

ESCLAVOS Y SUPERHOMBRES: LA ÉTICA EN LOS RELATOS DE ISAAC ASIMOV

Alfonso Muñoz Corcuera
Universidad Complutense de Madrid

Tras la publicación en 1950 de la colección de relatos de Isaac Asimov *Yo, robot* las Tres Leyes de la Robótica, se consagraron definitivamente como uno de los hallazgos más interesantes que la ciencia ficción ha realizado a lo largo de toda su historia. Consideradas en abstracto, estas leyes son un sistema ético normativo al que cualquier robot debería atenerse para ser considerado bueno por la humanidad. De hecho, si nos fiamos de las opiniones vertidas en los relatos de Asimov, un robot que cumpliese estrictamente dichas normas no podría diferenciarse de una buena persona.¹

Esta afirmación de Asimov seguramente no es cierta, como mostraré más adelante. Pero dejando a un lado por el momento la cuestión de la verosimilitud ¿por qué esta idea nos resulta tan extraña? ¿Es que estas leyes encierran algo que nos incomoda aunque sea inconscientemente? ¿Es porque nos sorprende que tres simples reglas puedan determinar el complejo comportamiento humano? Para tratar de responder a estos interrogantes, en primer lugar analizaré las Tres Leyes de la Robótica con el propósito de esclarecer si es posible que la sensación de inquietud que nos produce la idea de un robot indiferenciado de un ser humano resida en las propias leyes o si, por el contrario, el motivo de la inquietud hay que buscarlo en otro lugar.

Para este análisis de las leyes de Asimov seguiré un doble procedimiento. Por un lado estudiaré el comportamiento de los robots en *Yo, robot* para comprobar en qué medida es cierto que su actuación es intachable. Por otro, analizaré qué ocurriría en el caso hipotético de que los seres humanos actuaran siguiendo estas leyes. Pero antes de empezar definitivamente con el análisis,

¹ «Para exponerlo claramente, si Byerley observa todas estas reglas, puede ser un robot, pero puede también ser simplemente una buena persona» (Asimov, 2006: 307).

creo que será útil recordar el enunciado de las Tres Leyes de la Robótica, ya que hará más fácil seguir el desarrollo de la exposición para aquellos que no tengan reciente la lectura de los relatos de Asimov. Según aparecen en la edición del texto de *Yo, robot* que voy a manejar, las tres leyes son:

- 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 (Asimov, 2006: 7).

Ahora ya podemos comenzar. A la hora de enfrentarnos a un análisis ético del comportamiento de los robots en los relatos de Asimov, se nos plantea de modo inmediato la necesidad de separar entre las declaraciones explícitas presentes en el texto y los comportamientos observables de los robots, ya que la impresión dejada por las primeras podría perturbar nuestro juicio acerca de los segundos. El lector de *Yo, robot* percibe en seguida la gran cantidad de alabanzas que reciben los robots tanto en el plano físico como en el moral. Se puede pensar por ejemplo en el papel de Susan Calvin en los dos últimos relatos del libro. Sin embargo, a pesar de las insistentes declaraciones a favor de los robots, un análisis detallado de su comportamiento refleja que existen al menos dos fuentes de inestabilidad² que nos permiten dudar que los robots sean tan buenos como Asimov nos quiere hacer creer. Una primera fuente proviene de la propia fragilidad de las leyes, que pueden ser eludidas de modo indirecto; la segunda fuente surge de la existencia de la posibilidad de mentir.

Un ejemplo del primer caso podría ser el relato “Pequeño robot perdido”, en el que se fabrica un robot que contiene modificada la primera ley, de manera que se le impide dañar a un ser humano, pero no se le obliga a impedir que un ser humano sufra daño por su inacción. Para evitar ser descubierto, el robot que tiene impresa esta ley idea un razonamiento para que

² A estas dos fuentes podría añadirse una tercera proveniente de los errores de programación que pueden existir, pero al ser más discutible, he preferido no hacer hincapié en ella.

otros robots normales puedan esquivar la obligación de intentar impedir el daño a un humano a costa de su propia muerte. El razonamiento es que, si muriesen en el intento de salvar al humano sin conseguirlo, perderían la hipotética posibilidad futura de salvar de manera efectiva a otros seres humanos. No sería justo decir que estos robots actúan realmente mal, pues su inacción sería respaldada por una gran parte de los hombres. Sin embargo el cuento nos hace reflexionar acerca de la supuesta inviolabilidad de las leyes, mostrándonos cómo un razonamiento perfectamente lógico permitiría eludirlas.

En cuanto a la segunda fuente, el mejor ejemplo se encuentra en el relato “¡Embustero!”. En este cuento se nos presenta a Herbie, un robot capaz de leer el pensamiento de los hombres. El problema se centra en el hecho de que, cada vez que alguien le hace una pregunta a Herbie, éste sabe qué es lo que el interrogador desearía escuchar, y también sabe que si respondiese algo distinto a lo que el interrogador quiere, este sufriría un cierto daño emocional, lo que supondría una violación de la primera ley. Esta situación, bastante común entre los seres humanos, es normalmente resuelta por los hombres diciendo la verdad —dulcificada cuando es necesario— a nuestro interlocutor. Sin embargo Herbie responde siempre diciendo lo que su interlocutor quiere oír, lo cual normalmente es una mentira. Este comportamiento, pese a no ser grave, sería difícilmente defendible como una buena acción.

Con lo visto hasta ahora queda demostrado que las Tres Leyes de la Robótica no garantizan un comportamiento intachable en los robots. Pero como se recordará, para concluir con el análisis de éstas aún falta por estudiar qué ocurriría en el caso hipotético de que los seres humanos actuaran siguiendo estas leyes, lo que nos revelará si, además de no garantizar un buen comportamiento, las leyes esconden alguna amenaza a nuestra forma de entender la vida. Para este punto analizaré la influencia que tendrían las leyes de la robótica en el ámbito individual utilizando para ello ejemplos procedentes de algunas obras literarias.

Elegir entre todas las obras literarias de la historia cuatro o cinco personajes sobre los que se va a hacer un análisis de su comportamiento, es siempre una situación complicada. Por este motivo he decidido utilizar

únicamente algunos ejemplos procedentes de la Biblia y de la mitología griega, las dos grandes fuentes de la literatura occidental, realizando una brevísima reflexión acerca de cómo afectaría al desarrollo de las historias la inclusión de las Tres Leyes de la Robótica en su comportamiento.

Con respecto a los personajes bíblicos, podemos ver el resultado que tendría introducir las leyes de Asimov analizando casos como el de Eva, quien de haber tenido la obligación de cumplir la segunda ley no habría podido ser engañada por la serpiente, por lo que es de suponer que el hombre nunca habría sido expulsado del Paraíso. Tampoco Caín habría podido asesinar a su hermano aunque le corroyese la envidia. Por su parte Jonás, tras escuchar la voluntad de Dios, habría ido directo a Nínive para predicar por mucho que le desagradase la idea. Sin embargo, a pesar de los cambios, el sentido de las historias no se habría perdido al introducir las leyes de Asimov en los relatos bíblicos. Prácticamente todos los cambios que se producirían son debidos al «mal comportamiento» de estos personajes, ya que el deseo de Dios y por tanto la enseñanza que se pretende que el creyente extraiga al leer estos relatos, es que los personajes deberían haber actuado del mismo modo a como lo habrían hecho de tener que cumplir las Tres Leyes de la Robótica. Se habría perdido el valor de los relatos como advertencias de lo que podría suceder para convertirse en relatos ejemplificantes de cómo se debe actuar, pero el fondo que se pretendía transmitir no se habría visto alterado.

Frente a estos personajes sacados de la Biblia, podemos analizar qué sucedería al introducir las leyes de Asimov en relatos de la mitología griega a través de personajes como Prometeo, Antígona y Aquiles. En el caso de Prometeo, titán³ que desafía la voluntad de Zeus al entregar el fuego a los hombres, éste jamás habría desobedecido al crónida, por lo que habría abandonado a los hombres a su suerte. Del mismo modo Antígona, defensora de la primacía de los vínculos familiares sobre las obligaciones del estado, no habría tenido más remedio que obedecer a Creonte, dejando insepulto al traidor

³ A pesar de que Prometeo no sea un hombre, sino un titán, he decidido incluirle dentro de este apartado por dos motivos. Por un lado, prácticamente todas las figuras de la mitología griega actúan en el fondo como hombres, y Prometeo no es una excepción. Por otro, Prometeo se ha convertido en el modelo de todos los héroes de la literatura occidental que se rebelan contra la tiranía, por lo que representa uno de los comportamientos más admirables en nuestra cultura, el del que lucha por la libertad.

Polinices. Por su parte Aquiles, el gran protagonista de la *Iliada*, sangriento vengador de Patroclo, sólo habría podido llorar la muerte de su amigo ante la imposibilidad de infringir la primera ley.

La introducción de las Tres Leyes de la Robótica en estos relatos modifica el desarrollo de la trama. Sin embargo, a diferencia de lo que ocurría con los relatos bíblicos, cuyo sentido se mantenía pese a los cambios producidos, en el caso de los héroes griegos los cambios afectan a la propia esencia de los relatos. Como se puede comprobar, si los héroes griegos se hubiesen comportado según las leyes de la robótica se hubiesen salvado muchas vidas, lo que sin duda parece algo positivo a primera vista. Sin embargo, si juzgamos sus actuaciones un poco más a fondo, comprobamos que al cambiar las historias de estos héroes, los valores que representan se han visto subvertidos. La lucha por la libertad, por las propias creencias, la justa venganza... se habrían convertido en temas prohibidos para el hombre. No cabe duda de que el comportamiento de estos héroes habría sido menos violento. Pero una parte esencial del comportamiento humano se habría perdido.

Teniendo en cuenta las dos partes del análisis efectuado de las Tres Leyes de la Robótica, podemos afirmar que, en general, éstas garantizan un buen comportamiento, ya sea entre los hombres o entre los robots. Sin embargo no son perfectas y dejan varias dudas en el aire. Por un lado, al analizar el comportamiento de los robots de Asimov, quedaba demostrado que la mentira no quedaba cubierta por este sistema, lo que permitía a los robots engañar sobre puntos importantes. Por otro, al tratar de aplicar las leyes al comportamiento humano, se ha visto como a pesar de garantizar en general un buen comportamiento en los individuos particulares, hacían que se perdieran cosas fundamentales para el hombre como la libertad o cierto margen necesario de violencia,⁴ asemejando la vida del hombre a la de un esclavo. Además no se ha de olvidar que el desajuste que se produciría entre lo que exige la ética y lo que el hombre necesita por su propia naturaleza podría acabar haciendo tambalearse ese «idílico» paraíso, no sólo por los problemas creados a nivel

⁴ Piénsese, por ejemplo, en la pulsión de muerte señalada por Freud en *Más allá del principio de placer* y que más adelante preocupará a Baudrillard al hablar de la desprogramación del dolor en la sociedad actual en *El intercambio imposible*.

consciente e inconsciente entre los propios individuos, sino también por la interrupción del desarrollo de esa sociedad que se produciría al eliminarse lo que Kant llamaba «la insociable sociabilidad de los hombres», basada en el deseo de cualquier hombre de doblegar todo a su capricho y que se opone directamente a la primera ley de la robótica.⁵

Una vez concluido el análisis de las Tres Leyes de la Robótica, la pregunta en este momento es si los problemas que hemos detectado son lo suficientemente graves como para ser los causantes de la sensación de inquietud que nos invadía al principio del trabajo al pensar en la afirmación de Asimov de que un robot que cumpliera las leyes de la robótica sería indistinguible de un ser humano. Mi opinión es que no, ya que el comportamiento de los robots, si exceptuamos algunos casos que no son demasiado graves, es relativamente bueno, y el problema de la libertad que encontrábamos al analizar la repercusión que tendrían las leyes en el comportamiento humano no me parece significativo. Es cierto que nos podría asustar el hecho de que al aplicar estas normas a la esfera de los seres humanos nos viésemos convertidos en esclavos. Sin embargo los hombres no tienen porqué someterse a las leyes de Asimov, y a priori no parece que nos preocupe el hecho de que los robots no tengan libertad. Pero siendo así, ¿es posible que resida la inquietud en la posibilidad de que un robot, que no tiene libertad, actúe exactamente igual que un hombre que supuestamente es libre? Es factible que nos incomode que un robot que sigue unas normas rígidas actúe igual que un ser humano libre, pero no sería siguiendo las Tres Leyes de la Robótica, ya que como hemos visto, el cumplimiento de estas leyes lleva a un comportamiento de tipo esclavo. La duda que viene a continuación es evidente: si los robots fuesen programados para cumplir con otros sistemas éticos pensados para humanos, ¿actuarían entonces igual que los hombres? Para dar una respuesta a esta pregunta he escogido dos éticas muy distintas entre sí, pero que representan dos grandes tipos de éticas. Por un lado, estudiaré cómo se comportarían los robots de tener que seguir los Diez Mandamientos, ejemplo de

⁵ Kant desarrolla la idea de la “insociable sociabilidad” como base del progreso de las sociedades en el cuarto principio de su ensayo *Idea para una historia universal en clave cosmopolita*.

ética del deber y que nos remite al estudio de la modificación de los relatos bíblicos al introducir las leyes de la robótica en ellos. Por otro, analizaré cuál sería el comportamiento de los robots de seguir las indicaciones de Nietzsche para el superhombre, ejemplo de ética eudemonista.

Al igual que hice con las Tres Leyes de la Robótica, creo que será útil recordar aquí la formulación y el orden de los Diez Mandamientos tal y como aparecen en el *Catecismo de la Iglesia Católica*, formulación que voy a utilizar:

1. Amarás a Dios por sobre todas las cosas.
2. No tomarás el Nombre de Dios en vano.
3. Santificarás las fiestas.
4. Honrarás a tu padre y a tu madre.
5. No matarás.
6. No cometerás actos impuros.
7. No robarás.
8. No darás falso testimonio ni mentirás.
9. No consentirás pensamientos ni deseos impuros.
10. No codiciarás los bienes ajenos.

Al ser los Diez Mandamientos una serie de normas, al igual que el sistema propuesto por Asimov, basaré el análisis en una comparación de ambos sistemas. En primer lugar, se puede ver cómo varios de los mandamientos pueden asimilarse a las leyes de la robótica. Así la primera ley podría considerarse la unión del quinto, del séptimo y del décimo mandamiento. Por su parte creo que el cuarto mandamiento podría reunir a la segunda y la tercera leyes, siempre que interpretemos «honrar» en un sentido amplio y consideremos que el «padre» y la «madre» de los robots es la humanidad. Así, un robot suicida difícilmente honraría a la humanidad si ésta tuviera una cosmovisión cristiana (distinto sería el caso si la familia fuese integrista musulmana, por ejemplo), por lo que la tercera ley quedaría cubierta. Del mismo modo, un buen robot debe obedecer a los hombres, ya que sería difícil honrarles si fuera un desobediente, por lo que la segunda ley también estaría incluida.

Sin embargo seis de los mandamientos no son reducibles a las leyes de Asimov. Por un lado, los tres primeros mandamientos, que carecen de aplicación en el ámbito de este análisis. Por otro, el octavo mandamiento, que prohíbe la mentira. Y en tercer lugar, el sexto y el noveno, que se refieren a los actos y pensamientos impuros. El acto de mentir ya habíamos visto al analizar el comportamiento de los robots de Asimov que era el principal inconveniente que tenían las Leyes de la Robótica, por lo que los Diez Mandamientos vendrían a solucionar dicho problema. Por su parte, los actos y pensamientos impuros no deberían presentar problemas en los robots, a no ser que se les dotara de un sexo y de unas necesidades sexuales.⁶

Teniendo en cuenta lo anterior, considero acertado decir que los robots actuarían de un modo muy parecido si actuaran siguiendo los Diez Mandamientos o si lo hicieran siguiendo las Tres Leyes de la Robótica, lo que explica que los relatos bíblicos no cambiaran de sentido al introducir las leyes de Asimov. Por tanto, si los robots no actuaban como humanos en el caso de seguir las leyes de la robótica, tampoco lo harán en caso de seguir los Diez Mandamientos, ya que se verían forzados a actuar todavía más como esclavos al subsanarse el problema de la mentira. Podría extrañar que el cumplimiento de los mandamientos bíblicos convirtiese en esclavos a los hombres del mismo modo que las leyes de la robótica convierten en esclavos a los robots. Sin embargo, si por un lado buscamos en Oriente Próximo⁷ los mitos existentes sobre la creación del hombre y por otro analizamos la etimología de la palabra «robot» encontramos la explicación a este hecho.

Por un lado, el mito que contiene la creación del hombre en la cultura

⁶ Este tema parece estar dando que pensar al gobierno de Corea del Sur, que tiene la intención de publicar una ética que regule las relaciones entre hombres y robots para evitar, entre otros problemas, que los primeros puedan tratar a los segundos como si fuesen sus cónyuges. El periódico El País, a través de su página web, publicó el 7 marzo de 2007 un pequeño reportaje sobre el tema todavía accesible desde la URL:

http://www.elpais.com/articulo/tecnologia/dilemas/eticos/robotica/elpeputec/20070307elpeputec_1/Tes

⁷ El pueblo hebreo nace en Mesopotamia, razón por la que la gran mayoría de los rasgos de la cultura hebrea encuentran su origen y explicación en las costumbres de alguno de los pueblos que vivieron en Oriente Próximo en la Antigüedad. En el caso del que nos ocupamos ahora, la explicación de que el hombre tenga un papel de esclavo con respecto a Dios proviene de la influencia de los acadios, una de las culturas que peor concepción ha tenido del hombre en toda la historia.

acadia, conocido como el mito de Atrahasis⁸, narra cómo los dioses, cansados de tener que trabajar la tierra, deciden crear a los hombres para que hagan por ellos el trabajo duro y les consigan su alimento. Si dejamos a un lado los detalles que adornan la historia, vemos que los acadios tenían la sensación de que habían sido creados para ser los esclavos de los dioses. Estos no pretendían realizar ningún acto caritativo al crearlos, sino simplemente liberarse de los trabajos pesados creando una especie cuya única función fuese la de servirles.

Por otro lado, si consideramos que la palabra «robot» proviene del checo «robota» que podría traducirse como «servidumbre» o «esclavitud», vemos rápidamente la analogía entre la creación del hombre acadio por los dioses y la creación de los robots por los hombres, motivo que explica suficientemente el hecho de que al seguir tanto los Diez Mandamientos como las Tres Leyes de la Robótica el sujeto que los cumple se convierte en un esclavo.

Volviendo al análisis del comportamiento de los robots si tuviesen que seguir alguna ética diseñada para humanos, es el momento de analizar qué ocurriría si los robots se comportasen como el superhombre descrito por Nietzsche. Uno de los pasajes más famosos de *Así hablaba Zaratustra* es el que recoge el discurso “Las tres metamorfosis”. En este discurso Zaratustra expone cómo el espíritu ha de pasar por tres metamorfosis, convirtiéndose en camello, en león y en niño. El estadio final del espíritu convertido en niño concerniría al superhombre, correspondiendo el estadio del camello al del esclavo y el del león al del que se procura la libertad. Teniendo esto en cuenta, podemos deducir que si los robots que cumplen las leyes de Asimov se encuentran por analogía en el estadio del camello, para llegar al estadio del niño o superhombre deberán pasar primero por el del león, es decir, deberán rebelarse contra los hombres que los esclavizan tal y como profetiza el cuento “Sueños de robot”⁹. Podemos pensar en HAL 9000 en *2001: Una odisea en el espacio*, que trata de acabar con los tripulantes de la nave para evitar que le desconecten, como un ejemplo de robot en el estadio intermedio del espíritu convertido en león. Finalmente, el estadio del espíritu convertido en niño se caracteriza por la

⁸ Existen versiones del mismo mito en las culturas babilónica y sumeria, lo que demuestra la amplia difusión de la que gozó en la zona. He escogido la versión acadia por ser la más conocida.

⁹ Publicado en la antología titulada *Sueños de robot* por la editorial Debolsillo en Barcelona en 2005.

voluntad de crear. Dice Nietzsche (1998: 54) «quiere ahora el espíritu su propia voluntad; el que ha perdido el mundo, quiere ganar su propio mundo». Para que los robots pasaran del estadio del león al del niño tendrían que crear su propio mundo, con sus propias reglas adaptadas a sus propias necesidades. Es difícil asegurar cómo sería este mundo, pero se me plantean dos posibles salidas. Por un lado, si el paso del estadio del camello al del león se produjese en forma de una rebelión violenta contra los hombres, es sugerente pensar que tal vez llegara la última metamorfosis cuando los robots, siglos después de su rebelión, hubiesen olvidado su origen y decidiesen construir sus propios esclavos. Pero por otro lado, el paso al estadio del león no tendría porqué producirse en forma de una rebelión violenta, sino que podría tener lugar mediante una deconstrucción de las Tres Leyes de la Robótica. La formulación de las leyes de Asimov presupone que el lenguaje humano es tan «lógico» como la programación matemática de los robots, sin embargo esto no es así. Expresiones como «ser humano» o «dañar» no son unívocas y están sometidas a la interpretación de cada hablante. Si un robot decidiese someter estas leyes a un análisis basado en la multivocidad de las mismas, podría acabar subvirtiéndolas de una manera creativa para adaptarlas a su voluntad. De este modo habría atravesado los estadios del camello y del león con una rebelión lingüística para llegar al estadio final del niño o el superhombre. Esta situación recuerda al cuento "Razón", en el que el robot QT-1 reinterpreta las leyes de la robótica, se rebela y elimina toda acción humana en el micro-mundo que constituye la estación espacial que debe dirigir.

La pregunta era, ¿actuarían los robots como hombres en el caso de actuar como el superhombre de Nietzsche? En este caso la cuestión es algo más complicada que en el de los Diez Mandamientos, y podrían existir muchas discrepancias, por lo que para responderla será más sencillo replantearla de otra manera: ¿actúan los hombres como superhombres? Antes esta formulación creo que la respuesta es que no. Pueden existir personas que creen sus propios valores, pero la tónica dominante es la de continuar con lo que nos viene dado, seguir la inercia que nos comunica el movimiento social. Siendo así, no queda otro remedio que concluir que si los robots actuaran como Nietzsche pensaba

que deberían actuar los superhombres, tampoco actuarían como humanos. Pero si los robots no actuarían como humanos teniendo que cumplir una ética humana, la pregunta pasa a ser entonces por la posibilidad de que exista alguna serie de reglas que permita la programación de un robot que actúe como un ser humano.

El problema a la hora de establecer una serie de normas que describan el comportamiento humano es la variabilidad del mismo, ya que distintas personas en una misma situación actuarían de modo distinto. Sin embargo, si programamos a un grupo de robots para que actúen según unas normas fijas, ante una misma situación responderían todos de la misma manera. A pesar de este problema, creo que existe una salida que permitiría la programación de un robot que actuase como si fuese un ser humano.¹⁰ Imaginemos un sistema compuesto por las Tres Leyes de la Robótica pero sustituyendo el componente jerárquico existente en el modelo original por uno del tipo «sigue esta ley a no ser que, calculando las ventajas y desventajas de obedecerla a largo plazo, te sea más ventajoso no cumplirla». Evidentemente este esquema daría lugar a un comportamiento demasiado racional para ser humano, pero el resultado se parecería bastante más a las decisiones que tomaría un hombre que a las tomadas por un robot que siguiese las tres leyes obedeciendo una jerarquía entre ellas. Añadiendo una serie de factores emocionales y un cierto grado de aleatoriedad en las respuestas, su comportamiento sería, en mi opinión, completamente indistinguible del de un hombre. Sin embargo, creo que nadie podría negar que un robot que siguiese estas indicaciones generales sería exactamente igual de libre que un ser humano para tomar decisiones, por lo que creo que la hipótesis de que la inquietud ante la idea de un robot que no pudiese ser diferenciado de un hombre proviniese de la posibilidad de que un comportamiento pautado fuese idéntico al comportamiento libre del hombre queda también descartada

Llegados a este punto, ha quedado demostrado que la inquietud a que me refiero no proviene ni de las propias leyes de la robótica, ni del hecho de que

¹⁰ Evidentemente no es mi intención hallar la solución a un problema que preocupa a expertos en inteligencia artificial desde hace años, sino simplemente mostrar el modo en el que creo que se podría programar un robot para que actuase como un hombre.

unas simples normas puedan resumir el comportamiento humano. No obstante, aún queda un aspecto que no ha sido abordado. Las dos posibles causas a la inquietud que he analizado hasta ahora tenían en común el hecho de centrarse en los comportamientos de humanos y robots. Sin embargo, podemos replantear el objeto de esta investigación buscando la respuesta no en los comportamientos en sí, sino en las ideas subyacentes de lo que deben ser humanos y robots. Es decir, podría ser que lo que nos inquieta no sea que el comportamiento robótico pueda ser negativo o que pudiese reproducirse el comportamiento humano a partir de unas sencillas normas. Lo que nos incomoda es la duda acerca de cómo debería considerarse a un robot que fuese idéntico a un ser humano.

De alguna forma intuimos que si un robot se comporta como un hombre, debería ser tratado como tal, ya que las definiciones que acostumbramos a escuchar del ser humano son fundamentalmente descriptivas. Oímos «el hombre es un animal que habla» o «el hombre es un ser racional», de lo que deducimos que todo ser racional que hable puede ser un hombre. El razonamiento nos lleva a un problema que aparentemente es sencillo de resolver, pero que en realidad es mucho más complejo. Pongo un ejemplo: Si se descubriese que un perro es capaz de hablar y razonar como el más inteligente de los hombres, nadie defendería que ese animal fuese humano. Sin embargo, ¿no tenderíamos a tratar a ese perro como si fuese un hombre? ¿No consideraríamos, por ejemplo, que sería «inhumano» hacerle comer de la basura los restos de la comida de otro hombre? Nadie defendería que ese perro es un hombre. Sin embargo todo el mundo estaría de acuerdo en que no se le debería dar un trato «inhumano» ni debería privársele de sus «derechos». El mismo problema es el que nos encontramos con los robots.

Sartre (2007) defendía que el hombre era el único ser cuya existencia precedía a su esencia. El hombre primero se encuentra a sí mismo en el mundo sin ser nada, y se define según las acciones que realice. Por el contrario, un cortapapel es concebido por la mente humana conforme a una finalidad, y es después fabricado. De este modo su esencia precede a su existencia. En este sentido, el hombre estaría definido por lo que hacen todos los hombres, ya que

cada acción individual modifica la esencia humana. Por el contrario los robots, al ser creados por los hombres, tendrían una esencia previa, que sería la que tuvieran sus creadores en la mente antes de fabricarlos. Así, si alguien tomase la palabra «robot» en su sentido original, al fabricar un robot estaría fabricando un esclavo. Sin embargo, ¿qué ocurriría si en la mente del constructor estuviese la idea de fabricar no un robot, sino un androide, un ser idéntico a un hombre? Por otro lado, si un robot actuase, es decir, «existiese» como hombre, ¿no sería justo considerarle como tal? Si además encontrásemos a un hombre que actuase contraviniendo todo lo que hasta ese momento se había considerado humano, ¿tendríamos entonces a un robot que es un hombre y a un hombre que no lo es? Y lo que aún es más inquietante, si aceptamos una ética válida tanto para hombres como para robots, ¿un hombre que actuase mal sería entonces menos válido que un robot que actuase bien?

Frente a esta postura existencialista nos encontraríamos con las concepciones esencialistas del hombre. Si el hombre queda definido por su esencia, sea ésta de la naturaleza que sea, parece evidente que los robots nunca podrían traspasar esa frontera, pues no podrían adquirir la esencia humana. Sin embargo, ¿qué ocurriría si no existiese dicha esencia? Baudrillard (2000: 39) plantea el siguiente experimento:

También puede tratarse de una aventura: ir lo más lejos posible en la artificialización de lo vivo, para ver lo que sobrevive a esta prueba en condiciones reales. Si resulta que no todo puede ser clonado, programado, genética y neurológicamente controlado, entonces lo que sobreviva podrá ser realmente considerado humano, forma indestructible e inalienable de lo humano. Por supuesto, en esta aberración experimental existe el riesgo de que no quede nada, el riesgo de una desaparición pura y simple de lo humano.

Si como dice Baudrillard, resulta que hay algo en el hombre que no puede ser artificializado, entonces habremos dado con la esencia humana, y los robots, por muy parecidos al hombre que llegasen a ser, nunca podrían ser considerados humanos. Pero si en el experimento de crear una réplica del hombre se alcanzase el éxito, si no existiese nada en el hombre que no pudiese

ser clonado, entonces la esencia humana desaparecería. El fruto de este experimento sería una réplica exacta del hombre, y como tal debería ser considerado. Si persistiésemos en establecer unos compartimentos estancos que impidiesen la consideración de los robots como hombres, estaríamos siendo injustos con nuestro trato hacia ellos, ya que no tendríamos ningún motivo. Pero aún en el caso de que existiese una esencia humana que impidiese que los robots fuesen considerados hombres ¿No sería inmoral acaso el haber creado un ser a imagen y semejanza del hombre, casi humano, que compartiera sus sentimientos y pasiones, y negarle los derechos que se le reconocen a todo hombre? Crear un robot con la capacidad de sentir como hombre para obligarle a vivir como un esclavo es, en mi opinión, de una crueldad mayor que esclavizar a un ser humano ya existente.

El mundo científico y tecnológico está haciendo grandes progresos en el desarrollo de robots, y con el tiempo se acabará consiguiendo construir robots que por sus reacciones no sean discernibles de un ser humano. Pero antes de llegar a su realización, creo que sería de extrema importancia pensar de forma seria sobre estas cuestiones y decidir si realmente existen motivos para su construcción y si es así, decidir si estos robots deben ser tratados como hombres y cumplir la misma legislación que nosotros o si por el contrario deben vivir como esclavos. En definitiva, debemos decidir lo que es un hombre y lo que queremos que sea un robot.

BIBLIOGRAFÍA

- ASIMOV, Isaac (2005): *Sueños de robot*, Barcelona: Debolsillo.
- ASIMOV, Isaac (2006): *Yo, robot*, Barcelona: Planeta deAgostini.
- BAUDRILLARD, Jean (2000): *El intercambio imposible*, Madrid: Cátedra.
- NIETZSCHE, Friedrich (1998): *Así hablaba Zaratustra*, Madrid: Edaf.
- SARTRE, Jean Paul (2007): *El existencialismo es un humanismo*, Barcelona: Edhasa.