

Pedro Francés Gómez*

Sobre si nos conviene ser moralmente buenos

RESUMEN

El contractualismo moral ha sido acusado de justificar racionalmente la moralidad al precio de identificarla con la «maximización a largo plazo». Pero si la moralidad equivale a la maximización a largo plazo, la ética carece de interés y debería ser sustituida por una teoría normativa de la racionalidad maximizadora. El presente artículo trata de refutar esta perspectiva ofreciendo una re-exposición del argumento contractualista moral y una explicación del concepto de agente moral que éste implica.

Palabras clave: INTERÉS, MORALIDAD, CONTRACTUALISMO, RACIONALIDAD.

ABSTRACT

Moral contractualism has been charged with justifying morality by identifying it with «long term maximization». If this were the case, ethics would lack any interest, and it should be replaced by a normative theory of maximizing rationality. This paper tries to refute such view, giving a restatement of the moral-contractualist argument and an explanation of the concept of moral agent involved.

Keywords: INTEREST, MORALITY, CONTRACTUALISM, RATIONALITY.

* Facultad de Filosofía. Universidad Complutense de Madrid.

1. Interés individual y moralidad

Una de las interpretaciones de la obra de David Gauthier *La moral por acuerdo*¹ que juzgo más exacta, la he escuchado de labios del profesor Ezra Heymann. Esa interpretación dice, simplemente, que la minuciosa y compleja obra del filósofo canadiense tiene como objeto principal mostrar esta gran cosa: que un agente moral no es completamente estúpido. A los espíritus bienpensantes quizá les parezca muy evidente que la «gente buena» no sea necesariamente imbécil. Mas los espíritus bienpensantes no tienen idea de lo complicado que resulta demostrarlo. De hecho, conforme a los criterios de la racionalidad instrumental generalmente supuestos en las ciencias sociales —y extendidos desde los años sesenta a ciencias normativas tales como la economía y la filosofía políticas— el agente moral, entendido no tanto como aquella persona simpática que *desea* hacer el bien al prójimo (el altruista puro), sino como aquella persona rigurosa, convencida de que hay límites absolutos en su *forma* de actuar impuestos por normas universalmente válidas, y que se obliga a sí misma a respetar esos límites en virtud de tal convicción racional; ese tipo de agente es, conforme a aquellos criterios, en efecto irracional.

Refutar esta afirmación sin renunciar al punto de partida sólido que ofrece la racionalidad instrumental² requiere básicamente demostrar que el agente moral actúa, en el fondo, *de un modo instrumentalmente* correcto: logra avanzar sus intereses (cualesquiera que estos sean) todo lo posible, dadas las condiciones que le imponen el mundo y los demás. El problema es que una justificación de la moralidad sobre estas líneas no parece una justificación de la moralidad en absoluto. Mostrar que todo el aplauso que merece el agente moral es debido a la gran astucia que demuestra al elegir la moralidad como medio para el logro de sus propósitos, no parece, en efecto, un buen modo de argumentar en favor del *verdadero* punto de vista moral, sino en favor de una idea convencional e instrumental de la moralidad, que la niega por completo. Según

¹ Barcelona, Gedisa, 1994; para las referencias, emplearé la edición inglesa, *Morals by Agreement*, Oxford, Blackwell, 1986.

² Esta refutación es, desde un punto de vista, más sencilla si se opta por un punto de partida diferente, argumentando, como hace Habermas, que hay otro tipo de racionalidad completamente ajeno a la instrumental (la racionalidad comunicativa) de donde se deriva la validez de los principios morales.

este argumento —copia del desplegado por Glaucón y Adimanto en el libro II de *La República*, referido a la justicia³—, el único valor de la moralidad se derivaría de las buenas consecuencias que trae para quien la practica. Los inmorales son obcecados maximizadores a corto plazo. Mientras, el agente moral parece perder ocasiones inmediatas de beneficiarse, pero sólo porque está invirtiendo inteligentemente para obtener una recompensa mayor en el futuro. Parece estúpido, pero no lo es. A largo plazo quedará claro que su estrategia supera a la de quienes desprecian las normas para atrapar ganancias inmediatas.

Pero, tal como Glaucón y Adimanto reclamaron a Sócrates, no puede ser esto lo que estamos buscando al demandar una demostración de que es racional ser moral. Una justificación simplemente utilitaria de las normas morales deja al agente moral, si bien se mira, *fuera* de la moralidad. El agente moral no sería más que un deliberador instrumental muy bien disfrazado; llevando al más alto grado de sofisticación su estrategia. Tal deliberador no tendría ningún interés *intrínseco* por la moralidad. Si se demostrase que hay un medio mejor para lograr sus fines, estaría obligado, como agente racional, a emplearlo (so pena de cometer, ahora sí, una necesidad). Sin embargo, esto es precisamente lo que creemos que distingue a la moralidad de otros cuerpos normativos: que su cumplimiento no está (o no debe estar) subordinado a ninguna consideración utilitaria. Ser moral no es sólo *cumplir* con las normas morales por cualquier motivo; sino cumplirlas *porque* son normas morales. Ser moral implica tener un interés *intrínseco* en la moralidad; un interés que no puede depender de las eventuales consecuencias del acto o modo de acción que la norma moral imponga.

Ante este resultado, una retirada prudente parece aconsejable, porque si el interés en la moralidad no puede depender de las consecuencias de su aceptación, ¿de qué podría depender? En un mundo laico, individualista, racionalizado y economizado (o monetarizado), las cosas tienen valor (y despiertan nuestro interés) sólo en la medida en que son buenos medios para nuestros fines individuales —es decir, en la medida en que juzgamos que sus consecuencias son buenas para nosotros o para quienes nos importan personalmente. Esta regla es poco menos que indiscutible en el ámbito público, y sólo

³ Cfr. Platón, *La República*, 358e-362e.

algunas personas la cancelan en el ámbito privado, donde conceden un valor absoluto a ciertos objetos del mundo en virtud de sus creencias metafísicas o religiosas. Pero en una sociedad liberal moderna, estas creencias no deben salir del ámbito privado, mientras que la moralidad —que es, por definición, universal— debería funcionar también en el ámbito público. Parece, por tanto, que estamos avocados a que un código utilitarista público reemplace a la moralidad. Sus normas, convencionales, dependerían de: *a*) nuestra *percepción* de cuáles son nuestros intereses (individuales y colectivos) en cada momento y *b*) nuestra capacidad de cálculo sobre sus consecuencias a largo plazo para el avance de esos intereses. Quizá podríamos llamar «moral» a esa cosa, pero en realidad no tendría nada que ver con la moral tal como la entendemos. Nada tendría que decir aquí el filósofo moral, que cedería su lugar al psicólogo, al político, al estadístico, al economista y al matemático. La prudente retirada de la moralidad al ámbito privado no es otra cosa que una renuncia total a la moral propiamente dicha⁴.

Quizá ahora se entienda mejor el empeño que el contractualismo mantiene en mostrar que es posible justificar la moralidad en términos de racionalidad instrumental *sin* reducirla a una mera «buena estrategia» maximizadora a largo plazo. Por un lado, la no reducción de la moralidad a la maximización a largo plazo es necesaria si queremos seguir hablando de moral en absoluto; por otro, un argumento en términos de racionalidad instrumental es el único capaz de mantener la esperanza de una justificación que pueda presentarse como universalmente válida, es decir, como una restauración no decisionista de la moral.

El problema con este intento es que, *en apariencia* (debido a la forma del argumento moral contractualista), el tipo de justificación ofrecida consiste en defender que la moralidad es una buena estrategia maximizadora a largo plazo. Y claro, si esta apariencia fuera realidad, el proyecto contractualista habría sido completamente fútil. Mi opinión personal es que el argumento contractualista

⁴ En este sentido, me parece vigente la opinión de Kant de que la reducción de la razón práctica a los fines empíricos (*eudaimonismo*) y a la decisión consecuencialista sobre los medios para alcanzarlos (mediante la formulación de imperativos hipotéticos) supondría la muerte callada (*euthanasia*) de toda moral. Cfr. el prefacio de la *Metaphysische Anfangsgründe der Tugendlehre*, en *Werke*, vol. ix (ed. Rosenkranz y Schubert), p. 221.

sí es significativo, y sí logra restaurar una moral racional universal —en la (modesta) medida en que dicha restauración es posible— y muestra, así, en efecto, que el agente moral no es estúpido. Pero esta es una opinión personal que con gusto someteré a un mejor juicio. En su apoyo, ofreceré una interpretación benévola del argumento contractualista que, por razones de espacio, no podrá considerar todos los aspectos relevantes para mi tesis. A pesar de ello, espero que ésta quede expuesta con suficiente claridad a través de los «retazos argumentales» que emplearé.

2. La interpretación moral del contractualismo

Como es sabido, el argumento contractualista propone justificar las normas apelando a un acuerdo hipotético entre los afectados por ellas. Adoptar ese punto de vista hipotético significa situarnos totalmente fuera de la realidad en que vivimos (que es siempre ya normativa) y que nos constituye. Es, por tanto, muy discutible que sea un ejercicio intelectual posible. Cualquier definición del «estado de naturaleza» (es decir, del punto de vista hipotético apropiado para juzgar de forma reconstructiva la validez de las normas) es criticable. Sin embargo, el argumento contractualista parte de la aceptación de alguna definición de ese estado pre-social. Por ello, las teorías contractualistas más potentes se caracterizan por postular estados de naturaleza (o «posiciones originales», usando el término que popularizó Rawls) definidos lo más genérica y débilmente posible, para minimizar el impacto de las eventuales críticas y permitir un amplio asentimiento a los mismos.

En esa línea, el contractualismo moral supone que el acuerdo hipotético que va a justificar las normas morales es un pacto entre personas exclusivamente auto-interesadas, y racionales en el sentido más débil posible: capaces de reconocer y emplear medios efectivos para realizar sus fines, que vendrán determinados por sus deseos y preferencias subjetivos. Desde luego que esta hipótesis es criticable, pero mucho menos que si se supusiera, por ejemplo, que las personas en la posición inicial son necesariamente altruistas, o que tienen «intereses objetivos» que cualquiera puede determinar, o que deben siempre obedecer los mandamientos de la «Ley de Dios», etc. Esas posibilidades, y muchísimas otras, pertenecen al acervo de creencias de algunos miembros de nuestras sociedades; mientras que la posición original intenta recoger aquellos rasgos que (aunque en última instancia son también creencias) pueden aceptarse

como descriptivos de cualquiera de nosotros en un nivel básico. Sólo así se puede esperar que el argumento contractualista resulte, en general, plausible.

Especificar una posición original ampliamente aceptable tiene el inconveniente de reducir, tanto los participantes como las circunstancias, a esquemas muy simples, de los que podría decirse que no representan ni nuestra esencia como personas ni el tipo de situación en que normalmente ejercemos nuestras capacidades de decisión. Este inconveniente está justificado precisamente por el objetivo que el argumento contractualista persigue. Es cierto que, en un sentido, las partes en la posición original no tienen casi nada en común con nosotros. Nosotros somos personas morales, en quienes las ideas sobre el valor moral y la justicia de las acciones, y el respeto aprendido a ciertos principios, forman parte del carácter y (para bien o para mal) juegan un papel en lo que decidimos y en cómo actuamos. Pero el objetivo de la filosofía moral es justamente aclarar la racionalidad de este componente de nuestra personalidad. Tal aclaración exige poner en cuestión la capacidad moral como un todo, para someter a crítica nuestra racionalidad tal como la percibimos, cargada ya de normas y principios. De ahí la separación analítica de algo que es parte integral de nuestra personalidad, y su puesta entre paréntesis en beneficio del argumento. Esto explica que las partes en la posición original sean agentes despojados de su personalidad moral. En este sentido, como decía, nuestros *representantes* en la posición original difieren mucho de nosotros, y de esta diferencia se deriva un problema, porque no es nada fácil imaginar cómo elegirían (y por qué razones) estos agentes incompletos.

En este punto, los filósofos contractualistas más perspicaces se han dejado guiar, desde Hobbes, por las ciencias empíricas. Actualmente, la economía nos provee de un aparato teórico que facilita sobremanera ese trabajo: la teoría de la decisión racional define con precisión el comportamiento de agentes auto-interesados que actúan en un medio en que las consideraciones morales no son relevantes⁵. Y, como es una teoría con pretensiones normativas, ha de resultar

⁵ Aunque no voy a elaborar aquí este punto, no hay que entender esta afirmación en el sentido de que el medio económico es un medio *inmoral*; sino más bien en el sentido de que es un medio institucional cuyo contenido moral está incorporado en sus reglas, las cuales permiten a los participantes que se «despreocupen» de sus intereses morales personales y actúen —dentro de ese marco institucional— sólo en vista de sus intereses económicos.

persuasiva, esto es, las personas concretas han de sentirse identificadas (en una faceta al menos) con la definición del agente propuesta, de modo que acepten la relevancia para ellas de las eventuales conclusiones normativas. Si la teoría logra esta identificación, su uso como parte del argumento contractualista garantizaría que, a pesar de las diferencias, aún pudiéramos considerar a los agentes en la posición original como verdaderos representantes nuestros.

Por eso, a la hora de imaginar cómo elegirían los agentes en la posición original, las versiones contemporáneas del contractualismo admiten casi sin discusión los supuestos de la teoría de la decisión racional; y suponen que las personas pueden identificarse con los agentes económicos, cuya norma de decisión racional consiste en maximizar la utilidad esperada. Cualquiera puede identificarse con un agente económico en cierta medida: en lo que respecta a satisfacción de preferencias e intereses, todos somos agentes económicos (de hecho, adoptamos en exclusiva este papel en una amplia gama de situaciones institucionales: cuando nos comportamos como consumidores, como trabajadores, como ahorradores y, en general, cuando vamos «al mercado»). El paradigma económico es tan básico que incluso los comportamientos animales pueden interpretarse como respuestas maximizadoras con relación al medio. Por otro lado, la parte de nosotros que *no* es un agente económico (y que en forma legítima se revela contra esta identificación simplista), no necesita una justificación de la moralidad. Si las obligaciones morales nos resultan a veces cuestionables, es precisamente porque las evaluamos desde un punto de vista externo a su propia lógica (y ese punto de vista sólo puede ser el auto-interés). Por decirlo así, el hecho de que estemos embarcados en esta investigación es prueba suficiente de que tenemos más de agentes auto-interesados de lo que quizá estemos dispuestos a reconocer. Y en esa medida, nos es posible identificarnos con las partes en la posición original —definidas, en principio, como «agentes económicos».

La posición original, o estado de naturaleza, está, por tanto, formada (en el caso ideal de justificación de un principio moral universal) por todos los agentes racionales así descritos. En la medida en que la posición se considera pre-social, no pueden ser tenidas en cuenta las diferencias de poder, categoría, riqueza, estatus o educación que se deben a condiciones sociales presentes o heredadas. En ese sentido, las partes están en una posición de extraordinaria igualdad, como ya advirtió Hobbes.

Por otro lado, hay que suponer que los agentes en la posición original son *perfectamente racionales*. El alcance de esta suposición es malinterpretado con frecuencia, por lo que requiere un breve comentario. Es evidente que no todos los seres humanos tenemos las mismas capacidades. Y no es necesario imaginar, contra esta realidad, que en la posición original, todos poseen exactamente la misma cuota de esta importante capacidad de previsión, reconocimiento y cálculo que llamamos racionalidad. Ahora bien, si *cualquiera* quiere comportarse de manera racional, en el sentido instrumental, ha de *suponer* que los demás son tan racionales como él. No se trata de una suposición ontológica, sino estratégica. Sería simplemente un error táctico suponer que los demás son menos racionales (menos hábiles) que uno mismo. Sería como andar con la guardia baja. Así pues, lo de menos es hasta qué punto somos iguales en nuestras capacidades (aparte de que, despojados de las diferencias sociales, se puede pensar que lo somos bastante); lo que importa es que podemos legítimamente conjeturar que, al tener que interactuar (decidir, negociar, argumentar) en la posición original, cada cual supondrá que todos los demás son perfectamente racionales, y basará su acción en esta suposición.

Con esto concluye la definición de las partes, en los aspectos que más nos interesan. Sólo hay que recordar que en la posición original hay varios agentes y que lo que se trata de aclarar es el modo (reglas) de su interacción. Añado este comentario en apariencia innecesario porque tiene una consecuencia importante: a diferencia de otros enfoques morales, el contractualismo no puede justificar directamente obligaciones con uno mismo (salvo quizá las derivadas de las normas de la razón instrumental). Si fuera posible una situación de absoluta no-interacción (caso impensable en nuestras condiciones habituales, y que suele por ello identificarse con la ficticia situación de Robinson Crusoe en su isla), entonces, en verdad, sólo la maximización a largo plazo, en conjunción con las necesidades, deseos y preferencias del sujeto en cuestión, nos ofrecerían un criterio evaluativo sobre la racionalidad de sus acciones. En la situación de no interacción, el mundo real y la hipótesis de justificación normativa se colapsan, por así decir; y un agente solitario que quisiera justificar en forma racional sus acciones, no tendría por qué recurrir a pacto hipotético alguno, sino sólo al ideal de la razón maximizadora en condiciones paramétricas aplicado a su situación real. El problema de justificación normativa surge cuando ese ideal paramétrico no puede aplicarse porque enfrente no hay sólo objetos, sino sujetos que pueden

oponer su propia racionalidad instrumental a dicha aplicación. En ese momento, esto es, en interacción, surge la necesidad de criterios de acción distintos a la maximización de utilidad esperada, que resulta insuficiente en esa nueva condición. Y la justificación de tales criterios (y las normas en que se expresan) es el objetivo del argumento contractualista. Por eso no es trivial recordar que la posición inicial debe representar a varios agentes (tantos como sean los afectados por la norma o principio sometido a crítica). Al representar a esos agentes en las condiciones hipotéticas de la posición original, resulta que las normas que tienen sentido para todos desde esa perspectiva, pueden no coincidir con las que se están aplicando de hecho en la situación real (es decir, la hipótesis de justificación y la situación real no se colapsan, con en el caso de la decisión paramétrica, sino que pueden diferir). Y esto proporciona a todos el criterio evaluativo para la interacción. Un criterio que, decantado en principios y normas de interacción racionalmente justificadas, rebasará la norma deliberativa individual a la que cada quien fía la evaluación de las acciones que no implican a otros agentes.

El modelo contractualista de justificación de normas, instituciones o principios (en el caso del contractualismo moral nos interesan estos últimos, obviamente) se basa en la deliberación de cada individuo: «cada uno debe juzgar, en actitud deliberativa, si las normas sociales y las prácticas que originan son aquellas que tendría sentido para él acordar, si él y sus conciudadanos tuvieran la oportunidad de decidir conjuntamente los términos de su interacción»⁶; suponiendo, claro está, que esa decisión conjunta respondiese a las condiciones hipotéticas ya mencionadas, lo que garantizaría que es expresión de un acuerdo racional⁷.

Este modelo de justificación permite distinguir de forma procedimental qué normas o principios de interacción tendrían validez universal. Sólo las

⁶ Gauthier, D., «Political Contractarianism», *The Journal of Political Philosophy*, vol. 5, n° 2, 1997, pp. 132-148; p. 133.

⁷ Una versión altamente sofisticada de qué puede considerarse un acuerdo racional, puede encontrarse en Gauthier, D., *Morals by Agreement*, cit., cap. V. El propio autor ha reconocido que su elaboración sobre el acuerdo racional no es correcta en todos sus detalles, pero aún conserva su valor heurístico. Cfr., en este sentido, Gauthier, D., «Uniting Separate Persons», en Gauthier y Sudgen, R. (eds.), *Rationality, Justice and the Social Contract*, Ann Arbor, University of Michigan Press, 1993, pp. 176-192; p. 177-178.

normas que admiten una justificación contractual parten de una normatividad básica que todos aceptan y ejercitan: la incorporada en la deliberación racional. Y, por otro lado, el procedimiento contractual, esto es, el acuerdo unánime como mecanismo de decisión conjunta, garantiza que los intereses de todos los afectados son tenidos en cuenta. Ello es tanto más así en cuanto que el procedimiento para acordar se defina de forma apropiada como una negociación entre personas con intereses no necesariamente compartidos, pero que han de ceder en beneficio del interés subyacente de todos (en cuanto agentes racionales) que es mantener una interacción mutua beneficiosa (cooperación) y su medio necesario: el acuerdo normativo.

Aunque sería demasiado largo elaborar aquí este punto, puede mostrarse (a) que en una situación no constreñida por normas intersubjetivas de ningún tipo (como hay que suponer que es la posición original), la interacción no-cooperativa produce un resultado sub-óptimo; y (b) que la respuesta racional (optimizadora) a esa situación sub-óptima (dada la aproximada igualdad de los agentes) es la cooperación. A su vez, la condición necesaria para la cooperación es que cada agente desarrolle la disposición a contribuir a la creación de la estructura (normativa) que la hace posible⁸.

Esto significa que cabe mostrar de manera consistente la posibilidad de un acuerdo normativo a partir de una definición simple de las partes y su interacción en la posición original, y que ese acuerdo normativo incluye la idea de una disposición racional en cada agente a aceptar y eventualmente cumplir las normas que se reconocen como condiciones necesarias para tal acuerdo (y son a la vez producto suyo).

Debe ser evidente que en este punto, el argumento contractualista moral va más allá que sus antecesores. Según Hobbes, el estado de naturaleza hace que sea racional acordar, pero no haría surgir ninguna disposición nueva en las partes, es decir, no garantiza que sea racional cumplir lo acordado. Conscientes de ello, las partes autorizarían a un soberano con poder absoluto para imponer el acuerdo. Al mantener cada uno su disposición natural (egoísta), sólo la fuerza efectiva del poder político garantiza que todos tendrán un motivo para cumplir las normas acordadas.

⁸ Para una exposición detallada de este nudo argumental, cfr. Gauthier, D., «The Incomplete Egoist», en *Moral Dealing*, Ithaca, Cornell University Press, 1990, pp. 234-273.

Frente a este resultado, el contractualismo moral defiende —sacando todas las consecuencias de la suposición de perfecta racionalidad— que la lógica del acuerdo demanda una transformación en las partes que, de algún modo, «aprenden» a considerar racional el cumplimiento no necesariamente coactivo de las normas. En esquema, esta necesidad puede expresarse diciendo que para un agente racional sólo tiene sentido acordar si puede esperar que cada una de las demás partes va a cumplir; pero, por otro lado, cada uno de los demás mantendrá esta misma actitud condicional hacia él: sólo cumplirán si esperan que él lo haga. Para la teoría bayesiana estándar de la decisión racional, esto es una especie de callejón sin salida; y tal vez lo sea en ciertas situaciones reales que tienen esa estructura. Pero, en las condiciones ideales que plantea el argumento contractualista, es conocimiento público que la incapacidad de salir de ese callejón resulta en una situación que es peor para todos. Esta conciencia se conjuga con el hecho de que una transformación en la disposición racional de los agentes es posible (nada hay en la definición de la situación que la prohíba). Así, las partes en la posición original percibirían la inferioridad pragmática de la incapacidad (o negativa voluntaria) de modificar su disposición racional. Desde ese momento, comprenderían que una disposición meramente maximizadora es, en un contexto de interacción, un modelo limitado de racionalidad. Sin embargo, hemos supuesto que la racionalidad de las partes es perfecta y que ésta es una suposición mutua que todos han de conceder y tener en cuenta al actuar. Esta suposición crucial hace que las partes desechen cualquier alternativa estratégica que no consista en llevar a cabo la transformación que les permite aceptar como racionales las normas que incorporan y expresan el acuerdo cooperativo y estar sinceramente dispuestos a cumplirlas, al verlas como parte integral de las normas deliberativas individuales⁹.

Una precisión inmediata es que, al igual que el hecho de reconocer la racionalidad de la maximización de utilidad esperada en contextos paramétricos, no significa que siempre actuemos con éxito en ese tipo de contextos (de hecho, en ocasiones decidimos de manera voluntaria no obedecer los dictados de la racionalidad deliberativa); el reconocimiento de la racionalidad de cumplir

⁹ Para un detallado argumento formal que explica plausiblemente por qué no hay alternativa estratégica a esta transformación, cfr. McClennen, E., *Rationality and Dynamic Choice*, Cambridge, Cambridge University Press, 1990.

las normas contractualmente justificadas en contextos de interacción tampoco conlleva necesariamente su cumplimiento. Debe estar claro que lo que el argumento contractualista proporciona es una piedra de toque normativa, no una motivación psicológica, una terapia conductual, o una garantía del éxito de la interacción.

Ahora bien, como tal piedra de toque normativa, el argumento contractualista representa un *test* procedimental ideal de las normas que regulan la interacción real. Y, en la medida en que la lógica del procedimiento muestra que la coacción no es un componente necesario para la justificación de la existencia y vigencia de las normas (lo cual equivale a mostrar que éstas son racionales en sí mismas, y no sólo como medio para otros fines), dicho procedimiento puede considerarse legítimamente la expresión de un principio moral. Al asegurar la validez de las normas (o su crítica), el procedimiento contractualista representa un *test* moral al que podemos someter toda la normatividad social. Y en ese sentido funciona como principio moral que puede ser internalizado por todas las personas. Además, si las normas que sometemos a crítica son lo que habitualmente llamamos reglas u obligaciones éticas (que en general orientan las normas legales y políticas), entonces, el principio moral cifrado en el procedimiento contractualista alcanza su máximo contenido, al intervenir en forma directa como modelador de reglas e ideales que, a su vez, conforman el carácter moral de los miembros de una sociedad.

De nuevo, estos apuntes necesitarían un desarrollo que no pertenece a este lugar. Lo importante es que el procedimiento contractualista de validación de normas permite establecer un nexo entre la razón individual y el respeto a ciertos principios de interacción. Las personas pueden reconstruir retrospectivamente el procedimiento ideal mediante el cual una simple disposición (la racionalidad instrumental individual) exige su propia transformación en el contexto de la interacción. Como resultado, el componente moral de la personalidad, hipotéticamente excluido en el argumento, es recuperado, no ya como un elemento más del complejo normativo del «mundo de la vida» internalizado acríticamente mediante la socialización, sino como una parte integral de una constitución ética reflexiva y autónoma. Por otro lado, en esa reconstrucción se descubre el fundamento racional de ciertas normas, y la base para la crítica

racional de otras. Y es de este modo que el procedimiento contractualista se convierte en la ley moral del agente racional¹⁰.

El significado normativo de esta ley o principio moral debe entenderse como muy restrictivamente. Como tal, la conciencia del nexo necesario entre la racionalidad deliberativa y un principio moral, no obliga a nada. Esa conciencia sólo sirve para reconocer, especulativamente, que existen restricciones racionalmente justificadas con carácter intersubjetivo (pues dependen del hecho de la interacción). El modo en que esas restricciones se hacen efectivas es contingente. En nuestros días, adoptan con frecuencia la forma de normas jurídicas¹¹. En sociedades suficientemente homogéneas, pueden adoptar otras formas de normatividad más difusa. En cualquier caso, algún tipo de sanción suele asociarse a su incumplimiento, lo cual prueba que la conciencia de su justificabilidad racional (y de su legitimidad democrática formal en el caso de las normas legales y políticas) no tiene gran ascendiente sobre el ciudadano medio. Sin embargo, es crucial que el reconocimiento racional de la validez intersubjetiva de las normas sea, al menos, posible. De hecho, el miedo al castigo no sería explicación suficiente del orden observado en nuestras sociedades, lo que parece apuntar a que un elevado número de personas actúa conforme a normas y principios internalizados. Para tales personas (que en uno u otro grado, somos todos) es primordial contar con un procedimiento de validación y crítica intersubjetiva de esas normas (es decir, con un principio moral). De otro modo, el aprendizaje y el progreso ético serían imposibles, y los patrones sociales heredados bien se reproducirían de forma inevitable, bien sufrirían la pérdida total de validez y asistiríamos a la *eutanasia* de la moral de que la hablaba Kant.

¹⁰ El sentido en que llamo «ley moral» al procedimiento contractualista de validación de normas prácticas es precisamente el mismo sentido en que Kant denomina «ley moral» a la ley fundamental de la razón pura práctica (cuyo papel es paralelo al de nuestro procedimiento: servir como punto de referencia crítico para las máximas o reglas de acción particulares). Cfr. *Kritik der Praktischen Vernunft (Crítica de la razón práctica)*, *Werke*, (ed. de Rosenkranz y Schubert) vol. VIII, p. 143.

¹¹ Esta proposición es a la vez de hecho y normativa. Quiere decir que, en el tipo de sociedades liberales en que vivimos, la forma de afirmar una restricción moral consiste en reclamar un derecho, o la imposición de una obligación legal. Y la forma de expresarla, una vez reconocida democráticamente, es mediante un texto legal.

Personalmente, mi interpretación del argumento contractualista moral —una interpretación que podría llamarse epistemológica— es la siguiente: el argumento representa el modo de conocer las demandas normativas intersubjetivas de la razón práctica en condiciones de interacción, es decir, en las condiciones de la justicia. Si esas condiciones no se dan —porque no existe interacción o porque (debido a cualesquiera otras circunstancias) su estructura no implica que el uso de la deliberación individual sin restricciones produzca un resultado sub-óptimo—, aquellas demandas normativas no juegan papel alguno. Pero en las circunstancias de la justicia, tales demandas normativas representan límites intersubjetivos absolutos: límites que (actualizados en forma de principios, obligaciones o derechos subjetivos) un agente racional no puede dejar de reconocer —aunque, obviamente, siempre puede dejar de respetar en uso de su libertad. El contenido de esas restricciones no puede anticiparse, pues dependerá de la situación de interacción concreta y de las personas afectadas, pero eso no nos permite pensar que se trata de límites convencionales o arbitrarios. Ello está garantizado por el carácter intersubjetivo del procedimiento adecuado para su elucidación.

3. Racionalidad moral y racionalidad económica

Si el contractualismo es aceptable como filosofía moral, es evidente que el agente moral *no* es un maximizador a largo plazo. Por el contrario, es una persona que reconoce que la racionalidad no exige sólo el cálculo de las consecuencias de sus acciones y el empleo de la mejor estrategia para realizar sus intereses, sino además el respeto a reglas y principios legítimos (que aunque no son accesibles a través de un argumento que parte de la racionalidad medios/fines de cada persona, tal vez comporten sacrificios ocasionales no justificables desde esa perspectiva puramente instrumental. Ser racional, por tanto, no es ser un maximizador a largo plazo, sino una persona moral, lo cual no significa negar que cada quien tenga sus intereses y lleve a cabo las mejores estrategias posibles (dentro del respeto a las normas) para avanzarlos.

Ahora bien, este no es un resultado concluyente. La compleja justificación contractualista de la moralidad parece mostrar que es racional ser moral *por motivos auto-interesados*. Lo que el argumento parece decir es que, al agente moral *le va mejor* que al agente maximizador. El tipo de reconstrucción racional en que el argumento consiste se dirige al agente económico y, por tanto, lo que

muestra es el sentido en que es económicamente racional *convertirse* en un agente moral. Según esa lógica, el agente moral resulta ser el *verdadero maximizador*. Esta apariencia da pie a un argumento escéptico, que defendería que no hay, en realidad, criterio de racionalidad superior a la maximización de la utilidad esperada; que el agente moral sólo es racional en la medida en que, consciente o inconscientemente, se comporta como un maximizador a largo plazo; y, por tanto, que la moral es una impostura.

Hay básicamente dos momentos en que el escéptico puede lanzar este desafío. Un primer momento ocurre en el interior del propio argumento contractualista (y en ese caso, el escéptico estaría concediendo, al menos, que tal argumento es un modo posible de justificar demandas normativas válidas). Su crítica consistiría en sostener que el único principio normativo que obtiene, de hecho, el asentimiento de todos en la posición original, es una versión refinada (y tal vez colectiva) de la normatividad ya incorporada en la racionalidad consecuencialista individual. Tal «refinamiento» no requeriría ninguna transformación en los agentes, que seguirían ejerciendo, tras el acuerdo, el mismo tipo básico de racionalidad supuesta al inicio¹². La idea subyacente a esta posición es que, incluso en las condiciones ideales del acuerdo normativo, es más beneficioso para cada uno *fingir* adhesión al acuerdo con la doble intención de cumplirlo cuando sea inevitable y violarlo cuando sea posible. Tal disposición (derivada del principio de maximización supuesto como premisa) es la más beneficiosa para cada agente dadas las circunstancias y, por tanto, no habría base alguna para introducir un criterio alternativo de racionalidad.

Hasta aquí el argumento del escéptico tiene cierta plausibilidad, pero ninguna novedad —es el mismo argumento que llevó a Hobbes a negar toda validez al pacto en tanto no se establece la autoridad política. En mi opinión, tampoco es del todo plausible, pues parece desconocer el alcance y consecuencias del supuesto de perfecta racionalidad de las partes, comentado anteriormente¹³. Sin embargo, no creo que sea necesario reproducir los argumentos ya

¹² Creo que esta es la posición de algunos utilitaristas, como J. Harsanyi; cfr. su «Morality and the Theory of Rational Behaviour», en Sen, A. y Williams, B., *Utilitarianism and Beyond*, Cambridge, Cambridge University Press, 1982, pp. 39-62.

¹³ De hecho, tampoco el argumento de Hobbes es plausible. Un análisis del estado de naturaleza en términos de teoría de juegos revela que un Estado-Leviatán no es su consecuencia necesaria. Cfr. Hampton, J., *Hobbes and the Social Contract Tradition*, Cambridge, Cambridge University Press, 1986.

expuestos arriba sobre la justificación, entre agentes maximizadores, de una disposición a cumplir acuerdos racionales¹⁴. En última instancia, el escéptico puede siempre enroscarse en una posición dogmática (y carente de soporte empírico, por cierto), según la cual la maximización de la utilidad esperada es, por definición, el único criterio de elección racional; a partir de ahí, toda discusión es inútil.

El segundo momento en que el escéptico puede desafiar el argumento moral es el momento de la aplicación de las restricciones morales, esto es, en la vida cotidiana. En este caso el escéptico no presta atención al argumento moral en sí mismo. Simplemente apunta que, en algunas situaciones, mientras el agente moral abraza de manera racional ciertas restricciones que le impiden de hecho elegir estrategias maximizadoras, un agente a-moral despreciaría esas restricciones, desplegaría las mejores estrategias maximizadoras, y satisfaría en mayor medida sus intereses¹⁵.

Si, después de todo, el agente moral logra resulta avanzar sus intereses aún más que el directamente maximizador, nadie podrá negar que la moralidad es la mejor estrategia maximizadora a largo plazo. Pero si no lo hace, nadie podrá negar, tampoco, que no fue tan racional (en sentido instrumental) como podría haber sido si hubiera despreciado la moralidad en esta ocasión¹⁶.

En el primer momento, el escéptico pasaba por alto las implicaciones de la perfecta racionalidad, de modo que su crítica al argumento moral quizá no fuera concluyente. En este segundo momento, lo que hace es admitir como un hecho que *hay* agentes que creen que la moral está justificada y mostrar que sería fácil para otras personas aprovecharse de su «simplicidad». El escéptico nos trae

¹⁴ Sobre esto Cfr. Gauthier D., *Morals by Agreement*, cit., cap. VI, en especial pp. 174 y ss.

¹⁵ Hay que precisar que, en un sentido trivial, el agente moral también elige una estrategia maximizadora, en la medida en que, para él, la sujeción a las normas tiene más valor que las estrategias alternativas. A lo que se refiere esta distinción es al sentido no trivial en que el agente moral es capaz de ver cuál sería su mejor estrategia si no hubiera normas morales (o estas no fueran de aplicación al caso) y, a pesar de todo, no la pone en práctica para evitar el conflicto con la norma moral.

¹⁶ Evidentemente, nos referimos a un «desprecio de la moralidad» que resulta exitoso, y por ello logra avanzar eficazmente los intereses del agente. Este agente eficazmente inmoral es magistralmente descrito por Glaucón en *La República*, 360e1-362c8.

de nuevo la cuestión con que comenzábamos este ensayo: ¿es el agente moral el más hábil maximizador a largo plazo, o el más completo estúpido?

En este punto, debo reconocer que el análisis de la lógica del desafío del escéptico (que desafortunadamente es otro aspecto que no puedo desarrollar aquí) me ha llevado a convencerme de algo que, hace sólo unos meses, habría rechazado. Mi respuesta a este desafío habría sido que, en efecto, la moralidad, bien entendida, es la mejor estrategia a largo plazo. No, desde luego, en el sentido trivial de que el cumplimiento instrumental de las normas éticas asegure mejor el logro de los intereses a largo plazo del agente, sino en el sentido de que sólo un agente moral puede esperar alcanzar la felicidad con certeza, mientras que el agente económico, por muy prudente, hábil y previsor que sea, está sujeto a la contingencia de la cadena causal que liga sus decisiones y acciones con el resultado (fin) que espera obtener y que las orienta¹⁷.

Y si la moral, bien entendida, es la única garantía de felicidad, entonces el agente económico, el puro maximizador de utilidad, el frío calculador consecuencialista, sería, estrictamente hablando, irracional; a pesar de sus desafíos.

Sin embargo, hoy no estoy tan convencido de esto. Mostrar por qué, no me resulta fácil. Emplearé el socorrido paralelo de la epistemología para aclarar de dónde provienen mis dudas. Es evidente que, cuando hablamos del mundo, todos estamos suponiendo alguna teoría de la verdad (de otro modo, sería imposible cualquier argumentación). Esto quiere decir que, explícita o implícitamente, todos hemos de admitir que debe existir algún método que podría justificar la verdad de ciertas proposiciones (y la falsedad de otras). Sin embargo, de aquí no podría concluirse que todos seamos realistas, o que todos abracemos una concepción de la verdad como correspondencia. De hecho, si en vez de hablar del mundo nos pusiéramos a discutir de epistemología, quizá resulta que todos (o algunos de nosotros) somos escépticos, o convencionalistas, etc.

En el campo de la acción, es también evidente que todos decidimos entre los cursos de acción que nos parecen posibles, tanto en el sentido físico, como

¹⁷ La conexión entre moralidad y felicidad de que estoy hablando aquí es de inspiración kantiana (Cfr. *Kritik der Praktischen Vernunft*, cit. p. 252), salvo que Kant está en posición de defenderla, y yo no; porque él supone el concurso providencial del «Autor de la Naturaleza».

en el normativo y, al hacerlo, todos estamos suponiendo que algunos criterios de elección (uno al menos) son válidos (de otro modo sería imposible cualquier decisión). Esto quiere decir que, explícita o implícitamente, todos hemos de admitir que hay algún procedimiento que podría justificar la validez de ciertas normas (y, consiguientemente la justicia —entendida como adecuación a la norma válida— de ciertos modos de interacción, y no de otros). Dicho de otra forma, el nexo entre la deliberación individual y la validez normativa que el argumento contractualista hace explícito, puede verse como un supuesto general de la interacción. Por lo tanto, al igual que la mera participación en una discusión científica pre-supone alguna noción de verdad, parece que la mera participación en la interacción pre-supone alguna noción de justicia, es decir, alguna ley moral. Sin embargo, no podemos concluir de aquí que todos abracemos la *misma* moralidad. De hecho, si en vez de actuar en el mundo, discutimos sobre filosofía moral, quizá resulte que algunos somos convencionalistas (o escépticos), y empleamos la moralidad como un medio conveniente para avanzar nuestros intereses.

Podríamos preguntarnos, entonces, por la escogencia del argumento contractualista. Y la mejor respuesta emplearía, quizá, las palabras de Kant: « la realidad objetiva de la ley moral no puede probarse deductivamente mediante la razón teórica, ni en su uso especulativo, ni en el empírico »¹⁸ Con Kant, opino que la ley moral está establecida por sí misma (interpretése esto social, cultural, histórica, psicológica o metafísicamente), y no podría introducirse como una conclusión de ningún razonamiento deductivo. Ello no significa, y esto es lo que el argumento muestra, que el agente moral carezca de justificación racional para sus acciones. Tal justificación existe y, además, debido a la forma del argumento, es eficaz para denunciar la inadecuación de la maximización directa en sus propios términos.

Por eso se puede esperar que entre agentes racionales (incluso concebidos inicialmente como meros individuos económicos) la moralidad impere como ley (aunque no siempre se obedezca, o incluso aunque nunca se hubiera obedecido). Pero quienes no la reconocen como ley, no pueden ser considerados irracionales sin más. Porque si bien la racionalidad como maximización es

¹⁸ *Kritik der Praktischen Vernunft*, cit., p. 163.

inadecuada, no es inconsistente ni contradictoria en sus términos. Si todo lo que hubiera en el mundo real fuese maximización (bien miope, bien a largo plazo), sería imposible juzgarla como irracional. Y, aunque de hecho en el mundo no hay sólo maximizadores, sino también agentes morales, es dudoso que tal juicio (que siempre es posible) tenga sentido. Hay que tener en cuenta que la maximización es obviamente racional desde la perspectiva maximizadora, de modo que su calificación como irracional habría de hacerse desde un paradigma axiológico ajeno a la misma. Pero ese paradigma no sería legítimo para el agente maximizador, que —en el caso extremo— podría igualmente calificarlo de irracional. Esta descalificación mutua carente de sentido confirmaría simplemente que se trata de paradigmas inconmensurables.

Como tal, el agente maximizador es racional. Inmoral, pero racional. Nuestra organización social así lo sostiene: si considerásemos que el puro auto-interés irrestricto es una forma de irracionalidad, no habría cárceles; sólo manicomios o escuelas. Sin embargo, estamos concediendo, como participantes en esa organización, que la maximización, y la inmoralidad que a veces conlleva, es racional. Nos protegemos de la inmoralidad (incluso de la propia) mediante leyes, precisamente porque la vemos como una alternativa racional.

Pero de esto no se sigue que el agente moral sea irracional. Ahí reside la falacia del desafío del agente económico. El argumento contractualista muestra no sólo que la moralidad no es contradictoria en sus términos, sino que es racional, incluso juzgada desde el criterio económico más elemental.

Que la moralidad no sea contradictoria en sus términos implica que para un agente moral sería simplemente irracional comportarse como un maximizador puro de la utilidad esperada. Lo que para el agente económico es la expresión directa de la racionalidad, carece de sentido para el agente moral. Como opción ética, el agente moral ha desarrollado una personalidad (una estructura de preferencias, diría el economista) en la que se concede siempre más valor a la acción que se considera correcta, *por el hecho de ser correcta*. Cuando actúa, el agente moral exhibe una coherencia entre sus sentimientos, preferencias y creencias que no puede ser disputada desde el punto de vista económico.

Y aquí se produce una diferencia que rompe el paralelo (y la inconmensurabilidad) entre ambas concepciones de la racionalidad, cuando se trata de juzgar una en términos de la otra. Ambas son internamente coherentes y, en ese sentido, racionales en sus propios términos. Además, la moralidad es irracional

para el hombre económico, y la maximización directa es irracional para el agente moral. Esto expresa la racionalidad interna de cada paradigma y es lo que provoca que la precaución social contra la inmoralidad sea con más frecuencia la amenaza que la educación.

Ahora bien, mientras que el agente moral ha de ser considerado racional (dada su personalidad ética) incluso desde el punto de vista de la Teoría Económica de la Decisión Racional; lo mismo no ocurre desde el otro lado: el agente económico sólo puede ser considerado racional en términos económicos. En términos morales *no es racional*. Y no porque no comparta la «preferencia» por la justicia, sino porque no ha reconocido la totalidad de las implicaciones normativas de la interacción en la que se halla inmerso como agente racional.

4. Conclusión

En la medida en que el argumento contractualista justifica la moralidad en términos de racionalidad maximizadora, el agente moral está autorizado a criticar el concepto puramente maximizador de la racionalidad; mientras que el agente económico no está autorizado a la crítica contraria.

Sólo el astuto inmoral que trata de aprovechar la apariencia de hombre decente en su propia ventaja sería criticable desde el punto de vista económico, en cada ocasión en que fallaran sus previsiones y no lograra su objetivo a largo plazo. Si fuésemos necesaria y únicamente individuos económicos, y la moral (así como los demás códigos sociales) fuesen aceptables sólo como «males necesarios», o «engaños beneficiosos»¹⁹, impuestos por la necesidad de la interacción, entonces la conciencia de su verdadero significado socavaría nuestra inclinación aprendida hacia la moralidad.

Mas si la reflexión que dejo anotada arriba es acertada, lo que el argumento contractualista muestra no es la utilidad de la moralidad, sino su necesidad como punto de referencia normativo último de cualquier interacción que pueda ser considerada racional en absoluto. Así, la perspectiva moral, de la que tenemos noticia inicialmente por la experiencia de nuestros sentimientos y nuestra aventura ética, porque permea nuestras decisiones y la concepción de nosotros mismos; esa perspectiva, desafiada por el componente instrumental de

¹⁹ Cfr. Gauthier, D., *Morals by Agreement*, cit., p. 339.

la racionalidad, que también reside en nosotros, resulta justificable incluso en términos instrumentales. Pero esa conciencia no conlleva la difuminación (eutanasia) de la moralidad, porque si hemos llegado a tener noticia de la moralidad en primer lugar, es precisamente porque el agente moral no actúa en términos de auto-interés. El nivel de la justificación no afecta, por así decir, al nivel de la motivación (excepto en lo que pueda indirectamente ayudar a que los agentes morales se sientan más firmes en sus convicciones).

Así pues, el agente moral no es el verdadero maximizador a largo plazo, por muy bien que le vaya en la vida. El análisis económico confirmará, en muchas ocasiones, que el agente moral llevó en efecto una vida alta (incluso máximamente) valiosa en términos de sus preferencias. Pero si lo logró, fue justamente porque no se lo propuso como fin. De hecho, en la medida en que nos comportamos moralmente —es decir, en la medida en que somos auténticamente racionales al interactuar— el criterio de la maximización de utilidad no tiene más aplicación que la que le conceden ciertos contextos institucionales legítimos. Por decirlo de manera gráfica, el agente moral es, simplemente la persona capaz de comprender que, en un marco social, hay motivos racionales para ser moral; pero consciente de que, en la práctica, sólo los motivos morales pueden dar sentido a las acciones moralmente correctas.