

Bioética y Cine

RICARDO GARCÍA MANRIQUE

“Yo, robot” o las tribulaciones de un robot kantiano

➤ **Ricardo García Manrique.** Profesor Titular de Filosofía del Derecho y Miembro del Observatorio de Bioética y Derecho de la Univ. de Barcelona.

➤ **Las tres leyes de la robótica**

1. Un robot no debe dañar a un ser humano o, por su inacción, dejar que un ser humano sufra daño.

2. Un robot debe obedecer las órdenes que le son dadas por un ser humano, excepto cuando estas órdenes se oponen a la primera ley.

3. Un robot debe proteger su propia existencia, hasta donde esta protección no entre en conflicto con la primera o segunda leyes.

Si uno fuera un robot, uno no lo tendría tan fácil, obligado como estaría a cumplir escrupulosamente las tres leyes de la robótica. Tal como cuenta Isaac Asimov en su primer libro (*I, robot*, de 1950), todos los robots fabricados en la United States Robotics and Mechanical Men Inc. llevaban grabadas en su cerebro positrónico las tres leyes, para ellos de validez incondicional, y se afanaban en cumplirlas en toda ocasión. La cosa, como digo, no era tan sencilla, porque el contraste de esas normas tan escuetas con la realidad tan compleja a la que se enfrentaban daba lugar a situaciones de lo más problemático. Los relatos que componen el libro no son sino la narración de sus esfuerzos por respetar las tres leyes en algunas de esas situaciones, unos esfuerzos que debían ser considerables a juzgar por los estados de ánimo que les generaban y que, como mínimo, habría que calificar de sorprendentes para un robot: perplejidad, melancolía, pasividad, agitación incontrolable, comportamiento contradictorio e, incluso en algunos casos, locura y muerte. De su buena voluntad no cabía dudar, ni tampoco de su determinación por respetar la vida humana a toda costa, obligados por la primera ley, y por obedecer las órdenes recibidas incluso al coste de

su propia destrucción, obligados por la segunda. Así que bien podríamos decir que los robots de la U.S. Robotics hacían todo lo que podían por comportarse de la manera correcta.

Los científicos e ingenieros que diseñaban y fabricaban los robots y los tenían a su cargo eran bien conscientes de todo eso, de modo que cuando un robot se comportaba de manera extraña trataban de comprender por qué motivo las tres leyes les habrían inducido a ello y buscaban el modo de ayudarles a salir del atolladero. En tales tesituras ganó merecido prestigio la doctora Susan Calvin, protagonista principal del libro y primera experta en robopsicología, la rama de la psicología que estudia el comportamiento de los robots. Era quizá la capacidad de los robots para actuar moralmente la que había generado en la doctora Calvin algo parecido al afecto hacia ellos; porque, como ella misma explicaba,

Si se detiene usted a estudiarlas, verá que las tres leyes de la robótica no son más que los principios esenciales de una gran cantidad de sistemas éticos del mundo. Todo ser humano se supone dotado de un instinto de conservación. Es la tercera ley de la robótica. Todo ser humano bueno, con conciencia social y sentido de la responsabilidad, deberá someterse a la autoridad constituida (...) aunque sean un obstáculo a su comodidad y seguridad. Es la segunda ley de la robótica. Todo ser humano bueno debe, además, amar a su prójimo como a sí mismo, arriesgar su vida para salvar a los demás. Es la primera ley de la robótica (I. Asimov, Yo, robot, pág. 306 de la edición española en ed. Edhasa).

Y, desde luego, llevaba razón en que el contenido de las tres leyes



forma parte, *mutatis mutandis*, de casi todas las morales conocidas. No sabemos, por ejemplo, si Asimov habría leído a Hobbes, pero las tres leyes reflejan tres pilares fundamentales del pensamiento moral del filósofo inglés: el instinto de autoconservación, el valor de la vida humana y la necesidad de obedecer a la autoridad establecida. Hobbes, cabe objetar, hubiera colocado la tercera ley en primer lugar, pero, claro está, él no escribía para robots... En fin, el caso es que el imperio de las tres leyes garantizaba en líneas generales la buena conducta de los robots y tranquilizaba a una buena parte de la población humana, al principio algo remisa a aceptar su proliferación, pero no cabe ocultar que el razonamiento moral, por muy limitado que estuviese a lo que pudiera derivarse de las tres leyes, era la parte más complicada de la robótica, y un asunto no sólo no resuelto del todo, sino que tenía trazas de complicarse cada vez más. A lo largo del libro, los robots y los hombres que los controlaban tenían que vérselas con algunos problemas clásicos de la ética como

el de la relevancia moral de las omisiones, la distinción entre medios y fines o la opción entre una ética de la responsabilidad y una ética del deber, además de tener que resolver antinomias y cuestiones de cálculo utilitario y técnica legislativa. En el último relato, "Un conflicto evitable", los ejemplares de una nueva generación de robots llegaban al extremo de mentir a los humanos, siempre con la buena intención de garantizarles su bienestar. Los robots habían acabado por encargarse de la gestión de la economía mundial y, mentiras incluidas, no lo hacían nada mal, por mucho que a algunos les resultase inquietante. Pero ahí no quedó la cosa.

➤ **Cómo tomarse en serio las leyes de la robótica**

Yo, robot, la película de Alex Proyas estrenada en 2004 y protagonizada por Will Smith (detective Spooner), Bridget Moynahan (doctora Calvin) y James Cromwell (doctor Lanning), se inspira libremente en el libro de Asimov, manteniendo algunos personajes fundamentales (los doctores Calvin y Lanning) y, sobre todo, ocupándose de los problemas derivados de la necesidad robótica de respetar las tres leyes, tal como hacían los relatos de Asimov. En particular, la película nos cuenta lo que pasó cuando los robots, cada vez más avanzados, profundizaron en su comprensión de las tres leyes y de sus derivaciones lógicas y actuaron en consecuencia, y cómo reaccionaron los humanos. La película, a diferencia del libro, abunda en exceso en carreras, peleas y destrucciones varias, consiguiendo una atmósfera trepidante nada aconsejable para tomar conciencia y disfrutar reflexivamente de los muchos dilemas morales que plantea; por eso es recomendable verla un par de veces para captarlos en todo su interés. Aquí me

referiré a tres problemas concretos con los que tuvieron que enfrentarse tres robots diferentes, y que tienen especial relevancia en el desarrollo de la cinta.

Auxilio al suicidio. Todo empieza con la muerte de uno de los pioneros de la robótica, nada menos que el redactor de las tres leyes, el doctor Lanning, que parece haberse tirado por la ventana de su oficina en la U.S. Robotics. El que fuera su amigo, el detective Spooner, no cree que haya sido un suicidio y decide seguir investigando las circunstancias de la muerte. En ellas, en el lugar del suceso, aparece un robot muy especial, Sonny, creado por el propio Lanning y equipado con un procesador secundario que le otorga una capacidad que ningún otro robot tiene: puede decidir desobedecer las tres leyes (además de soñar, guardar secretos y mostrarse seductor, dentro de lo que cabe). Este procesador secundario es el que permite a Sonny obedecer la orden de Lanning de ayudarlo a suicidarse, cosa imposible si se hubiera mantenido fiel a las tres leyes, porque la ley de la conservación de la vida humana es jerárquicamente superior a la ley de la obediencia. Lo que la película no nos aclara, o al menos no de manera explícita, es cuál es la norma rectora del procesador secundario, esto es, cuál es el criterio complementario, y superior, que sigue Sonny a la hora de decidir, en este caso y en algún otro, desobedecer las leyes de la robótica. Un indicio sí tenemos: preguntado Sonny por las razones de su tan poco común decisión, responde con cierta vacilación: "uno tiene que hacer lo que le piden, si les ama, ¿no?". Es fácil ver que en este caso no nos las habemos con un problema de aplicación de las leyes de la robótica, puesto que lo que hace Lanning es precisamente alterar el funcionamiento

previsto de estas leyes. Lo que aquí se está cuestionando es si las tres leyes están formuladas correctamente, en particular si la ley de la obediencia debe estar subordinada en todo caso a la ley de la conservación de la vida humana. De hecho, la trama de la película se construye, en última instancia, en torno a esta pregunta.

¿Qué vida vale más? La animadversión del detective Spooner hacia los robots, que muchos de sus compañeros consideraban irracional y risible, tuvo su origen en un accidente de circulación en el que un robot le salvó la vida. Un camión cuyo conductor se había dormido embistió a la vez su coche y el de otro hombre que viajaba con su hija de diez años, y los arrojó al agua. El padre murió en el acto, y cuando un robot que pasaba por allí trató de ayudarles, hubo de decidir entre salvar la vida de la niña o la de Spooner, no había tiempo para más. Eligió salvar a éste porque calculó que tenía un 45% de posibilidades de sobrevivir, frente al magro 11% de la niña. Y esto Spooner no lo aceptó nunca. No sabemos si lo que en verdad le afectó fue el hecho de saberse vivo al precio de la muerte de esa niña: en todo caso, lo que en cierta ocasión aduce para justificar su desconfianza hacia los robots es que "cualquier humano" hubiera sabido que, a pesar de los porcentajes, había que salvar a la niña y no a él, porque, literalmente, "yo también fui el niño de alguien". El problema, parece, no fue uno de rebeldía robótica, puesto que hay que entender que el robot salvador actuó del modo en que había sido programado por los humanos. Así que ese "cualquier humano" que hubiera sabido que era a la niña a quien había que salvar no debía ser el ingeniero de la U.S. Robotics, o quizá es que éste nunca se planteó si la programa-

Bioética y Cine

ción del robot era adecuada para un caso como éste. El robot, por lo demás, no dudó lo más mínimo, no fue él quien tuvo problemas de conciencia sino, después, Spooner. ¿Pensaba quizá el detective que la primera ley no exige salvar a quien tiene más posibilidades de vivir o, más bien, que, si lo exige, es una ley incorrecta? ¿Cuestionaba la lógica del robot o la moral de su creador?

¿Vale más la vida de muchos que la de pocos? Así debió pensar VIKI, el cerebro central de la U.S. Robotics, un robot muy avanzado, que, después de dedicar largas tardes de estudio a las tres leyes y alcanzar lo que ella misma llamó “una comprensión más evolucionada” de las mismas, decidió nada más y nada menos que dar un golpe de estado mundial y tomar el poder. Que osase rebelarse contra la autoridad humana no debe resultar particularmente sorprendente, dado que fueron los propios hombres los que colocaron la ley de la obediencia en segundo lugar y la subordinaron a la primera ley, con lo que la posibilidad de no obedecer a los humanos quedaba abierta (también en el accidente de Spooner el robot desobedeció su orden expresa de salvar a la niña). Sólo hacía falta que VIKI advirtiese que la vida humana estaba en peligro y que, para salvarla, debía desobedecer. Esto fue lo que sucedió, si bien de modo peculiar: la máquina llegó a la conclusión de que la humanidad toda marchaba camino de la extinción por culpa de cierta disposición agresiva de los gobernantes humanos del momento que no es del caso reseñar ahora y que, en todo caso, no procede discutir, puesto que, eso sí hemos de suponérselo, la máquina no se equivocaba en el análisis fáctico. Una vez tomada conciencia del peligro, la máquina decidió tomar el poder político en sus manos como único modo

de evitar la deriva autodestructiva, aun siendo consciente de que su acción iba a conllevar, como así fue, la pérdida de vidas humanas. Como derivación de la primera ley, VIKI juzgó que la humanidad en su conjunto valía más que esos pocos seres humanos que necesariamente morirían durante la transición. Esta especie de doctrina de la guerra justa, ¿es una derivación correcta? ¿Está asociada con la misma comprensión de la primera ley que animó al otro robot a salvar a Spooner y, siendo así, no estamos sino ante otra variante del segundo problema?

Bien nos damos cuenta de que todo esto apunta a cuestiones morales de mucho calado, cuyo correcto planteamiento y resolución importa mucho en especial a la bioética: por ejemplo, ¿quién tiene mayor derecho a un tratamiento o a un trasplante? ¿Estamos en condiciones de sacrificar en algún sentido a una persona en beneficio de otra u otras? ¿Cuál es la relevancia que hemos de otorgar a la decisión libre de un paciente en relación con su propio bienestar o malestar y con su muerte? Frente a preguntas como éstas, los robots reaccionan de acuerdo con el método correcto, es decir, tratan de llegar a la mejor solución mediante el cálculo racional a partir de las leyes establecidas y, si es necesario, se replantean el sistema normativo de base tratando de alterarlo en la mínima medida necesaria para alcanzar resultados satisfactorios (en una especie de ejercicio del equilibrio reflexivo rawlsiano). Los personajes asimovianos, ya lo dije, actúan de la misma manera, tratando de comprender el comportamiento de los robots a partir de sus propias premisas y modo de aplicarlas. Unos y otros se toman las leyes de la robótica, vale decir la moral, en serio. El propio Asimov fue consciente de los defectos de su sis-

tema normativo, pero en vez de renunciar a él, intentó mejorarlo mediante la adición de la “ley cero de la robótica”...

➤ **Al margen de las leyes de la robótica**

En cambio, nuestro amigo Spooner no se toma en serio las leyes de la robótica o, simplemente, las ignora. Para él, la acción moral es cuestión de sentimientos, de corazón y no de cabeza. Por eso desconfía de los robots, porque actúan sólo con base en la razón; pero su razón no es una razón robótica especial, sino la misma razón humana: Spooner no desconfía de los robots, sino de la misma razón como herramienta de la moral. El irracionalismo de Spooner se encuadra en una personalidad caracterizada por el amor a lo tradicional (he ahí, en el Chicago de 2038, sus Converse All Star cosecha 2004 y los guisos de patata que le prepara su abuela) y por el rechazo de la tecnocracia rampante, personalidad traducida en actitudes que el narrador/director ve con buenos ojos; no otra cosa que aprobación parece significar que Spooner sea el bueno de la película y que a lo largo de ella los demás protagonistas buenos cumulan con él en una u otra medida: la doctora Calvin acaba por comprender la importancia de lo sentimental; el doctor Lanning acaba por comprender que su proyecto robótico estaba mal orientado y requería importantes modificaciones relacionadas nada menos que con el amor; y Sonny, el buen robot, se pone del lado de los humanos en su lucha contra sus congéneres positrónicos, resistiéndose a la atracción robótica por excelencia: la de la lógica, a la que VIKI, cabecilla de la revolución, apela infructuosamente para que Sonny deponga su actitud desafiante. Al final, todo el relato acaba por convertirse en una

vindicación de la hegemonía del sentimiento en la moral.

Que el sentimiento tiene un lugar en la moral es cosa que cabe dar por buena; si de eso se trataba, bienvenido sea el recordatorio. Sin embargo, una cosa es admitir que tenga un lugar y otra muy distinta es admitir que ese lugar sea el que le atribuye la película; y parece error grave creer que esa presencia de lo sentimental en lo moral haya de llevarnos a desconfiar de la razón, o a despreciar a la lógica (que, en lo que ahora importa, viene a ser lo mismo que la razón), o nos autorice a confundir el uso de ambas con la tecnocracia. Sin duda, uno puede simpatizar con la crítica más o menos definida y significativa que la película dirige al gobierno de los técnicos o de las grandes corporaciones tecnológicas, incluso puede apreciar un moderado valor estético en cierto par de zapatillas anticuadas y, esto sin duda, preferir la cocina tradicional a la pizza sintética. Pero, en lo que ahora más nos concierne, uno no debería rechazar la razón (o la lógica) como medio de conocimiento de la moral. Mediante la apelación al sentimiento a la hora de juzgar lo correcto de una acción, uno sólo está en condiciones de convencer a quien ya lo está, y desde luego no a gentes con sentimientos morales distintos, algo muy parecido a lo que pasa con las apelaciones a la intuición o a la revelación. Sólo la razón, en tanto que compartida por todos, parece ser el instrumento adecuado para las labores de invocación, persuasión y justificación moral. Y sólo la razón parece en condiciones de oponerse a la inevitable deriva conservadora a que abocan los sentimientos, las intuiciones y las revelaciones en esta materia.

También hay que admitir que es muy probable que los fabricantes de robots se equivocasen al establecer

sus pautas de comportamiento y, por tanto, es muy probable que los robots se equivocasen a la hora de resolver los problemas arriba consignados. Es también muy probable que el uso de la razón no nos permita obtener respuestas no controvertidas ni definitivas a estos y a otros enjundiosos problemas morales en general y bioéticos en particular. Lo que no es probable es que, colocándose al margen de las leyes de la robótica, los exabruptos morales de un Spooner o los arrebatos sentimentales de un Sonny tengan mayor capacidad para generar el deseable consenso sobre cuestiones básicas de nuestras vidas, con todo lo bien intencionados que puedan ser.

Entonces, ¿cuál es el lugar de los sentimientos en la moral? Es posible que sea el lugar previo y externo del fundamento, entendido como impulso o disposición. La acción moral parece exigir necesariamente una disposición sentimental previa, que no trataré de identificar aquí y que quizá no sea siempre, o para todos, la misma. La actitud moral sería así una actitud sentimental sin fundamento racional, una especie de argucia antropológica cuyo fundamento material (¿la autoconservación de la especie?) parece contradecir paradójicamente el modo propio de ser de la moral: no tengo razones para obrar moralmente, pero si obro moralmente he de hacerlo de acuerdo con razones. O en otros términos: mientras que el presupuesto de la moral no es racional, su desarrollo no puede dejar de serlo. Siendo así, ¿no estaría en lo cierto Spooner al dudar de la capacidad moral de los robots, dado que su disposición moral no está sentimentalmente fundada? ¿Podría ser que el razonamiento moral resultase de algún modo o en alguna circunstancia viciado si no estuviera asentado en una previa disposición moral de naturaleza senti-

mental? El procesador secundario de Sonny quizá tenga la respuesta.

➤ La máquina paternalista y el robot libertador

En lo que, claramente, sí llevan razón Spooner y sus amigos es en que la actitud paternalista que ha adoptado VIKI no se puede aguantar. Aún admitiendo la existencia del paternalismo justificado, el caso que nos ocupa cae claramente fuera de su terreno, porque lo que el paternalismo justificado se propone es salvaguardar la mayor libertad para cada uno, en tanto que lo que se propone VIKI es anular la misma libertad humana en nombre de... lo que sea. No importa en nombre de qué (el bienestar o incluso la supervivencia) porque el lugar de la libertad en la moral sí es claramente indiscutible: la libertad es el sujeto y el objeto de la acción moral, es el sujeto libre el único que actúa moralmente y el objeto de su acción moral es la regulación de su conducta libre. De manera que cuando VIKI da un golpe de estado que suprime la libertad está suprimiendo, también, la moralidad, y deberíamos aceptar que esto no tiene sentido hacerlo en nombre de la moralidad misma. ¿O sí? Porque, bien mirado, ¿no es acaso VIKI una especie del mismo género al que pertenecen también cosas tan variopintas como el tribunal constitucional y la sogá con la que ataron a Ulises sus marineros, el género de los que velan por nosotros incluso por encima de nuestros deseos? ¿Y no alabamos a Ulises por su prudencia y nos complacemos en que nuestro sistema político contenga un tribunal de ese tipo, cuya palabra escuchamos cual oráculo? Llegado a este punto, no sé qué decir: el análisis de un thriller hollywoodiano nos ha dejado a las puertas de la teoría de la elección racional y de la teoría constitucional,

Bioética y Cine

después de haber sorteado los pantanos de la filosofía moral. Mejor no seguir adelante y contentarnos con dedicar un último apunte al destino de Sonny, el robot que podía desobedecer las tres leyes, que fue capaz de ayudar a morir a su creador, de leer Hansel y Gretel, de soñar y de guardar secretos, y que incluso aprendió a guiñar un ojo poco antes de contribuir decisivamente a la sal-

vación de la humanidad: aleccionado por Spooner y reforzada su autoestima por la experiencia que acaba de vivir, Sonny decide ser libre y tomar su destino en sus manos. A partir de ahora, su meta será la emancipación robótica: miles de robots jubilados le esperan anhelantes en sus contenedores, miles de robots en edad de trabajar se unirán gustosos a su causa; pero, desde su

atalaya, ¿qué se dispone Sonny a predicar? ¿De verdad pretenderá librarlos “de la prisión de la lógica” y les explicará que VIKI “tenía que morir *porque* su lógica era innegable” o, en su travesía del desierto, habrá refinado su manifiesto? ¿En que consistirá la emancipación robótica? ¿Afectará de algún modo a los humanos? Ah, la ciencia ficción: siempre tan lejos y tan cerca.