de quién debe cambiar los pagos y las reglas que rigen el juego, tenga inevitablemente que debatirse entre el peligro del autoritarismo, y los riesgos derivados de la falta de coercitividad.

Ahora bien, ya se trate de una decisión tomada por una autoridad superior, o de una regla construida a través del proceso de regateo, el contenido de las restricciones que se aceptan inicialmente establecen límites a los resultados posibles, y también en cierta forma, los determinan. De allí la importancia que reviste el análisis de las mismas. Entre las condiciones generales acerca de las cuales conviene discutir en virtud de su decisiva influencia sobre el resultado del juego, se encuentra la relativa a los niveles de comunicación entre los jugadores, y especialmente a las formas de acceso a la información al momento de trazar sus estrategias; es evidente que un acceso diferencial a la información puede golpear la equidad del escenario, y distorsionar un comportamiento cooperativo. También es necesario detenerse a considerar críticamente la regla que determina quién es el que juega primero, cuál es la duración del juego, y el carácter único o iterado del mismo. Finalmente, las posibilidades de los jugadores de salir o no del juego, y de sacar o no las ganancias del escenario común, constituyen otro centro de interés para el análisis racional. En cuanto a la selección del mejor punto de equilibrio, ya que éste no es único, las restricciones formales que se establezcan, no son en lo absoluto asépticas; ya que bajo determinadas condiciones conducen a resultados que exhiben entre sí importantes diferencias ético-políticas. Los debates acerca del cuarto axioma de Nash para determinar el punto de equilibrio, son una prueba de que una discusión de este tipo es imprescindible cuando se pretende inducir una determinada conducta porque se la considera socialmente deseable<sup>15</sup>.

---

<sup>15</sup> Las soluciones de Nash y Harsanyi, por una parte, y la de Kalai, Smorodinski y Gauthier por la otra, intentan, mediante la selección de axiomas alternativos, una modificación de las reglas para determinar el punto de equilibrio en un proceso de regateo o en un arbitraje. Bajo determinadas circunstancias, unas reglas parecen ser más apropiadas que las otras, lo que implica que la selección de una de ellas conlleva una opción de valor. Por este motivo, aunque bajo cada grupo de axiomas la solución es única, al momento de decidir el empleo de alguna, es necesario evaluar la plausibilidad del sistema de axiomas en el escenario al cual va a ser aplicado.



Como se ha podido observar, cuando se toma en cuenta el uso normativo del dilema, el foco de la discusión se sitúa en los principios mismos del juego, que son considerados como las premisas prescriptivas del juicio de valor: «es deseable que se actúe cooperativamente». Este escenario, así como las herramientas utilizadas, que son las propias de la Filosofía Analítica, configuran una instancia particular de justificación, que se diferencia marcadamente de la correspondiente al uso analítico. Si se quisiera sintetizar esta cuestión, se podría decir respecto del Dilema, que un punto de equilibrio perfecto, considerado en términos analíticos, no corresponde necesariamente al punto más aconsejable en términos normativos. La regla de conducta que se deriva en un enfoque normativo del Dilema del Prisionero deberá estar en equilibrio, o en cuasi-equilibrio; pero además, deberá ser una consecuencia coherentemente obtenida a partir de principios prescriptivos plausibles y suficientemente evaluados en un estado de cosas determinado.

## **6. El lenguaje de la cooperación**

El análisis efectuado en las anteriores secciones nos permite abordar el tema de la cooperación desde una perspectiva más estricta y potencialmente más fecunda. La caracterización de los usos analítico y normativo del Dilema del Prisionero, y de los métodos empleados en los usos considerados, cumple una importante función en el análisis de temas cruciales para la Filosofía Moral y la Ética Normativa. Sólo mediante la aplicación concurrente de ambos métodos, es posible caracterizar tanto la red de propiedades formales que pretenden explicar determinados comportamientos, así como los problemas valorativos fundamentales implicados en la elección de estrategias alternativas, y el análisis del impacto ético que produce la selección de uno u otro marco de referencia. Se considera que ambos usos son metodológicamente complementarios, ya que ofrecen en dos esferas también complementarias, la estructura de justificación necesaria para un acercamiento más amplio a uno de los más importantes temas en el mundo de las relaciones sociales y políticas. Parece claro que ninguno de ambos enfoques metodológicos es por sí solo suficiente: sin el enfoque analítico se corre el riesgo de transitar innecesariamente por caminos oscuros y complicados que desaprovechan todo el potencial explicativo y la plenitud semántica de tales métodos; y sin el enfoque normativo, existe el peligro de olvidar que aunque a veces la selección de un

determinado curso de acción tiene la apariencia de una cuestión estrictamente técnica, siempre implica una definición valorativa que debe ser sometida al juicio de la razón moral.

A partir del análisis de las soluciones intentadas por diferentes autores, tanto en el plano analítico como en el normativo, se hace evidente que los comportamientos cooperativos no son posibles ni razonables, sino dentro de determinados marcos de significados. Por este motivo no tiene caso predicar en abstracto la superioridad moral de la cooperación, ni establecer reglas que prescriban tales conductas de manera general. Como pudo verse, la razonabilidad de utilizar una conducta cooperativa depende del marco dentro del cual se cumplen las acciones, y de las reglas establecidas en el mismo. Esta condición de «razonabilidad» cumple una función en extremo importante, ya que cuando en virtud del entorno ético y de las reglas del juego establecidas no es «razonable» la cooperación, la misma no estará en equilibrio, lo cual significa que no tendrá elementos suficientes para autosostenerse.

Tal vez la mayor importancia, tanto en el plano intelectual como de política práctica, que puede atribuirse al Dilema del Prisionero, radica en su capacidad para mostrar muy claramente las enormes dificultades que entraña lograr el éxito de la cooperación social en un mundo de actores comunicados y que persiguen sus propios intereses individuales, sin apelar a la intervención de un Ente que autoritariamente ordene las conductas. Al propio tiempo el dilema también ha permitido vislumbrar que bajo ciertas condiciones, los comportamientos cooperativos además de ser socialmente beneficiosos, llegan a beneficiar asimismo a los propios actores individuales que los producen. Si esto es así, parece importante en términos normativos articular mecanismos para que tales conductas cooperativas se generalicen. Esta empresa requiere, si pretende obtener un éxito real, que se acepten sinceramente aquellos puntos que han evidenciado ser angulares en el tema, aunque tal aceptación implique asumir dificultades teóricas y políticas de envergadura. Hay una tendencia bastante generalizada, y por cierto sumamente negativa, a pensar que todo elemento analítico que evidencie dificultades para lograr una salida cooperativa debe ser ignorado o distorsionado en homenaje a la pretendida superioridad moral de esa salida. Tales posturas no parecen estar interesadas realmente en un desarrollo de las soluciones cooperativas como hecho analítico y moral, sino en la defensa contra todo evento de un punto de vista, que empieza a parecerse peligrosamente a un mito o a un dogma.

En la estructura del Dilema del Prisionero, que como se dijo refleja de manera muy acertada el conflicto entre los intereses individuales y los comportamientos cooperativos, se ponen también de manifiesto aquellos puntos que se han calificado de angulares, y cuya consideración racional no puede evitarse si en efecto se desea la superación de tal conflicto.

El primero de ellos se refiere al carácter determinante de las reglas establecidas, y de los pagos otorgados. Es evidente que en el propio momento en que el fiscal establece una serie de restricciones, y señala los pagos correspondientes a cada intersección de estrategias, deja prácticamente signado el destino del juego. Cada jugador tiene un cuadro de alternativas posibles, y su capacidad de selección se limita a escoger la estrategia que, en virtud de la racionalidad, maximiza su utilidad esperada. En el caso del Dilema es indiscutible que el fiscal desea, y así se pone en evidencia a partir de las restricciones establecidas, favorecer la no cooperación; y esta tarea la cumple cuidadosamente, ya que una a una todas las condiciones que va elaborando tienden coherentemente hacia esa finalidad.

Sin embargo, si se observa con atención crítica la exitosa tarea que el fiscal ha cumplido, es posible extraer algunos elementos de gran utilidad para la caracterización de nuestro propio problema acerca de cómo promover comportamientos cooperativos estables. El primer elemento, y sin dudas crucial, es el del carácter que el fiscal atribuye a las reglas del juego. Estas reglas que definen la estructura del juego, son sus reglas constitutivas; y en consecuencia quedan fuera de toda posible discusión. El juego mismo, y su carácter no cooperativo, queda decidido en tales reglas, y por esa razón, toda vez que se intenta superar el dilema, se concluye fuera del campo de juego. No son reglas que puedan desarrollarse, sino solamente sustituirse. Este carácter asignado a las reglas tiene el inconveniente de que si las cosas son así, el conflicto entre los intereses individuales y la cooperación estable no tendría solución sino en el marco del autoritarismo.

En efecto, en virtud de estas reglas que constituyen el juego, las cuales incluyen tanto las restricciones como el sistema de pagos, es que el par en equilibrio del Dilema del Prisionero es no cooperativo. Este es otro de los puntos considerados angulares en el tratamiento del tema. Cuando se efectúa tal afirmación, se producen reacciones interesantes. Hay quienes se resisten a aceptar la situación, y buscan soluciones analíticas que demuestren la posibilidad de superar la afirmación; hasta el momento, según pudo comprobarse ampliamente, tales intentos no han logrado

éxito. Otro grupo de reacciones se orienta en el sentido de atacar desde un punto de vista ético la validez de la afirmación. Este enfoque equivoca su camino, ya que la afirmación se refiere a una situación que o bien es analítica, o bien es de hecho (en el plano de la práctica política); en consecuencia, su plano de justificación y de crítica debe corresponder al *status* analítico o empírico en cada caso, y nunca ubicarse en el plano valorativo. Bueno es señalar que quienes pretenden negar que la solución estable del dilema es no cooperativa, no son siempre quienes de manera más auténtica desean estimular las conductas cooperativas; por el contrario, su actitud, que en gran medida dificulta la búsqueda de soluciones, podría ser un indicador del deseo de no alcanzarlas. A este respecto puede hacerse una afirmación enfática: no es posible buscar soluciones al conflicto entre interés individual y cooperación estable, si no se acepta que tanto analíticamente como en términos de práctica política, el equilibrio es no-cooperativo. Esta afirmación es una prueba evidente de lo bien que ha trabajado el fiscal en el establecimiento de las reglas de este juego.

Estas reglas que el fiscal ha establecido tienen, tal como se dijo, el carácter de reglas constitutivas del juego mismo, y su contenido determina que la única solución estable sea no cooperativa. Si lo que se desea es estimular las conductas cooperativas, dotándolas de estabilidad, será de mucha utilidad preguntarse acerca de cuáles son esos contenidos que tan fuertemente sostienen el equilibrio no-cooperativo; porque una vez determinados los mismos será factible pensar en el sentido de los cambios que pueden efectuarse.

Resulta válido decir que dentro de los contenidos de las reglas del dilema, el referido a la relación de orden en los pagos asignados es el responsable de que en el plano analítico se presente el conflicto entre el interés individual y la cooperación estable. Esta regla de distribución de los pagos aprisiona a los jugadores en una disyuntiva que los lleva al equilibrio no-cooperativo. Si se recuerda, la relación que guardan los pagos establecidos, puede expresarse como  $T > C > NC > E$ , donde T equivale a traicionar, C equivale a cooperar, NC a no cooperar, y E equivale a sufrir explotación como consecuencia de la traición del contrincante; esto último lleva a que la relación traición-explotación sea inseparable. La ubicación de la C en el segundo lugar en el orden establecido, y de la NC en el tercero, puede llevar a pensar que la selección cooperativa ha sido privilegiada; pero en realidad, el tema de la estructura de los pagos no es tan simple. La selección de la intersección C, lleva implícito correr el

riesgo, ciertamente alto, de ser víctima de una explotación en virtud de la traición del contrincante. Ese riesgo se ha calificado de alto en razón de que al colocarse a la traición en el primer lugar del ordenamiento se la torna muy atractiva. Por otra parte, escoger la NC es también el único modo seguro de no ser explotado, que es la peor situación de pagos que pueda presentarse.

Como se observa, no sólo la relación de orden entre los pagos ofrecidos, sino la estructura de los mismos, tienden a definir un conflicto profundo entre la prosecución de los intereses individuales y el logro de una cooperación estable. En otras palabras, el sistema de pagos establecido por las reglas del fiscal es desequilibrador de la cooperación, y altamente reforzante del equilibrio no cooperativo. Aunque muchos no deseen admitirlo, bajo tales condiciones es perfectamente racional no cooperar, y muy poco plausible pensar en la inducción de conductas cooperativas estables, por muy deseables que las mismas sean desde el punto de vista de la ética social.

Así como la estructura de la matriz de pagos es la responsable del conflicto entre los intereses individuales y la cooperación estable, la regla relativa a la prohibición de efectuar acuerdos, y de contraer obligaciones previos al juego es la causante de que aquel conflicto no pueda resolverse en el marco del Dilema del Prisionero. En efecto, si los prisioneros pudieran establecer contacto previo a la selección de las estrategias, estarían en condiciones de analizar la relación entre los pagos, y seguramente acordarían, tomando algunas precauciones para evitar incumplimientos, concurrir a la casilla cooperativa, que les ofrece a ambos la expectativa de una alta ganancia. Para evitar la posibilidad de que los prisioneros contraigan obligaciones previas al juego, el fiscal los mantiene aislados. Este aislamiento abarca varios aspectos: primero, los prisioneros no pueden comunicarse directamente; en segundo lugar, tampoco pueden hacerlo a través de terceros, y, por último, no pueden comunicarse por medio de un universo normativo compartido, ya que esto implicaría la existencia de un acuerdo preexistente.

Eliminadas estas posibles formas de comunicación, el conflicto entre la búsqueda del interés individual y el mantenimiento de un equilibrio cooperativo se vuelve insoluble, ya que las reglas iniciales, y la estructura de pagos establecidas por el fiscal, no pueden ser corregidas o desarrolladas. La incomunicación, entonces, actúa reforzando el carácter constitutivo y determinante de las reglas del fiscal; ya que elimina cualquier instancia reguladora que no se cumpla en el terreno mismo del juego, con lo que

todo intento por superar el equilibrio no-cooperativo está condenado al fracaso.

Considerando el planteamiento del dilema en su versión original, puede verse entonces, que a causa de la estructura de distribución de las ganancias, la solución en equilibrio es no cooperativa; y que en razón de la incomunicación entre las partes, tal solución es fuertemente estable. Por otra parte, según se vio, la estructura del dilema tal como está planteada, refleja de manera certera la profunda dificultad que existe para armonizar los intereses individuales con un comportamiento cooperativo estable. Pero a pesar de que tengamos que reconocer este hecho, ¿el mismo resulta verdaderamente deseable? En otras palabras, ¿tiene algo de bueno la cooperación como para que merezca tanta atención analítica y política su estímulo?

Hay un hecho a este respecto que resulta irrefutable: cuando se produce el enfoque cooperativo de un juego, el espacio de soluciones factibles se amplía de manera considerable respecto de un enfoque no cooperativo; y tales ampliaciones del espacio se producen de una manera que no es siquiera pensable en el marco de un enfoque no cooperativo. Asimismo, como pudo observarse en la matriz del dilema, se obtienen mayores ganancias individuales que cuando no se coopera, y la suma de las ganancias colectivas es la máxima. Desde un punto de vista analítico estas ventajas parecen suficientes para preferir el enfoque cooperativo sobre el no cooperativo; siempre y cuando se pudieran suavizar o eliminar las restricciones que tornan altamente riesgosa, y en consecuencia inestable, la intersección cooperativa.

A pesar de que en el plano analítico es posible comprender perfectamente el sentido de la ampliación del espacio que se produce cuando se utiliza el enfoque cooperativo, no existe en este aspecto un ejemplo más claro que lo sucedido en los procesos de unificación europea cumplidos en lo que va de la segunda mitad de nuestro siglo.

Los deseos de producir una unificación de Europa no son nuevos. Ya en 1923, el austriaco Countenhove Kalergi, aspiraba a producir desde la Unión Paneuropea, que él mismo ayudara a fundar, la creación de los Estados Unidos de Europa. Sus modelos políticos eran el de la unificación lograda en Suiza en 1648, y la de los Estados Unidos en 1776. En el célebre discurso pronunciado en Ginebra ante la Sociedad de Naciones el 5 de septiembre de 1929, por el ministro de Asuntos Exteriores de Francia Aristide Briand, se propuso a los gobiernos la creación de la Unión Europea en el marco de la propia Sociedad de las Naciones. Si se analizan los discursos políticos de este momento, se

percibe que las buenas intenciones estaban sostenidas por conceptos que apelaban fundamentalmente al espíritu de colaboración, altruismo y humanitarismo. El hecho de que no fuera siquiera planteada la necesidad de autolimitaciones en los derechos e intereses de los Estados, coloca el intento en un terreno más propio del romanticismo moral que en el de la realidad política. A este respecto hay que señalar que en los procesos de tomas de decisión interactiva, no es posible olvidar que siempre nos encontramos en presencia de formas diversas de conflictos de intereses, que para ser resueltos cooperativamente exigen la limitación racional del principio de maximización individual. Si este punto no es considerado, y resuelto apropiadamente, no se puede pensar en una cooperación estable. Los discursos morales no fueron suficientes para evitar el drama de la guerra, en la que el espíritu de los juegos suma cero fue llevado a su expresión más aterradora.

En el confuso escenario de la postguerra surgen una serie de organizaciones que, alentadas por el ideal expresado en 1929, pero a la vez con mayor madurez en el enfoque metodológico, trabajan a favor del logro de una salida cooperativa estable. El tema de los intereses divergentes, y el de la necesidad de limitar la maximización individual de cada estado, esta vez no son eludidos. Tales limitaciones hacen posible logros del tamaño moral del Convenio Europeo de Protección de los Derechos Humanos y las Libertades Fundamentales, del 4 de noviembre de 1950, que a su vez da origen a la Comisión Europea de Derechos Humanos y al Tribunal de Derechos Humanos. Mediante estos organismos se establece una garantía mínima y claramente estable, que ha supuesto importantes limitaciones a la maximización de las ganancias individuales que cada Estado lograría con el ejercicio irrestricto de su soberanía. El escenario se ha ampliado claramente a causa del enfoque cooperativo, y muy pronto se evidenciaron las nuevas ganancias individuales (para cada Estado) y colectivas, que en razón de tal ampliación se produjeron.

Pero, según se señaló, la ampliación del espacio de soluciones posibles a partir del enfoque cooperativo no se limita al previsible escenario del juego mismo, sino que a su vez, posibilita nuevas ampliaciones. En el ejemplo que está en consideración tal ampliación se produce cuando, sobre la base del plan Schuman que proponía integrar la totalidad de la producción franco-alemana de carbón y acero bajo una Alta Autoridad Común, se firma el 18 de abril de 1951 el Tratado de la Comunidad Europea del Carbón y el Acero (CECA), que entra en vigor el 23 de julio de 1952. Aquí las limitaciones del principio de maximización

individual son de tal importancia que suponen la renuncia irreversible a muy importantes parcelas de la soberanía nacional, en la búsqueda de soluciones cooperativas estables y equitativas. Dichas limitaciones se irán incrementando con el tiempo, y produciendo nuevas y nuevas ampliaciones del espacio factible.

No es en absoluto exagerado considerar a esta nueva ampliación del espacio de soluciones como de orden superior, ya que da lugar a un nuevo concepto que aunque nacido de aquel enfoque cooperativo de primer orden, lo supera. Este es el concepto de integración, en el que se torna visible la estabilidad de la solución cooperativa, estabilidad derivada del hecho de que tanto las ganancias individuales como las colectivas se han incrementado.

De lo expuesto hasta este momento se infiere que tanto desde el punto de vista analítico, como del de la práctica política, el espacio de las soluciones factibles se amplía con el enfoque cooperativo; pero también se infiere que en ambas esferas, las dificultades para lograr soluciones cooperativas estables es grande. Indudablemente es conveniente cooperar, pero parece necesario determinar cuáles son las condiciones que pueden favorecer esas conductas en un marco de respeto de los intereses individuales y los colectivos.

Cuando decíamos que el fiscal hizo muy bien su trabajo para favorecer los comportamientos no cooperativos, señalábamos que simultáneamente ha puesto en evidencia las condiciones fundamentales que hacen que cooperar sea difícil y riesgoso. Parece sencillo derivar que si se desea favorecer o estimular las soluciones cooperativas, habrá que revertir aquellas condiciones establecidas por el fiscal, e intentar que el trabajo de reversión sea tan bueno como el original.

El primer aspecto a considerar es el relativo al carácter de las reglas del dilema. Según se puso ya en evidencia, tanto de las consideraciones analíticas como de las normativas, surge el carácter determinante de las reglas que enmarcan el juego, y de las premisas prescriptivas que bajo ciertas circunstancias hacen una conducta más recomendable que otras. Asimismo se ha visto que los procesos de regateo en un prejuego entre los jugadores, o el arbitraje por parte de una autoridad, parecen ser los mecanismos que pueden transformar en un momento dado las reglas del juego, modificar los pagos, y redefinir las premisas prescriptivas.

Si bajo la nueva óptica que ha surgido de los análisis efectuados en este trabajo, se reconsidera el tema inicial de la cooperación, puede observarse claramente una desmistificación del



asunto. Ya la cooperación no se vincula a comportamientos altruistas valiosos en sí mismos, sino a la búsqueda de una utilidad conjunta que se legitima sólo bajo determinadas condiciones, que tienen mucho que ver con complejos procesos de producción y distribución de los beneficios. La cooperación, entonces, es solamente una salida; aconsejable o no, posible o no, conforme a las reglas que sea factible establecer en determinados momentos. Las reglas se constituyen entonces, en el punto crucial a considerar en el ámbito de la Filosofía Moral, el de la Ética Normativa, y el de la Política Práctica.

A fin de determinar la índole de las reglas del dilema se seguirá el camino de caracterizar los diferentes tipos de reglas, para tratar de establecer analogías con alguno de ellos. Hay un primer tipo que son las llamadas reglas de la naturaleza, que tienen por característica fundamental la de ser descriptivas de comportamientos naturales, sobre cuya ocurrencia no exhiben facultades normativas. Las descripciones que estas reglas efectúan pueden producir una interpretación de tales hechos naturales (teorías), pero de ningún modo pueden alterar el acaecer de los mismos. Resulta fácil trazar la diferencia entre las reglas de la naturaleza y las que nos ocupan, ya que éstas en ningún momento son descriptivas del comportamiento de los jugadores y, desde luego, ejercen una notable influencia sobre el mismo.

Otro tipo de reglas son las llamadas reglas del Derecho. Estas reglas no describen comportamientos, sino que prescriben determinadas conductas, estableciéndolas como obligadas, prohibidas o permitidas. Este carácter prescriptivo de las reglas del Derecho hace que las mismas ejerzan una considerable influencia sobre los comportamientos, especialmente porque la regla de Derecho se encuentra asociada a la idea de alguna forma de sanción, la cual refuerza su poder modelador de conductas. A pesar de que la diferencia de este tipo de reglas con las del dilema no es tan evidente como en el caso anterior, sí puede ser claramente percibida, ya que en el caso de las reglas del dilema, no encontramos estrictamente hablando una prescripción de conductas, aunque de un cierto modo influyan sobre las mismas; y por cierto que no se vinculan a una idea de sanción.

Un tercer tipo de reglas son las llamadas reglas de un juego (como el ajedrez o el tenis), que no describen ni prescriben conductas, sino que de una cierta manera, las determinan. Las reglas del ajedrez, por ejemplo, determinan los movimientos permitidos y aquellos que no lo están, con lo que también determinan qué significa jugar correctamente, o incorrectamente,

o no jugar al ajedrez<sup>16</sup>. La diferencia entre estas reglas y las del dilema no es ya tan clara. En un cierto sentido, en el dilema las reglas determinan las alternativas permitidas pero de un modo diferente al del ajedrez; pues mientras las reglas del ajedrez determinan de una manera fuerte, ya que no sólo regulan la actividad de jugar, sino que la constituyen, las reglas del dilema sólo regulan de una manera débil un sistema de relaciones. De este modo la expresión «seguir la regla» tiene distinto significado en uno y otro caso. La diferencia se podrá percibir mejor si se consideran las reglas del lenguaje.

Como una forma especial de las llamadas reglas de juego, se encuentran las reglas del lenguaje<sup>17</sup>. Entendido como un juego, el lenguaje posee reglas con características especiales que se diferencian de las reglas de un juego como el ajedrez en varios aspectos: en primer lugar, se trata de reglas en la mente y no escritas; en segundo lugar, son reglas de significado y no justificatorias; y en tercer lugar, son reglas que pueden producirse a medida que se juega. Si consideramos estos rasgos del juego del lenguaje, encontramos mucha similitud con las reglas que nos ocupan.

Cuando dos actores sociales se enfrentan en el prejuego para ir fijando las reglas del juego y determinar la matriz de pagos, o cuando lo hace un árbitro después de una ronda de negociaciones, a la vez que están estableciendo reglas para el juego, están jugando. También en el ejercicio del propio juego se descubren los significados de determinadas acciones, y el alcance de ciertas restricciones. De esta manera, lo mismo que en el juego del lenguaje, a medida que se siguen las reglas se las va fijando; este hecho torna perfectamente clara la diferencia de estas reglas con las de un juego como el ajedrez. En las primeras, el jugador es, como en el juego del lenguaje, un productor de reglas; mientras que en el segundo, el jugador sólo puede sujetarse a las reglas que han constituido el juego.

Precisamente en el Dilema, el carácter que el fiscal ha asignado a las reglas es el propio de juegos como el del ajedrez, y esto limita las posibilidades de los jugadores para efectuar una decisión

---

<sup>16</sup> La caracterización de las reglas que incluyen a las del ajedrez como reglas constitutivas, es decir, que definen el ámbito del juego y también lo que está permitido o prohibido dentro del mismo, se debe a G. E. v. Wright.

<sup>17</sup> La concepción del lenguaje como un juego, y la descripción de los caracteres peculiares de las reglas del mismo se deben a Wittgenstein, quien proporciona claves muy importantes para comprender el complicado proceso mediante el cual aprendemos el significado de «hacer lo mismo».

que sea moralmente juzgable. Este hecho, y el carácter constitutivo que se asocia a tales reglas, no resulta plausible en un escenario de relaciones sociales y políticas, en el que los actores deben asumir responsabilidades morales, y en consecuencia tienen el derecho de procurar la transformación de los marcos del juego hacia variantes más equitativas y menos traumáticas. Si verdaderamente se desea alcanzar la armonización de los intereses individuales con el equilibrio cooperativo, las reglas deben ser consideradas como no-constitutivas; es decir, reglas en elaboración. Esto supone borrar los límites tan rígidamente trazados por el fiscal entre el juego y el prejuego. En ambas instancias, los jugadores están jugando: exhiben su lenguaje a través de los procesos de regateo o a través de la selección de estrategias; y de esa manera van configurando el marco de significados del lenguaje. Este es el único modo en que se puede establecer un diálogo que haga posible y menos riesgosa la cooperación. A este respecto, si se consideran las posiciones más decididas a aconsejar en términos normativos las conductas cooperativas (tales como las de Axelrod y la de Gauthier), puede verse que ninguna se inclina por una cooperación irrestricta, sino condicional a que la otra parte comparta ese lenguaje<sup>18</sup>; es decir, que consideran a la cooperación como un diálogo; y sólo de este modo puede llegar a ser «racional» en el sentido del juego. El único error de estos autores consiste en que pronunciándose por una defensa tan férrea de la cooperación en el propio marco del dilema no han prestado atención a la forma en que ese marco se construye; y no han podido ver que si se acepta que las reglas del fiscal como autoridad son constitutivas, sólo la acción de una nueva autoridad podrá modificarlas. Para que un jugador tenga la facultad, y la posibilidad de incidir en esas reglas, no puede ser un jugador de ajedrez o de tenis, sino alguien que está hablando un lenguaje.

Como puede observarse, la consideración de las reglas de la cooperación como reglas de lenguaje obliga a plantear una cierta disolución de los límites entre el juego y el prejuego, lo que nos lleva al tema de la comunicación; la cual, por las reglas del fiscal se encuentra excluida, bajo cualquier forma que pudiera asumir. El especial cuidado puesto por el fiscal en evitar formas de comunicación (directa, indirecta, etc.) evidencia la importancia

---

<sup>18</sup> Tanto la estrategia TIT FOR TAT de Axelrod, como la de maximización limitada de Gauthier, condicionan el uso de la cooperación a la conducta del oponente, ya que admiten que una cooperación incondicional produciría la explotación y el ulterior aniquilamiento de quien la practique.

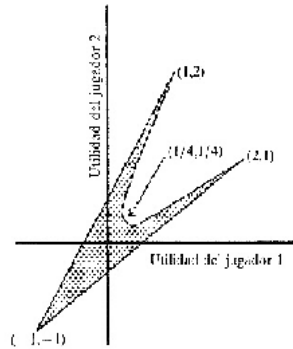
de la misma como estímulo de las conductas cooperativas. A pesar de que tal como lo sostiene Harsanyi la sola comunicación no puede garantizar el equilibrio cooperativo, sin embargo ella sí amplía el espacio de soluciones posibles y, naturalmente, baja las probabilidades de que se produzcan situaciones de traición-explotación. La comunicación no puede producir sus efectos si el dilema es jugado en una sola oportunidad; pero según ya se señaló, en el ámbito de las relaciones de política práctica, la situación más frecuente es la de interacciones reiteradas entre los mismos actores. Tal reiteración abre el camino a diferentes formas de comunicación, que hacen posible la ampliación del espacio de soluciones, al plantearse la visión conjunta del problema.

Hay un juego no-suma-cero-no-cooperativo que puede exhibir con la máxima claridad el poder de la comunicación como ampliadora del espacio de soluciones posibles. Su interpretación más usual es conocida como la Batalla de los Sexos: el jugador número 1 es el hombre, y el número 2 es la mujer. El hombre tiene preferencia por los espectáculos de lucha, y la mujer por el ballet; pero además, ambos otorgan importancia a disfrutar juntos de sus diversiones. La matriz de utilidades puede expresarse como sigue:

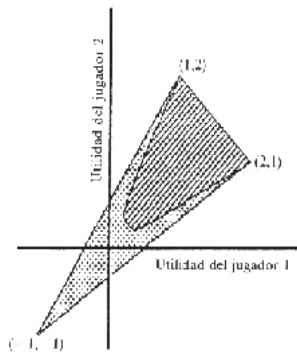
		Jugador 2	
		L	B
Jugador 1	L	2,1	-1, -1
	B	-1, -1	1,2

Naturalmente, el hombre estará muy feliz si ambos van a presenciar la lucha (2,1); y la mujer preferirá por encima de cualquier otra alternativa la de que ambos vayan al ballet (1,2). Si se supone que el juego se realice una sola vez y sin posibilidades de comunicación prejuego, la gráfica correspondiente al espacio de soluciones posibles es el siguiente:

Pero si el juego se repite, y la comunicación es permitida, los jugadores pueden,



mediante acuerdos que respeten los intereses de ambos, producir una ampliación del espacio de soluciones factibles. Bastará con que utilicen un método de aleatorización de las estrategias conjuntas (por ejemplo, arrojar una moneda insesgada para decidir cada vez a donde irán), y su espacio factible de utilidades se ampliará hasta los límites  $(3/2, 3/2)$  señalados en la siguiente figura:



Intuitivamente hablando, la comunicación ha permitido tender un puente entre las dos utilidades máximas individualmente esperadas, al producir la convexificación del espacio factible. Tal es la función que cumple la comunicación y el prejuego como modificadores de los sistemas de pagos, lo que implica modificar el universo simbólico que los jugadores comparten.

Si se observan los grupos humanos y políticos que han obtenido los mejores resultados en la articulación de los intereses individuales y la cooperación estable, se percibe que tales grupos han logrado, por lo general a través de una práctica colectiva, superar el escollo fundamental que conspira contra el desarrollo de todo lenguaje. Para que un lenguaje exprese al grupo que lo usa, hay que asumir el carácter especial de sus reglas, que se producen a medida que se usan. Esto supone, en primer lugar, no considerar a las reglas como inamovibles; si las reglas han sido construidas es posible, desde luego, modificarlas, y esa modificación supone el ensayo de nuevas reglas que permitan expresar mejor el juego que los jugadores desean jugar. En el plano analítico, muchos autores han trabajado en intentos por lograr reglas que definan puntos de equilibrio diferentes. Los trabajos de Nash, Harsanyi, Kalai y Smorodinski y recientemente Gauthier, constituyen búsquedas de reglas para resolver los procesos de regateo; y al mismo tiempo constituyen puntos de vista diversos de lo que puede considerarse «la mejor solución». Sin embargo, conviene recordar que el juego del lenguaje no consiste en su gramática, y que en tanto las soluciones no tomen la forma práctica del regateo efectivo por modificar pagos y condiciones, no es posible hablar de un uso y construcción de nuevas reglas. Los grupos exitosos en la obtención de acuerdos respecto de las reglas no son, por lo general, los que exhiben una mejor gramática (normas de sintaxis), sino los que a través de una práctica social intensa van produciendo reglas que mejor los expresan.

El otro rasgo que exhiben los grupos que logran mayor éxito en el establecimiento de un lenguaje de cooperación, es su disposición a considerar el carácter provisional de cada regla (solución), ya que admiten que el uso y la construcción de reglas de este tipo prácticamente se confunden en un mismo proceso. Esta disposición se expresa en el orden práctico como una propensión al regateo permanente en procura de ampliar los espacios de soluciones posibles. Al usar el regateo como modo de cooperación se van construyendo reglas y sistemas de pagos que favorecen el uso de un lenguaje común.

Si se acepta que la solución al Dilema del Prisionero debe ser un punto de equilibrio, y tal como está descrita la matriz de

ganancias el punto de equilibrio es no cooperativo; y si además a partir del análisis de las premisas valorativas se acepta la superioridad de la salida cooperativa bajo ciertas circunstancias, el procedimiento para lograrla consistirá en ir ganando terreno para que los pagos se modifiquen en el sentido definido como deseable. Tal modificación no puede ser planteada como drástica y absoluta, ya que eso supondría un viraje normativo difícil de explicar y muy difícil de lograr sin quiebras dramáticas en la trama social. La solución entonces parece ser la propia del jugador: calcular el riesgo, pero arriesgarse dentro del marco razonable de un planteamiento utilitario que le asigna valor al logro de un mejoramiento en las reglas o el sistema de pagos. Que la salida cooperativa sea estable depende de las reglas y de los pagos; en consecuencia, desarrollar un lenguaje de cooperación consiste en ir construyendo, mientras se juega, un sistema de reglas y de pagos que favorezcan y no desestime la cooperación.

El fiscal hizo muy bien las cosas a favor de la no-cooperación; pero al hacerlas tan bien ha dejado muy claros indicios de cómo dismantelar el dilema: si los jugadores (son al menos dos) aceptan el destino de jugadores de ajedrez o de tenis que el fiscal les ha asignado, sólo es posible esperar que otro fiscal (éste, procooperación) cambie las reglas y los pagos; pero si los jugadores consideran que para que este juego sea válido en el campo político y social las elecciones deben ser éticamente responsables, lo que incluye la posibilidad de modificar las reglas, entonces el prejuego será parte integrante del juego, ambos como instancias de producción de las reglas del lenguaje.

Para llevar adelante esta tarea se requiere aceptar un principio: el del carácter no constitutivo de las reglas del juego, lo que abre la posibilidad de modificarlas mientras se juega. Por otra parte, como método para producir tales modificaciones, el permitir la comunicación no sólo a través del juego, sino en el prejuego, ha probado analítica y prácticamente tener una excelente potencia de expansión del espacio de soluciones. La aceptación del principio, y el empleo del método, no parecen ser sinsentidos en muchos escenarios sociales y políticos de nuestro tiempo.

**BIBLIOGRAFÍA**

- Axelrod, R., *The Evolution of Cooperation*, Basic Books, New York, 1984.
- Barragán, J., *Hipótesis metodológicas*, Editorial Jurídica Venezolana, Caracas, 1983.
- , «Los juicios de valor en la planificación de políticas», *Planificación y Política*, núm. 3, año 1985.
- Flood, M. M., «Some Experimental Games», *Research Memorandum R.M. 789*, The RAND Co. Santa Mónica, 1952.
- Garzón Valdés, E., *El concepto de estabilidad en los sistemas políticos*, Centro de Estudios Constitucionales, Madrid, 1987.
- Gardner, M., «Mathematical Games», *Scientific American*, July, 1973.
- Gauthier, D., *Morals by Agreement*, Oxford University Press, New York, 1986.
- Harsanyi, J., «Communications», *American Political Science Review*, núm. 68, Dec. 1974.
- , *Essays on Ethics, Social Behaviour and Scientific Explanation*, D. Reidel Publishing Co. Dordrecht-Holland/Boston, USA, 1976.
- , *Reseña de Morals by Agreement by D. Gauthier*, University of California, Berkeley, Mimeo, 1988.
- Hoerster, N., *Problemas de ética normativa*, Ed. Alfa, Buenos Aires, 1975, traducción del alemán de E. Garzón Valdés.
- Howard N., *Paradoxes of Rationality: Theory of Metagames and Political Behaviour*, MIT Press, Cambridge, Mass., 1971.
- Kalai, E.; Somorodinski, M., «Other Solutions to Nash's Bargaining Problem», *Econometrica*, vol. 43, núm. 3, May. 1975.
- Levi, I., «Newcomb's Many Problems», *Theory and Decision*, núm. 6, May. 1975.
- Luce, D.; Raiffa, H., *Games and Decisions*, John Willey & Sons, New York, 1957.
- Nash J., «The Bargaining Problem», *Econometrica*, 1950.
- Nozick, R., «Newcomb's Problem and Two Principles of Choice», en *Essays in Honor of Carl G. Hempel*, ed. Nicholas Rescher (Dordrecht, Netherlands: D. Riedel Publishing Co., 1969.
- , «Reflections on Newcomb's Problem», *Mathematical Games Department*, Scientific American, March, 1974.
- Schlesinger, G., «The Impredictability of Free Choices», *British Journal for the Philosophy of Science*, núm. 25, September, 1974.
- V. Wright, G. H., *Norm and Action*, London, 1963.
- Wittgenstein, L., *Philosophical Grammar*, Blackwell Ed. London, 1974, traducción del alemán de A. Kenny.

**DOXA -6 (1989)**