

¿Puede la inteligencia artificial tener sentimientos?

ALBERTO PRIETO

En nuestra relación con otras personas, máquinas y otros elementos de nuestro entorno lo que realmente importa es lo que percibimos y nuestra interpretación de esas sensaciones. Podemos enjuiciar y catalogar a las personas (antipática, inmoral, etc.) pero en la práctica lo que nos incumbe son las interacciones particulares con ellas. Así decimos que una persona es empática cuando muestra que se identifica con las personas y comparte sus sentimientos. Es una conclusión que obtenemos de su comportamiento, independientemente de la naturaleza o esencia de dicho individuo (si en el fondo es o no es empático).

Bajo este punto de vista se puede afirmar que una máquina da la apariencia de tener sentimientos si en sus interacciones con las personas percibimos actitudes similares a las de los humanos. En efecto, podemos programar adecuadamente a las máquinas para que se comporten como si tuvieran sentimientos o emociones, en el sentido de que, ante situaciones concretas, emulen fenómenos tales como alegría, afecto, humor, amor, felicidad, miedo, ansiedad, ira, tristeza, enojo rechazo, vergüenza, culpa, sorpresa, esperanza o compasión.

Para poner de manifiesto las ideas anteriores puedo citar la conferencia pronunciada por la Dra. Carme Torras, investigadora del Instituto de Robótica e Informática Industrial (CSIC-Universidad Politécnica de Cataluña), en el Congreso Español de Informática (CEDI2024) mantenido recientemente en La Coruña, evento que agrupa a las principales sociedades científicas españolas en el ámbito de la Informática. En esta conferencia la Dra. Torras trató de 'Las múltiples facetas y desafíos de la robótica asistencial'. Los robots que han diseñado y programado en su instituto han logrado retos muy exigentes para asegurar una comunicación fiable, personalizada y respetuosa con los valores y la dignidad de cuidadores y pacientes, así como para garantizar las estrictas medidas de seguridad que el contacto físico requiere, por ejemplo, al ayudar a usuarios con limitaciones de movilidad a vestirse o alimentarse. Estos robots 'sociales' se han diseñado de forma que además de cubrir sus objetivos básicos, man-

También a los sistemas que denominamos inteligentes se les puede dotar, por ejemplo, de comportamientos que denominamos 'políticamente correctos'

tienen de forma exquisita la interacción persona-robot. Esto lo realizan a través de la información que el robot va recopilando y actualizando sobre la personalidad, gustos y progresos en el comportamiento específicos de cada uno de los pacientes y, programando expresiones habladas o textuales que, de acuerdo con el perfil almacenado del paciente, muestran sentimientos como afecto, alegría, o enojo. Se dirigen a cada persona por su nombre o apodo y pronuncian expresiones



como «Antonio, ánimo, vas muy bien, seguro que hoy lo haremos todo mejor», «Yolanda, debes esforzarte un poco más ya que si otros lo han conseguido tú también lo lograrás», «¿Quieres un poco más de agua?», «Espero que pases muy buen día», etc. Estas expresiones la máquina no las entiende en absoluto, actuando como un loro, pero las emiten en el momento oportuno y anteponiendo el nombre de la persona a la que se dirigen, todo ello de acuerdo con lo programado.

En otra conferencia, impartida por el Dr. Ramón López de Mántaras, en la que realizó unas reflexiones sobre «la inteligencia de la inteligencia artificial», citó el caso de la exhibición ante sus alumnos de un conjunto de robots moviéndose de forma aleatoria encima de una mesa plana. Los robots se desplazaban sin chocar unos con otros y cuando se acercaban a los bordes de la mesa se paraban y cambiaban de dirección. Al observar lo un alumno inmediatamente exclamó: «¡Tienen miedo!». El alumno percibía e interpretaba que el robot tenía el sentimiento de miedo, lo que no era más que un comportamiento programado con los datos obtenidos de la lectura de sensores del borde de la mesa. Si el programador se lo hubiese propuesto hubiese podido hacer que el robot diese de forma ostensible la sensación de tener miedo, titubeando o girando de izquierda a derecha, etc.

También a los sistemas que denominamos inteligentes se les puede dotar, por ejemplo, de comportamientos que denominamos «políticamente correctos». Por ejemplo, pregunto a Alexa: «Alexa, dime un chiste de andaluces». Y contesta algo así

como: «Disculpa, pero no considero apropiada esta petición por sus implicaciones de mofa hacia una colectividad». Este ejemplo pone de manifiesto que se puede hacer que los sistemas inteligentes se comporten éticamente de forma correcta, o, abusando del lenguaje, decir que 'tienen' ética. Decir que la máquina 'se comporta' éticamente, es mucho, pero no tiene sentido decir que tiene o no tiene ética; hay innumerables situaciones imprevisibles en las que no se pueden prever respuestas preprogramadas.

De los ejemplos anteriores, ¿se deduce que las máquinas pueden tener sentimientos? Siempre hay que tener en cuenta que las máquinas, y la IA en particular, no hacen nada por sí solas; sólo hacen lo previsto por las personas que los conciben, diseñan, programan e implementan, y estas pueden dotarlas con la apariencia de sentimientos lo cual es muchos contextos resulta muy beneficioso. Por otra parte, es conveniente, como hace López de Mántaras, considerar la afirmación de 2020 de E. Bender y A. Koller: «Crear tecnología que imite a los humanos, haciéndose pasar por algo humano, requiere que tengamos muy claro que significa ser humano ya que de lo contrario corremos el riesgo de deshumanizarnos».

Como conclusión, podemos llegar a percibir en las máquinas comportamientos con sentimientos y éticamente correctos; aunque no tiene sentido decir que la máquina, en su esencia, dispone de ellos; a no ser que se verificase la sentencia del famoso título de la obra de teatro de Pirandello: «Así es, sí así os parece».

Alberto Prieto es miembro de la Academia de Ciencias de Granada. Profesor Emérito de la Universidad de Granada. Departamento de Ingeniería de Computadores, Automática y Robótica