

LA ESTRUCTURA DE LA ÉTICA Y LA MORAL

RICHARD M. HARE

UNIVERSIDAD DE FLORIDA

Debo empezar diciendo lo que creo que es el objeto de la empresa llamada filosofía moral: encontrar una forma de pensar mejor, es decir, más racionalmente, sobre las cuestiones morales. El primer paso hacia tal meta es *entender las preguntas que uno se formula*. Aunque ello podría parecer obvio, difícilmente se hallará quien trate de hacerlo. Tenemos que entender lo que queremos decir con expresiones como “debo” [*I ought*].* Y entender el significado de una expresión como ésta supone la comprensión de sus propiedades lógicas o, en otras palabras, qué es lo que implica o a qué nos comprometemos al usarla. Así, si vemos que no podemos aceptar aquello a que nos compromete, habremos de renunciar a decirla. Y en esencia, en eso consiste la argumentación moral. La ética, estudio de la argumentación moral, es pues una rama de la lógica. Éste es uno de los niveles de pensamiento que interesan al filósofo moral. Los otros se refieren a asuntos más sustanciales; pero este primer nivel lógico o, como a veces se le llama, metaético, es el fundamento de los otros.

Puesto que el tipo de ética que nos ocupa es un tipo de lógica, tiene que emplear los métodos de la lógica. ¿Pero cuáles son éstos? ¿Cómo encontramos qué se sigue de qué, qué implica qué o a qué me compromete decir, por ejemplo, “Debo alistarme en el ejército” o “Debo unirme a la revolución”? Esto se refiere, en parte, a un cuestionamiento general acerca del método lógico, del cual no podré ocuparme en profundidad por falta de espacio. Tan solo declararé cuál es mi posición respecto de algunas cuestiones decisivas. En primer lugar, creo que es útil —es más, esencial, y espero que no se vea como rasgo de pedantería— distinguir entre dos tipos de cuestiones. A las del primer tipo las llamaré cuestiones *formales* y a las del segundo cuestiones *sustanciales*. Las cuestiones formales son aquellas que pueden responderse recurriendo sólo a la forma, es decir, a las propiedades puramente lógicas, de las respuestas que se han

* Hemos traducido, a lo largo de todo el artículo, “ought” como “debo” o “deber”. [*Ésta y las demás notas marcadas con ‘*’ son del traductor.*]

propuesto para ellas. A esta clase pertenecen las cuestiones de que nos ocupamos en la metaética. En esta parte de nuestro trabajo no nos está permitido traer a cuento ninguna suposición sustancial.

Ilustraré la distinción mediante un ejemplo que nada tiene que ver con la ética, porque es un ejemplo más claro y no da por resuelta ninguna cuestión ética. El ejemplo proviene de un conocido artículo de los profesores Strawson y Grice donde se refuta (con acierto, a mi modo de ver) una afirmación hecha en un artículo aún más conocido del profesor Quine (algunos de los asertos de Quine pueden ser correctos, pero está bien claro que Strawson y Grice han refutado éste en particular).¹ Supongamos que digo "Mi niño de tres años de edad comprende la Teoría de los Tipos de Russell". Cualquiera estará seguro de que mi afirmación es falsa. Pero lógicamente podría ser verdadera. Por otro lado, supóngase que digo "Mi niño de tres años de edad es un adulto". Sabemos que no puedo decir esto consistentemente con sólo saber el significado de las palabras; y es obvio que éste no es el caso de la primera proposición.

Esto ilustra la distinción entre lo que llamo cuestiones formales y lo que llamo cuestiones sustanciales. Se aplica de igual forma a las cuestiones morales, que también pueden dividirse en estos dos tipos. Supongamos que digo "No hay nada malo [*wrong*"]* en azotar a la gente por diversión". Las razones de la gente para estar en desacuerdo conmigo (y más tarde volveré sobre lo que estas razones puedan ser) son de un tipo muy diferente del que serían si yo hubiera dicho "No hay nada malo en hacer lo que uno no debería hacer". Sabemos que no puedo decir consistentemente esto último con solo saber los significados de las palabras "deber" y "malo"; sí puedo, en cambio, formular *consistentemente* la primera proposición; todos consideramos que ésto es algo terrible de decir —sólo lo diría una persona muy perversa—, pero decirlo no sería *lógicamente* inconsistente.

Aquí no es necesario discutir el rechazo de Quine de la noción de analiticidad y el de la de sinonimia. Strawson y Grice pueden estar en lo correcto al defender estas nociones; pero aunque no lo estén, los asertos formales que yo necesito hacer acerca de los conceptos morales no necesitan formularse en términos de ellas, sino sólo en términos de la noción de verdad lógica, que Quine en ese artículo acepta. Ello se debe a que los conceptos morales son formales en un sentido todavía más estricto del que hasta ahora he sostenido.

¹ H. P. Grice, y P. F. Strawson, "In Defense of a Dogma", *Ph. Rev.* 65 (1956); W. V. Quine, "Two Dogmas of Empiricism", *Ph. Rev.* 60 (1951), reproducido en su *From a Logical Point of View* (Harvard University Press, 1953).

* Traducimos en todo el artículo "*wrong*" como "malo".

Esto es, su explicación no requiere de ninguna estipulación semántica material, sino únicamente de una referencia a sus propiedades puramente lógicas. Las propiedades semánticas de las palabras morales tienen que ver solamente con sus significados descriptivos particulares, que no forman parte de su significado en sentido estrecho y no afectan su lógica, aunque el hecho de que tienen que tener algún significado descriptivo sí la afecte.²

Se advertirá que el ejemplo que acabo de dar de un enunciado que todos nosotros debemos rechazar por razones lógicas, “No hay nada malo en hacer lo que no debemos hacer”, es un enunciado cuyo contradictorio (“Hay algo malo en hacer lo que no debemos hacer”) es una verdad lógica. Esto se debe a que “deber” y “malo” son interdefinibles puramente en términos de sus propiedades lógicas, sin traer a colación sus significados descriptivos o su semántica, tal como lo son “todos” y “algunos” en la mayoría de los sistemas de lógica cuantificacional.

A continuación, debo tocar un punto que cobrará importancia fundamental para la argumentación moral. Cuando resolvemos cuestiones del segundo tipo en cada caso (es decir, en las cuestiones formales) no estamos autorizados para echar mano de ningún otro tipo de consideración excepto las que pueden establecerse con base en nuestra comprensión de las palabras o conceptos usados. Para seguir con nuestros dos ejemplos: sabemos que no podemos decir “No hay nada malo en hacer lo que no se debe hacer” porque sabemos lo que significan “malo” y “deber”; y sabemos que no podemos decir “Mi niño de tres años de edad es un adulto” porque sabemos lo que significan “niño” y “adulto”. Si, para establecer que no podemos decir estas cosas, tuviéramos que hacer cualesquiera otras consideraciones distintas, las cuestiones de si podemos o no decirlas no serían cuestiones formales.

En general, establecemos tesis en lógica (o en el tipo de lógica al que me refiero, que incluye el tipo que usamos en ética) recurriendo a nuestra comprensión de los usos de las palabras y nada más. La razón de que sea tan importante comprender las palabras radica en que esta lógica es el fundamento de la argumentación moral. Debemos advertir que ésta es una característica del método por el cual me pronuncio, la cual lo distingue muy radicalmente de casi todas las teorías éticas que se proponen en nuestros días. Todas estas teorías se remiten en un momento u otro, y a menudo con gran frecuencia, a las convicciones morales sustanciales que tienen sus proponentes, y que éstos esperan que compartan sus lectores. Aunque reconozco que

² Véase sobre esto mi *Moral Thinking (MT)*; Oxford University Press, 1981), p. 2 y s., y mi *The Language of Morals* (Oxford University Press, 1952), pp. 122 ss.

muchas personas no creen en la distinción que he venido haciendo entre las cuestiones morales formales y las sustanciales, y que por lo tanto se sienten en libertad para usar en apoyo de sus teorías lo que yo llamaría sus convicciones sustanciales, sigo creyendo que mi forma de proceder proporciona una base más firme a la ética.

Permítaseme exponer brevemente mis razones para pensar así. Si discutimos sobre alguna cuestión moral (por ejemplo sobre la cuestión “¿Debo alistarme en el ejército?” o “¿Debo unirme a la revolución?”), una de las cosas que desde el principio hemos de tener claras es la *cuestión* de que se trata. Es decir, si no queremos enfrascarnos en una conversación de sordos, debemos querer decir las mismas cosas mediante las palabras con las que formulamos nuestra pregunta. Pero si *en efecto* queremos decir lo mismo mediante las palabras, tendremos una base sólida de acuerdo (aunque este acuerdo sea formal y no sustancial) sobre la cual podremos basar nuestros argumentos futuros. Si se mantiene la distinción entre formal y sustancial, entonces podemos tener este acuerdo formal a pesar de nuestro desacuerdo sustancial. Entonces podemos, como espero mostrarlo, usar el acuerdo formal para someter a prueba los argumentos que usa cualquiera de nosotros para apoyar sus puntos de vista. Podemos preguntar “¿Puede él *decir* consistentemente esto?”, o bien “¿Puede él decir consistentemente *esto*, si también dice *aquello*?”

Por otro lado, si la gente introduce sus propias convicciones sustanciales en los fundamentos mismos de sus argumentaciones morales, no será capaz de argumentar sólidamente en contra de nadie que no comparta tales convicciones. Esto es lo que hacen los filósofos llamados *intuicionistas*; y yo puedo afirmar con cierta confianza que mi posición es mucho más sólida que la de ellos, puesto que yo no me apoyo en nada salvo en lo que tienen que estar de acuerdo todos los que formulen las mismas preguntas que yo trato de responder. Ésta es la razón por la que he afirmado que antes de empezar una discusión tenemos que ponernos de acuerdo sobre el significado que le damos a nuestras cuestiones. Eso es *todo* lo que requiero para empezar.

Baste por ahora con lo que se ha dicho sobre el método de la ética. Podría decir mucho más al respecto, y para llenar todos los huecos tendría que hacerlo, como lo he hecho en otro lugar,³ pero ahora prefiero seguir adelante y decir qué espero establecer mediante este método y de qué manera nos sirve para tratar con cuestiones reales. En el nivel formal o metaético, necesito establecer sólo dos tesis, y por razón de sencillez las formularé como tesis a propósito de la

³ Véase *MT*, esp. cap. 1.

palabra “deber” y su lógica. Podría haber hablado en su lugar de otras palabras, como “correcto” y “bueno”. Pero prefiero hablar de “deber”, porque es la palabra más simple que usamos en nuestros cuestionamientos morales.

Así pues, tenemos aquí dos características lógicas de la palabra “deber”, como se da en las preguntas “¿Debo alistarme en el ejército?” y “¿Debo unirme a la revolución?” Son partes de aquello con lo que me comprometo si digo “Sí, debo hacerlo”. A la primera se le llama a veces *prescriptividad* de los juicios morales. Si digo “Sí, debo alistarme en el ejército”, y lo digo sinceramente y en su sentido pleno —si realmente creo que debo hacerlo—, entonces me alistaré en el ejército. Existen, naturalmente, sentidos no plenos de “deber” o de “creer que debo”, en los cuales podría decir “Creo que debo pero no lo haré”, o bien “Debo, pero ¿qué más da?”. Pero cualquiera que haya estado en esta situación (como yo: ésa fue una de las cosas que me llevaron a emprender el estudio de la filosofía) sabe que el meollo de preguntarse “¿Debo hacerlo?” consiste en ayudarnos a decidir la pregunta “¿Lo haré?”, y la respuesta a la primera pregunta, cuando se formula en este sentido, implica una respuesta a la segunda. De no ser así, ¿qué sentido habría tenido formularla?

La segunda característica de la palabra “deber” en que me apoyaré se llama generalmente *universalizabilidad*. Cuando digo que debo, me comprometo con más que ese *yo* debo. La prescriptividad exige que el hombre que dice “Yo debo” actúe en consecuencia siempre que el juicio se aplique a él y si él puede actuar de tal forma. La universalizabilidad significa que, al decir “debo”, el individuo se compromete a estar de acuerdo en que cualquiera que se encontrara en las mismas condiciones debería hacer lo mismo. Si digo “Debo, pero existe alguien exactamente en las mismas circunstancias, que trata de tal manera a alguien totalmente semejante a la persona a la que yo debo tratar de tal manera, pero él no tiene el deber de tratarla así”, entonces los lógicos me mostrarán su extrañeza. En el caso de dos personas exactamente comparables en situaciones exactamente iguales, es lógicamente inconsistente decir que la primera debe hacer algo en tanto que la segunda no debe hacerlo. Debo explicar que la semejanza de las situaciones se extiende a las características personales, y en particular a las preferencias y aversiones de las personas que se hallan en esas situaciones. Si, por ejemplo, la persona a la que yo azotaba gusta de ser azotada (así sucede con algunas personas), ello significaría que la situación no era exactamente semejante al caso normal, y esa diferencia podría ser digna de ser tomada en consideración.

De esta forma, reuniendo estas dos características de prescriptividad y universalizabilidad, vemos que si digo "Debo tratarlo a él de tal manera", me comprometo a decir no sólo que lo trataré a él así (y en consecuencia a tratarlo así), sino a decir que él me debería tratar a mí de tal manera si nuestros papeles fueran precisamente inversos. Tal es, como veremos, la forma en que el argumento moral adquiere su fuerza. Debo repetir que no es una parte esencial de mi argumento que *todos* los usos de "deber" tengan estas características. Todo lo que sostengo ahora es que, en ocasiones, cuando nos preguntamos "¿Debo?", usamos la palabra de esta manera. Me dirijo a quienes se hacen tales preguntas, como estoy seguro que mucha gente se las hace. Si alguien deseara formular preguntas *diferentes*, tendría que emplear una lógica *diferente*. Pero estoy muy seguro de que a veces nos planteamos cuestiones prescriptivas universales y estamos en desacuerdo sobre ellas —es decir, a propósito de qué prescribir universalmente para todas las situaciones de cierto tipo, independientemente de quién sea el agente o la víctima. Me contentaré con poder mostrar cómo podemos argumentar válidamente sobre tales cuestiones, cuyo carácter lógico está determinado por ser *tales* cuestiones, es decir, cuestiones prescriptivas universales.

Debo agregar que todo el sentido de tener un lenguaje moral con estas características —un lenguaje cuyo significado está determinado únicamente por sus características lógicas, y que por tanto puede emplearse en la discusión de dos personas que poseen convicciones morales sustanciales muy diferentes—, está en que las palabras significarán lo mismo para ambas, y ambas estarán regidas en su discusión por las mismas reglas lógicas. Si sus diferentes convicciones morales llegaran de alguna manera a incorporarse en los significados mismos de sus palabras morales (como sucede con *ciertas* palabras morales, y como erróneamente creen algunos filósofos que sucede con todas ellas),⁴ entonces desde el principio los protagonistas de la discusión estarían hablando de cosas diferentes; muy pronto terminaría su discusión moral y tendrían que dirimirla luchando. Debido a este carácter formal de mi teoría sobre las palabras morales, pienso que es más útil que las teorías de otros filósofos que intentan incorporar sus propias convicciones morales en los significados de las palabras o en las reglas del argumento. Esto es particularmente cierto cuando nos ocupamos de los tipos de problema moral acerca de los cuales la gente tiene convicciones radicalmente diferentes. Si las convicciones

⁴ Véase mi "A *Reductio ad Absurdum* of Descriptivism", en S. Shanker (ed.), *Philosophy in Britain Today*, Croom Helm, Londres, 1986; aparecerá también, al igual que el presente artículo, en R. M. Hare, *Essays in Ethical Theory* (Oxford University Press) (en prensa).

han infectado las palabras, éstas no podrán servir para que los interlocutores se comuniquen racionalmente entre sí. Esto es en verdad lo que vemos que ocurre en todo el mundo (pensemos, por ejemplo, en Sudáfrica, donde presenté una versión anterior de este artículo). En esta situación, las teorías que critico no tienen utilidad alguna, puesto que la gente apelará a sus convicciones contrarias, sin que pueda siquiera iniciarse una discusión seria y fructífera.

En seguida debo aclarar el papel que desempeña la apelación a los hechos en la argumentación moral. Si pregunto “¿Debo alistarme en el ejército?”, lo primero que necesitamos es, como ya se ha dicho, que esté claro lo que quiero decir con “debo”. Pero eso no es suficiente. Debo tener claro lo que pregunto; pero otro aspecto importante de esto es qué es lo que implican las palabras “alistarse en el ejército”. En otras palabras, ¿qué haría si me alistara en el ejército? Hay algunos filósofos que usan el epíteto “consecuencialista” como denuesto en contra de sus oponentes. Ahora bien, estoy dispuesto a conceder que puede haber un sentido en que debemos hacer lo correcto y lamentar las consecuencias. Pero estos filósofos están realmente muy confundidos si creen que al decidir lo que debemos hacer podemos pasar por alto lo que haríamos si hiciéramos una u otra de las cosas que podemos hacer. Si alguien piensa que no debe alistarse en el ejército, o que estaría mal hacerlo, su razón (lo que hace malo al hecho) debe tener algo que ver con lo que él haría si se alistara en el ejército. Se trata del acto o de la serie de actos cuya moralidad nos preocupa. Alistarse en el ejército significa, en sus circunstancias (si tal es el régimen en que vive, como en Sudáfrica), comprometerse a matar gente en las calles si así lo ordena el gobierno. Esto es lo que entraña ser soldado en su situación presente. Cualquiera que piense que en este sentido las consecuencias no tienen que ver con las decisiones morales no puede haber entendido cuál es el objeto de la moralidad: la moralidad se ocupa de los actos; es decir, de lo que hacemos; y esto es lo nosotros hacemos que ocurra —la diferencia que introducimos en el curso de los acontecimientos. Estos son los hechos que tenemos que conocer.

Existen algunas teorías éticas, conocidas en general como teorías *naturalistas*, para las cuales los hechos son pertinentes para las decisiones morales de una manera muy directa. Afirman que lo que las palabras morales *significan* es algo factual. Para dar un ejemplo muy burdo: si “malo” *significara* “tal que pone en peligro al Estado”, entonces obviamente estaría mal hacer cualquier cosa que pusiera en peligro al Estado, y no deberíamos hacerla. El problema con tales teorías es, en efecto, el mismo que con las que mencionamos anteriormente. Si el autor incorpora a la teoría misma sus propias

convicciones morales, la teoría es inservible por lo que toca a la argumentación moral; y una de las formas de hacer esto es incorporar esas convicciones en los significados de las palabras morales. Esto hace imposible la comunicación y la argumentación racional entre personas de diferentes convicciones morales. Si "malo" significara lo que se acaba de sugerir, alguien que pensara que, desde el punto de vista moral, existen algunas cosas más importantes que la salvaguarda del Estado, no podría usar la palabra para discutir con un partidario del régimen. Uno de ellos tendría que alistarse en el ejército y el otro unirse a la revolución: no tendrían más remedio que luchar entre sí.

Yo no digo lo mismo que los naturalistas, puesto que la explicación que he dado del significado de "deber" en términos de prescriptividad y universalizabilidad *no* incorpora ninguna convicción moral sustancial; es *neutral* respecto de los dos participantes en tal disputa, y por lo tanto, si no me equivoco, ambos podrán usar las palabras al ventilar su desacuerdo. Así pues, tengo que dar, y en parte ya la he dado, una explicación diferente de cómo los hechos son pertinentes para las decisiones morales. Vaya esto a modo de explicación de la racionalidad misma —de la noción de razón.

Puede ser útil que empecemos con algo más sencillo que los argumentos morales: empecemos con los simples imperativos. No deben confundirse estos dos tipos de actos del habla, pues entre ellos hay diferencias importantes; pero los imperativos como "Alístate en el ejército" ilustran de manera mucho más sencilla lo que trató de explicar. Ellos son el tipo más simple de prescripción (los juicios morales son de un tipo mucho más complejo debido a la universalizabilidad). Para poner un ejemplo todavía más simple: supongamos que digo "Dame té" y no "Dame café". Digo esto en atención a *algo* que en ese preciso momento tiene que ver con tomar té o café. Esa es mi razón para decirlo. Si tomar té fuera de un carácter diferente, podría no haberlo dicho. Deseo, o elijo, tomar té y no café porque creo que así es como sería tomar té, es decir, debido a un hecho (supuesto) relativo a tal acto. Espero que, si está claro que, aun en el caso de los imperativos simples como éste, los hechos pueden ser razones para enunciarlos, estará igualmente claro que también los juicios morales pueden enunciarse debido a tales razones, aunque *en sí mismos* no sean (o no sólo sean) enunciados de hecho, sino de carácter prescriptivo. Sería irracional pedir té pasando absolutamente por alto lo que de hecho significa tomar té.⁵ Nótese que lo que acabo de decir no depende de ninguna manera de la universalizabilidad; deliberadamente

⁵ Véase mi "What Makes Choices Rational?", *Rev. Met.* 32 (1979).

he tomado el caso de los imperativos simples que *no* son universalizables. Más adelante introduciremos un elemento que depende de la universalizabilidad; pero a estas alturas no es necesario.

Ahora quisiera introducir otro elemento que tampoco depende de la universalizabilidad y que, ciertamente, es independiente de todo lo que hasta ahora he expuesto. También es un elemento lógico que depende de los significados de las palabras. Se refiere a la relación entre *saber* lo que es experimentar algo y experimentarlo. La importancia de esto para lo que hasta aquí he dicho es que, para conocer los hechos concernientes a lo que haríamos si hiciéramos algo, una de las cosas que tenemos que saber es lo que nosotros, o los demás, experimentaríamos en caso de hacerlo. Por ejemplo, si pensamos en azotar a alguien por diversión, es importante que lo que haríamos en caso de azotarlo le diera a él *esa* experiencia extremadamente desagradable. Si a esa persona no le importara, o incluso le gustara, nuestro acto sería diferente en un respecto moral importante. Así pues, es importante considerar cuáles son las condiciones para que se diga que en realidad conocemos cuál sería la experiencia (ya sea nuestra o de otra persona) resultante del acto que nos proponemos.

Supongamos que la experiencia en cuestión es (como en este caso) un *sufrimiento* de cierto tipo. Quisiera dejar sentado que no podemos sufrir sin saber que estamos sufriendo, ni saber que sufrimos sin sufrir. La relación entre tener experiencias y saber que las tenemos fue advertida ya por Aristóteles.⁶ Existen dos razones que apoyan la segunda mitad de la tesis que acabo de plantear. La primera es que no podemos saber cosa alguna sin que ésta se dé (éste es el tipo de palabra que “saber” es).^{*} La segunda razón es una razón particular y es más importante para nuestra argumentación: si no tuviéramos la experiencia del sufrimiento, no habría nada que conocer ni medio para conocerla. El conocimiento de que tenemos la experiencia de sufrir es un conocimiento *directo*, no algún tipo de conocimiento por inferencia, y por tanto no puede existir sin que el objeto de conocimiento (es decir, el sufrimiento) esté presente en nuestra experiencia. He aquí por qué es tan difícil saber cómo son los sufrimientos de otras personas. Como veremos, la imaginación tiene que tomar, de manera inadecuada, el lugar de la experiencia.

Además, no podemos sufrir sin preferir, *pro tanto*, hasta donde cabe hacerlo, que no sufriéramos. Si no prefiriéramos, en igualdad de circunstancias, que cesara, tal estado no sería sufrimiento. La preferencia de que cesara es lo que expresaría, si se expresara en el len-

⁶ Aristóteles, *Ética Nicomaquea*, 1170a 29.

^{*} Utilizamos en todo el artículo las palabras “conocer” o “saber” como traducciones de “*to know*”, guiándonos en cada caso por las exigencias del contexto.

guaje, la prescripción de que cese. Así pues, resumiendo, si estamos sufriendo, y por lo tanto sabemos que sufrimos, estamos destinados a aceptar la prescripción de que, en igualdad de circunstancias, tal estado cese. Tenemos que querer que cese. De otra suerte, el estado no sería de sufrimiento.

Dejemos, por ahora, nuestros actuales sufrimientos. Ahora tenemos que considerar lo que supone el conocimiento de que *experimentaremos*, o en ciertas circunstancias *experimentaríamos*, sufrimiento, o de que lo experimenta *otra persona*. Ocupémonos primero del último caso. ¿A qué me comprometo si en verdad afirmo que sé cómo sufre otra persona, o qué es para ella sufrir así? Me parece que la piedra de toque de esto es la siguiente pregunta: “¿Cuáles son mis preferencias (o, en otras palabras, qué prescripciones acepto) respecto de una situación en la cual se *me* fuera a poner de inmediato exactamente en *su* misma situación, sufriendo de la *misma* manera?”. Si ella sufre así, sabe que lo hace, y tiene la preferencia de que cese (una preferencia de cierta intensidad que dependerá de la gravedad del sufrimiento). Por lo tanto, ella acepta, con cierto grado de intensidad, la prescripción de que tal estado cese. Esta preferencia y esta aceptación forman parte de su situación, y por ende, parte de lo que tengo que imaginar que yo experimento si se me fuera a poner de inmediato en tal situación.

Acabo de preguntar: “¿A qué me comprometo si afirmo en verdad que sé cómo sufre la otra persona?” Mi tesis será que me comprometo a tener la preferencia de que, si se me fuera a poner en seguida en su situación con sus preferencias, el sufrimiento terminará; y la fuerza de esta preferencia que estoy comprometido a tener es la misma que la fuerza de su preferencia.

He dicho que ésa sería mi tesis. Pero aún no he argumentado en su apoyo, y estoy seguro de que el argumento será motivo de controversia. La jugada que voy a introducir es la siguiente: el que realmente piense que la persona que de inmediato sería puesta en esa situación soy *yo* mismo, depende de que yo me asocie con las preferencias que tendría tal persona (es decir, yo mismo), o las asuma como propias. Claro que, como antes, tendremos que añadir “en igualdad de circunstancias”; bien puede haber otras cosas que yo prefiera con tal intensidad que superen a mi preferencia de que se satisfaga la preferencia de aquella persona (la persona que me imagino ser).⁷ Pero, *en igualdad de circunstancias*, debo preferir que se satisfaga, con la misma intensidad de preferencia que tendría si me encontrara en esa situación con esas preferencias. Si no lo estoy, entonces o realmente

⁷ Véase mi “Some Reasoning about Preferences”, *Ethics* 95 (1984).

no sé qué es estar en tal situación con tales preferencias (en realidad no me la represento plenamente), o no pienso realmente que soy yo la persona que estaría en esa situación.

Permítaseme dar un ejemplo para ilustrar todo esto. Supóngase que alguien ha sido atado con una llanta en torno a su cuello, y que se prende fuego a ésta con petróleo. Esta persona sufre en extremo, y por lo tanto sabe que está sufriendo. ¿Qué es para *mí* saber lo que para ella es sufrir así? O, para plantearlo en función de las preferencias: ella prefiere sin duda que la dejaran de quemar de esa manera; ¿qué es para *mí* saber lo que es tener una preferencia como ésa respecto a ese desenlace, o de esa intensidad (decir, por ejemplo, “¡Ya no, ya no!” con ese grado de angustia)? Y bien, supongamos que afirmo saber precisamente lo que para él es tener esa preferencia y sufrir de esa forma; y luego supongamos que alguien ofrece hacerme a *mí* sin mayor demora exactamente lo mismo, y que yo digo “No me importa; me da exactamente igual”. Eso sin duda demostraría que yo no sé lo que es eso para la otra persona. Esto supone que si eso se me hubiera hecho a *mí*, yo tendría las mismas experiencias y preferencias que la persona a quien se le hace. Creo que lo mismo podría mostrarse con ejemplos menos dramáticos y que una vez que se han considerado las posibles variantes de las circunstancias (es decir, las preferencias en conflicto), podría mostrar convincentemente que la tesis se sostiene; pero no procederé a defenderla ahora, pues quisiera sacar conclusiones de ella en favor de mi argumento principal.

Primero permítaseme resumir la tesis que he planteado hasta ahora. Tenemos la prescriptividad y la universalizabilidad de los juicios morales, las cuales, sostengo, pueden establecerse mediante argumentos basados en el significado de las palabras —argumentos lógicos. Luego tenemos la necesidad de una información factual correcta, si es que hemos de asentir racionalmente a las prescripciones, incluidas las prescripciones expresadas mediante “deber”. Por último tenemos la tesis de que *no* estamos en posesión de información factual correcta sobre el sufrimiento de otra persona, o en general sobre sus preferencias, a menos que nosotros mismos tengamos preferencias de que, de estar *nosotros* en su situación con sus preferencias, éstas se satisficieran. Nótese también que, aunque he afirmado poder establecer la universalizabilidad, aún no la he usado en mi argumento. Y eso es lo que ahora voy a hacer, conjuntamente con las otras tesis.

Supongamos que soy yo el que hace sufrir a la víctima. Ella desea con toda su alma que la desate. Es decir, ella apoya, con una fuerza de asentimiento muy elevada, la prescripción de que yo la desate. Podemos decir ya, sin traer a colación la universalizabilidad y con

apoyo en la fuerza de nuestras tesis previas, que, si sé qué es para ella estar en ese estado (y si yo no lo sé, entonces mi juicio moral adolece de falta de información), entonces debo preferir que me desaten en caso de encontrarme en tal estado. Es decir, debo prescribir que, en tales circunstancias hipotéticas, me desaten. Supongamos ahora que me pregunto a cuál prescripción *universal* estoy dispuesto a asentir en lo concerniente a mi conducta actual que está haciendo sufrir a esa persona; es decir, qué estoy dispuesto a decir que *debo* hacer ahora respecto de esa persona. El “debo” expresa aquí una prescripción universal, de modo que si digo “No debo desatarlo”, me comprometo a avalar la prescripción de que en circunstancias similares no se me desatara.

Naturalmente puedo decir que no estoy dispuesto a asentir a *ninguna* prescripción universal. Esa es la posición de la persona a quien en otro lugar llamé *amoralista*, e indique cómo trataría con ella.⁸ Pero supongamos que no soy amoralista, y por lo tanto estoy dispuesto a aceptar alguna prescripción universal tocante a personas que se encuentren precisamente en la presente situación. La cuestión es cuál será esta prescripción. Si universalizo la prescripción de seguir haciendo sufrir a alguien, ello presupone prescribir que si alguien estuviera haciéndome sufrir precisamente en una situación análoga, entonces ese alguien siguiera haciéndome sufrir. Pero eso contravendría la prescripción a la cual, como ya hemos visto, yo debería asentir si supiera qué es encontrarse en la situación de mi víctima: la prescripción de que si yo estuviera en esa situación *no* se me mantuviera en ella, sino que se me desatara. Así pues, me hallo en el predicamento que Kant llamaba contradicción en la voluntad.⁹

¿Cómo resolver esta contradicción? La respuesta se torna evidente si nos percatamos de que lo que sucede (lo que tiene que suceder si pretendo universalizar mis prescripciones) es que estoy obligado a tratar las preferencias de otros como si fueran mías. Ésta no es más que otra forma de plantear el requisito de universalizar mis prescripciones. Pero si en esta situación las dos preferencias que han entrado en contradicción fueran mías, lo que yo haría es dejar que la más fuerte de ellas se antepusiera a la más débil. Y eso es lo que me veo obligado a hacer en este caso en el que, a resultas del intento de universalizar, he llegado a la situación de encontrarme con dos preferencias o prescripciones contradictorias entre sí, en cuanto a lo que se me hiciera en la situación hipotética de encontrarme en los zapatos de otra persona. Así que, si la preferencia de mi víctima de que yo

⁸ Véase *MT*, pp. 182 ss.

⁹ Kant, *Grundlegung zur Metaphysik der Sitten*, parte media de la sección 2.

desista de atormentarla es más fuerte que mi propia preferencia de no desistir (como ciertamente lo será), la respuesta será: "¡desistir!".

Así pues, en este simple caso bilateral en que sólo intervienen dos personas, hemos llegado a lo que esencialmente es la respuesta utilitarista a nuestro problema moral; y hemos llegado a ella por una vía kantiana. La gente habla como si Kant y los utilitaristas se encontraran en polos opuestos de la filosofía moral; pero esto no hace más que demostrar cuán poco han entendido tanto a los utilitaristas como a Kant. Nos vemos llevados a ponderar las preferencias de todas las partes afectadas (en este caso dos) en proporción a sus fuerzas, y a decir que debemos actuar en consideración a la más fuerte. Yo podría, de haber espacio, mostrar cómo, generalizando este argumento hasta que cubriera situaciones multilaterales en que se afectarían las preferencias de muchas personas, deberíamos adoptar también las respuestas utilitarias, a saber, que en cada caso deberíamos actuar de manera que se satisficieran al máximo las preferencias de todas las partes afectadas considerándolas imparcialmente. Pero ahora no haré tal intento; ya lo he hecho en otro lugar,¹⁰ y debo pasar a explicar cómo va a funcionar esta forma de pensar en el transcurso de nuestra vida moral real, cuando tenemos que decidir sobre asuntos prácticos. Tendré que hacer esto sólo en términos muy generales.

Podría pensarse que hemos llegado a un tipo de utilitarismo de los actos; y esto es verdad. Pero no es el tipo de utilitarismo de los actos frente al cual se enseñen objeciones de cajón a quienes empiezan a estudiar filosofía. En seguida trataré de explicar en qué difiere del tipo burdo de utilitarismo de los actos el utilitarismo que ahora propongo. En rigor, la diferencia no es de carácter teórico. Más bien se deriva de una consideración del verdadero predicamento humano al que nos enfrentamos al pensar moralmente. Para apreciar esto, pensemos qué pasaría si no tuviéramos limitaciones humanas. Es decir, supongamos que nuestro conocimiento y claridad de pensamiento fueran infinitos y no fuéramos parciales en favor de nosotros mismos ni adoleciéramos de otras debilidades humanas. En otro lugar he dado el nombre de *arcángel* al ser que posee estas virtudes sobrehumanas.¹¹ Él realmente podría pensar de manera utilitaria. Pero sería desastroso, por razones obvias, que nosotros los humanos tratáramos de hacer lo mismo. En primer lugar, carecemos casi siempre de la información necesaria; en particular, somos muy malos para ponernos en el lugar de otra persona e imaginarnos qué sería

¹⁰ Véase mi *Freedom and Reason*, Oxford U. P., 1963, pp. 115 ss.

¹¹ Véase *MT*, cap. 3.

ser ella. En segundo lugar, no tenemos el tiempo para adquirir esta información y pensar en ella. Y finalmente, carecemos de la capacidad para pensar con claridad. Estas tres limitaciones nos llevan muy fácilmente a suponer que determinado acto nuestro probablemente sea el que tenga las mejores consecuencias (es decir, el que satisfará al máximo e imparcialmente las preferencias), cuando en realidad es nuestro propio interés el que nos lo aconseja. A cada paso vemos sucederse este tipo de demandas de consideración especial.

Supongamos que, conscientes de estas limitaciones, acudiéramos a un arcángel en busca de consejo, no sobre una situación particular (pues no siempre tendremos acceso a él, y en consecuencia querríamos que nos aconsejara para lo futuro), sino sobre cómo, en general, reducir al mínimo sus efectos negativos. La gente cree que puede recurrir de esta forma a Dios, aunque lo que afirman que Él les dice varía de una persona a otra. Pero supongamos que *tuviéramos* acceso inmediato a un ser supremo o al menos superior que nos pudiera aconsejar. Él señalaría que lo mejor que podemos hacer es, en cada oportunidad, procurar que la expectativa de satisfacción de preferencias que resulte de nuestras acciones sea lo más grande posible. Estoy seguro de que esto es lo que Dios haría, puesto que ama a sus criaturas y quiere que hagamos por ellas lo mejor que podamos.

La expectativa de satisfacción de preferencias (dicho brevemente, la expectativa de utilidad) es la suma de los productos de la utilidad y la probabilidad del resultado para todos los varios resultados posibles de la acción. A esto me refiero con la expresión "Actuar en vista de las mejores consecuencias". El problema es "¿Cómo lograr esto?" Dadas nuestras limitaciones, no lo lograremos haciendo en cada ocasión un cálculo utilitarista o un análisis de costo-beneficio. Antes bien, el arcángel nos dirá que cultivemos en nosotros mismos un conjunto de disposiciones o principios, junto con las actitudes o sentimientos o (si alguien quisiera emplear este término) intuiciones que los acompañan: un conjunto cuyo cultivo, tomado globalmente, es más probable que nos lleve a elevar al máximo la satisfacción de las preferencias. El arcángel, que puede obtener en cada ocasión particular la respuesta correcta, puede actuar mejor que nosotros; pero eso es lo mejor que podemos hacer.

Se verá que esto, aunque en cierto sentido es una forma de utilitarismo de las reglas, no es incompatible con el utilitarismo de los actos. Pues lo que el arcángel nos aconseja hacer es realizar ciertos actos, a saber, actos de cultivar disposiciones, y su razón para aconsejar esto es que tales actos son los que más probablemente lleven a actuar de la mejor manera, que es exactamente lo que aconsejaría un utilitarista de los actos.

Sin embargo, esta versión del utilitarismo garantiza las ventajas que formas más antiguas de utilitarismo de las reglas reclamaban para sí; en particular, garantiza la ventaja de inmunizar al sistema que proponemos contra las objeciones basadas en la contraintuitividad de sus consecuencias. Pues las intuiciones que nos aconseja cultivar el arcángel utilitarista de los actos son las *mismas* que aquellas a las que apelan sus detractores.

El efecto de este recurso es dividir en dos niveles el pensamiento moral (además del tercer nivel, o el nivel metaético, que se ocupa, no del pensamiento moral sustancial, sino de la forma del pensamiento moral, es decir, de la lógica del lenguaje moral). A estos dos niveles los llamo nivel *intuitivo* y nivel *crítico*. Si seguimos el consejo del arcángel, casi todo nuestro pensamiento moral se dará en el nivel intuitivo; de hecho, casi todo el tiempo nos comportaremos exactamente como los intuicionistas dicen que lo hacemos y deberíamos hacerlo. Con todo, la diferencia estribará en que, toda vez que —como todos podemos darnos cuenta— los buenos principios, disposiciones y actitudes que cultivamos correctamente son hasta cierto punto generales y simples e inespecíficos (si no lo fueran, serían inmanejables, inútiles e inenseñables), entrarán en conflicto en los casos particulares difíciles; y entonces, a diferencia de los intuicionistas, sabremos lo que debemos tratar de hacer, por difícil y peligroso que sea. Como de hecho no contamos con arcángeles que nos auxilien, tenemos que hacer todo lo posible para pensar críticamente como arcángeles en esas ocasiones problemáticas. Pero cuando nuestras intuiciones nos den una guía clara, las seguiremos —al menos eso es lo que nos aconsejaría nuestro arcángel utilitarista en la única entrevista que nos concediera en la vida.

Pero, se nos objetará, esto presupone que nuestras intuiciones son las correctas. Así es, en efecto. Ello nos da una razón más para emplear el pensamiento crítico. Es peligroso emplearlo durante una crisis; pero cuando ésta termina, o antes de que ocurra, puede ser esencial. De otra forma, ¿cómo podríamos tener la menor confianza en que las intuiciones que acertamos a desarrollar en el curso de nuestra vida son las mejores? Por ejemplo, las intuiciones respecto a cómo los blancos deben tratar a los negros, o los hombres a las mujeres. De esta suerte, lo que el arcángel prudente recomiende, y lo que practiquen los educadores y autoeducadores prudentes, será una mezcla juiciosa de pensamiento intuitivo y crítico, cada uno empleado en la ocasión apropiada. Y esto es lo que hacen ya las personas prudentes.