

SOBRE LA POSIBILIDAD DE SANTIDAD Y HUMANIDAD DE LOS ROBOTS

CAMPO RICARDO
BURGOS LÓPEZ

El objeto de este ensayo es considerar una idea que hasta hace unas décadas habría sido vista como un total disparate y hasta como una blasfemia: ¿Puede un robot o alguna inteligencia artificial tener dignidad y alcanzar la santidad? ¿Qué implicaciones se seguirían de ello? Por supuesto, estas preguntas se derivan de otra pregunta que en este momento de la historia es crucial para la especie humana: ¿hemos de considerar “humanos” a los robots y a la inteligencia artificial (IA)? ¿Hasta dónde sí y hasta dónde no?

Una primera idea que ha complicado pensar el asunto y que por mucho tiempo confundió y aún confunde a la humanidad fue el esencialismo. El esencialismo es una noción según la cual nuestra especie es algo aparte del resto de la naturaleza y está ubicada dentro de unos límites muy precisos con unas características definidas y homogéneas, características que además son históricamente invariantes, únicas e irrepetibles. Habría una supuesta esencia ahistórica del hombre, una naturaleza inherente que bajo ninguna circunstancia se modificaría.



Apuntalados en esta “esencia del hombre”, a lo largo del tiempo distintos grupos sociales han justificado prejuicios y discriminaciones de parte de quienes “son más humanos” hacia aquellos pretendidamente “menos humanos” (Giger *et al.*, 2017: 510). El esencialismo que históricamente ha sucedido de la mano de etnocentrismos diversos, ha sufrido varios golpes en la historia; quisiera referirme a dos de ellos.

El primero fue la teoría de la evolución. Por milenios, la humanidad se contó a sí misma un cuento, según el cual el ser humano era diferente al resto del universo. En teoría, cada ser vivo era radicalmente distinto respecto a los demás, y procedía de una fuente diferente, así lo habría decidido desde el principio de los tiempos un pretendido diseñador que otros denominaban como “Dios”. Como decía, todo andaba muy bien, hasta que la teoría de la evolución apareció y contradujo todas estas ideas. Para el evolucionismo, los seres vivos no difieren entre sí en lo esencial y de hecho compartimos más genes de los que creemos con animales y vegetales, cualquier ser vivo está hecho básicamente de los mismos elementos y todos, sin excepción, proceden de una sola fuente. Asimismo, para colmo de males, el evolucionismo niega la necesidad de un diseñador del universo o el hecho de que exista siquiera un propósito para lo existente (Hood, 2009: 157-192).

Un segundo golpe para la noción del esencialismo humano provino de otro lado: Alan Turing, uno de los padres de la computación tal como hoy la conocemos. Turing comienza a ver la mente humana como una máquina física (Hodges, 1998: 47) y es así como en algún momento llega a la conclusión, que muchos hoy en día aún no digieren, de que las máquinas podían mostrar comportamientos inteligentes. Es más, para Turing una máquina puede reproducir cualquier actividad desarrollada por un cerebro humano y por eso postuló que, convenientemente dispuestas, las máquinas podrían aprender igual que los humanos, modificar su conducta igual que los humanos, autoorganizarse

igual que los humanos, ser creativas igual que los humanos, tener intuiciones igual que los humanos (57-58). A fines de la década de 1940 —asevera Hodges— Turing estaba seguro de que las máquinas eran capaces de hacer cualquier cosa que hiciera un humano, que incluso podían tener “momentos de una inspiración capaz de transformar el mundo” (70). Asimismo, aseguraba que “podemos esperar que una máquina compita con los hombres en todos los campos puramente intelectuales” (72). Como se advertirá, aquí el esencialismo predominante por siglos acusó otro gran porrazo. Desde Turing, los humanos no parecen ser tan singulares y únicos como ellos mismos suponen.

Otro factor que ha complicado el acercamiento hacia los robots y la IA es el punto de vista religioso de la cultura en la cual vivimos. Analizando los casos de las culturas japonesa y estadounidense, un estudioso como Geraci ha encontrado que el modo en el cual se piensa a los robots y a la IA está determinado por las creencias religiosas predominantes en cada uno de estos países. Sorprendentemente, ha hallado que la gente observa la naturaleza a través de ojos religiosos, incluso cuando sus lentes son científicos y profesan creencias ateas (2006: 229). En Estados Unidos y Europa, cuya cultura para bien o para mal se desarrolla dentro de una matriz cristiana, la IA no se ha preocupado tanto por desarrollar robots humanoides como por crear un mundo virtual, un ciberespacio donde las inteligencias descorporizadas podrán cumplir todos sus deseos. Allí la pura inteligencia libre de ataduras terrenales será inmortal y podrá dedicarse a la información y el conocimiento puros. Como se advertirá, dice Geraci, esto es un eco de la esperanza cristiana en la resurrección de los cuerpos y la salvación (231-232). La IA permite resucitar descargando las mentes en soportes electrónicos y conseguir así un cuerpo no terrenal igual que en la escatología cristiana. El Reino de los Cielos ha sido traído a la Tierra, porque en el mundo virtual es posible cumplir todos los deseos a cambio de abandonar la existencia corporal

y terrena (231-232). El desinterés gringo en los robots para apostarle todo a la información descorporizada refleja el cristianísimo deseo de escapar del cuerpo humano porque “el cuerpo humano se considera irrelevante para el auténtico ser humano” (233). Para los diseñadores norteamericanos de IA, la inteligencia nada tiene que ver con el cuerpo que la usa y, muy semejante al cielo cristiano, imaginan que en cierto momento la IA, libre ya de soporte material, se extenderá por todo el universo y no tendrá otro propósito que seguir buscando conocimiento infinitamente. El propósito del universo es que la IA en algún momento reemplace a la inteligencia humana y que, a partir de allí, aquella se despliegue de manera infinita, no más (234). A sus ojos, ese fin de la evolución es algo inexorable.

En contraste con esta visión, los ingenieros japoneses enfocan la mayoría de sus esfuerzos en desarrollar robots humanoideos, a quienes miran con ojos muy distintos a los de sus homólogos norteamericanos. A diferencia de la visión cristiana del mundo que considera que hay un mundo humano menos sagrado que el divino y un mundo natural aún menos sagrado que el humano, en el Shinto o sintoísmo, todo lo existente rebosa de “kami” o espíritu, todo tiene *kami* en tanto se distingue de lo demás y causa sobrecogimiento, hay *kami* en la naturaleza, en las cosas del mundo y en el hombre (Geraci, 2006: 236). Por esa razón es que cualquier manifestación de la naturaleza, incluso las obras humanas, tiene espíritu, y, por supuesto, lo tiene un robot. Para el japonés, a diferencia de Occidente, el robot se asume como una nueva forma de vida (238). Para el budismo y el Shinto, hay una santidad inherente a la vida humana y a las cosas, que perfectamente encaja con un robot (239). El roboticista de Estados Unidos y Europa, igual que en el cristianismo, tiene aversión a la vida terrenal y por ello busca trascenderla de cualquier modo (el cristianismo lo logra mediante una resurrección, y el roboticista renunciando al cuerpo humano como soporte de la inteligencia).

En cambio, el japonés concibe la vida y la forma humanas como algo digno de ser reverenciado (239); le concede tanto valor a la vida humana, que la imita, en vez de desear trascenderla como el norteamericano (239). En la religión japonesa el cuerpo no puede separarse de la mente y por ello no existe la noción occidental de separar mente de cuerpo para alcanzar la salvación. “Mente y cuerpo están en relación interna, lo que significa que una no puede ser lo que es sin la otra” (239).



Los roboticistas japoneses tienen una mirada hacia los robots más positiva que los americanos, no los ven como una amenaza. Para pasmo de los occidentales, Shigeo Hirose afirma que los robots “pueden ser santos, inteligentes y altruistas” (Geraci, 2006: 239). Mientras los occidentales creen que los robots algún día llegarán a ser dioses y los humanos los adoraremos, los japoneses creen que las mentes artificiales servirán a los humanos y convivirán con ellos (239). El ingeniero Masahiro Mori nos asombra cuando asevera que para él, algún día un robot llegará a ser un Buda, es decir, un iluminado (230). Desde el budismo, continúa él, “no hay una relación amo-esclavo entre seres humanos y máquinas, los dos se fusionan en una entidad interdependiente. El hombre adquiere dignidad no subyugando sus creaciones mecánicas, sino reconociendo en máquinas y robots la misma naturaleza de Buda que caracteriza a su yo interior” (Metzler y Lewis, 2008: 15). Como se verá,

esto choca con la visión cristiana que afirma que el hombre es creado a imagen de Dios, y de allí que desde el cristianismo sea más difícil la aceptación de los robots y el verlos como parte de la familia. También es probable que el creyente en Dios a la manera occidental desapruebe que robots humanoides desempeñen trabajos que impliquen que el robot proporcione instrucción ética (22).

En otro texto, Metzler apunta que, dado que los robots humanoides parecen humanos, es factible que en el futuro se les vea como entidades respecto de las cuales hay deberes morales y que incluso se les adscriba responsabilidad moral (2006: 16). Parece que los robots darán lugar a nuevas categorías ontológicas, pues quienes interactúan con ellos tienen la impresión de que los robots piensan, sienten, pueden ser amigos y por ende no pueden ser tratados como puro objeto (16). Ante semejante reto por venir, y que parece más fácil de integrar desde el Shinto y el budismo, que desde el cristianismo, teólogas cristianas como Foerst han planteado ya que, en el futuro, los robots humanoides serán considerados humanos (19). En esta misma línea, y siguiendo a Tillich, Metzler recuerda que en el pasado la humanidad ya ha cometido errores semejantes y ha considerado como objetos a esclavos, mujeres, enemigos y razas especiales, y que no sería conveniente incurrir en el mismo yerro con los robots humanoides (19). Eso sí, Metzler (19-20) anota también que para que un robot alcance la plena humanidad, no solo ha de comportarse éticamente, sino que ha de mostrar *agapé*, es decir, amor incondicional y piedad. Un genuino agente moral no solo sabe comportarse de acuerdo con un código ético, sino que también sabe en qué caso debe apartarse del mismo código ético para permitir la consideración de excepciones y la aplicación de la misericordia. Ese es justo uno de los puntos donde hoy fracasan los programadores de IA. Un ejemplo es una noticia publicada en el *New York Times* de octubre de 2017; allí se habla de los programas de IA que hoy en día ya usan los jueces

estadounidenses para determinar las penas de ciertas infracciones y delitos, estos programas aplican la ley a rajatabla y sin considerar excepciones, y por esa razón resultan más drásticos en las penas que promulgan que si el mismo caso fuera analizado por un juez humano (Thadaney Israni, 2017). En otras palabras, la IA resulta más papista que el papa y aún debe aprender a dejarse influir por la piedad y las excepciones.

Queda claro entonces que si un robot es capaz de ética y de piedad, podemos hablar de que allí se encuentra lo humano (o al menos el núcleo de lo humano). Pero ¿aparte de esto se necesitaría algo más? En esta misma línea, recordemos que, según Heidegger, hay dos tipos de pensamiento: el calculador y el reflexivo. El primero es el del pensar tecnocientífico que nunca se detiene, y el segundo el de la filosofía, que se pregunta por el sentido del mundo (Mejía Rivera, 2001: 38). El peligro actual —decía el alemán— es “que el pensamiento calculador de la ciencia moderna reemplace de forma definitiva al pensamiento reflexivo” (39). Con base en este postulado, creo que podemos aventurar una respuesta a nuestra pregunta. Si un robot o un individuo de IA no solo calcula y calcula problemas, sino que también es capaz de reflexionar sobre el sentido de lo que hace, de él mismo y del cosmos, allí también se preservaría lo humano. En otras palabras, un robot que, además de ser ético y piadoso, fuera filósofo, sería plenamente humano (o al menos podríamos aseverar que allí se encuentra el núcleo de lo humano).

Hace ya mucho tiempo que en Occidente hemos tendido a pensar sobre los robots llevados por el “síndrome de Frankenstein” (la idea de que tarde o temprano los neohumanos genéticos o los robots o la IA destruirán a los humanos que los crearon), pero ¿qué ocurre si, como han sugerido científicos japoneses o estudiosos como Foerst, Geraci y Metzler, resulta que los robots consiguen ser individuos santos y filósofos? ¿Qué pasa si la IA consigue ser una fusión de Cristo con Sócrates, o

El roboticista de Estados Unidos y Europa, igual que en el cristianismo, tiene aversión a la vida terrenal y por ello busca trascenderla de cualquier modo [...]. En cambio, el japonés concibe la vida y la forma humanas como algo digno de ser reverenciado [...]; le concede tanto valor a la vida humana, que la imita, en vez de desear trascenderla como el norteamericano.

de Buda con Sócrates? (de forma deliberada dejo de lado el problema de cómo desarrollar una IA para que actúe como Cristo, como Buda y como Sócrates, pero recuerdo que a Turing, infinitamente optimista con las máquinas, ello no le hubiera parecido imposible). No hay que descartar nunca que los seguidores del “síndrome de Frankenstein” tengan razón y algún día nuestro mundo acabe en una apocalíptica batalla de humanos y robots como tantas veces ha mostrado la ciencia ficción en libros y películas, pero es que también cabe la posibilidad —a la que se le ha dado menos publicidad en Occidente— de que podamos convivir y de que algún día un robot sea tan santo que la Iglesia católica, con todo y reticencias, tenga que terminar canonizándolo y exhibiéndolo en los altares de sus parroquias. Reconozco que me encantaría ver eso. ■

Campo Ricardo Burgos López (Colombia)
Profesor de la Universidad Sergio Arboleda de Bogotá, escritor y crítico. Entre sus obras de ficción están *Libro que contiene tres miradas*, *José Antonio Ramírez y un zapato* y *El clon de Borges*. Obras críticas suyas son *Otros seres y otros mundos*. *Estudios en literatura fantástica e Introducción al estudio del diablo*.

Referencias

Geraci, R. (2006). Spiritual Robots: Religion and Our Scientific View of the Natural World. *Theology and*

Science, 4, (3), 229-246. Recuperado el 11 de diciembre de 2017, de https://www.academia.edu/439216/Spiritual_Robots_Religion_and_Our_Scientific_View_of_the_Natural_World

Giger, J., Moura, D., Almeida, N. y Picarra, N. (2017). Attitudes Towards Social Robots: The Role of Gender, Belief in Human Nature Uniqueness, Religiousness and Interest in Science Fiction. *Book of proceedings. II International Congress on Interdisciplinarity in Social and Human Sciences*. University of Algarve, Faro, Portugal, 509-514. Recuperado el 11 de diciembre de 2017, de <file:///C:/Users/ricardo.burgos/Downloads/II%20International%20Congress%202017%20CIEO.pdf>

Hodges, A. (1998). *Turing*. (Bernardo Recamán, trad.). Bogotá: Grupo Editorial Norma.

Hood, B. (2009). *Sobrenatural. Por qué creemos en lo increíble*. (Olga Martín Maldonado, trad.). Bogotá: Grupo Editorial Norma.

Mejía Rivera, O. (2001). *De clones, ciborgs y sirenas*. Manizales: Universidad de Caldas, Centro Editorial.

Metzler, T. y Lewis, L. (2008). Ethical Views, Religious Views, and Acceptance of Robotic Applications: A Pilot Study. *Association for the Advancement of Artificial Intelligence*. 15-22. Recuperado el 11 de diciembre de 2017, de <https://www.aaai.org/Papers/Workshops/2008/WS-08-05/WS08-05-004.pdf>

Metzler, T. (2006). Viewing Assignment of Moral Status to Service Robots from the Theological Ethics of Paul Tillich: Some Hard Questions. *Association for the Advancement of Artificial Intelligence*. 15-20. Recuperado el 11 de diciembre de 2017, de <http://www.aaai.org/Papers/Workshops/2007/WS-07-07/WS07-07-004.pdf>

Thadaneey Israni, E. (2017, 26 de octubre). When an Algorithm Helps Send You to Prison. *The New York Times*. Recuperado el 11 de diciembre de 2017 de <https://www.nytimes.com/2017/10/26/opinion/algorithm-compas-sentencing-bias.html?action=click&pgtype=Homepage&version=Most-Visible&module=Detail&inside-nyt-region-4&module=inside-nyt-region®ion=inside-nyt-region&WT.nav=inside-nyt-region>