

de nuestras vidas. De hecho, el control parece ser el tema común, tanto para evitar la frustración, la alienación y la irritación que experimentamos hacia las tecnologías contemporáneas, como para permitirnos interactuar con los demás de una manera fiable o para mantener los vínculos que nos unen con nuestros familiares, nuestras amistades y nuestros colegas.

No todas las interacciones tienen que producirse a tiempo real, con sus participantes, siempre asequibles y disponibles, a punto de responder, interrumpiéndose unos a otros. Las tecnologías que permiten almacenar y luego responder, como, por ejemplo, el correo electrónico y el mensaje de voz, es decir, que permiten dar salida a los mensajes cuando le vaya bien a quien los envía, y a su vez, que su receptor los escuche, lea o responda también cuando más cómodo le sea. Nos hacen falta formas de combinar entre sí los métodos de comunicación distintos, de modo que se pudiera escoger el correo normal, el electrónico, el teléfono, la voz o el texto en función de lo que exija cada ocasión. Tenemos asimismo necesidad de reservarnos un tiempo para concentrarnos sin interrupciones, durante el cual podamos dedicarnos plenamente a una cosa.

La mayoría ya lo hacemos. Apagamos los móviles y deliberadamente a veces no los llevamos encima. Revisamos y filtramos las llamadas recibidas, y no respondemos a menos que veamos —o escuchemos— que la ha hecho alguien con quien realmente queremos hablar. Nos vamos a lugares privados, para escribir o pensar mejor, o simplemente para relajarnos.

Hoy por hoy, las tecnologías luchan para asegurarse su omnipresencia, de modo que estén disponibles con independencia del lugar en que nos hallemos o de lo que estemos haciendo. Esto es perfecto mientras la decisión sobre si son utilizadas o no recaiga en el individuo que se halla en el extremo receptor. Tengo una gran fe en la sociedad y estoy convencido de que alcanzaremos una solución sensata de convivencia con estas tecnologías. En los primeros años de la implementación de cualquier tecnología, los inconvenientes demasiado evidentes igualan a sus potenciales aplicaciones, generando aquella relación de amor-odio que tan habitual resulta en el despliegue de las nuevas tecnologías. De amor hacia su potencial y de odio hacia la realidad de los hechos. Pero, con el tiempo, a través de un diseño mejorado de la tecnología, así como del modo en que es empleada, es posible minimizar aquel odio y transformarlo en una relación de amor.

## Capítulo 6

### Máquinas emocionales

*Dave, no continúes. No continúes. No continúes, Dave. ¿Continuarás Dave?... Dave, no continúes... Tengo miedo. Tengo miedo... Tengo miedo, Dave... Dave... No puedo pensar con claridad... Lo siento... Lo siento... No puedo pensar con claridad... No cabe duda... Lo siento... Lo siento... tengo... miedo.<sup>1</sup>*

HAL, el omnipotente ordenador de la película  
2001: Una odisea del espacio

HAL está en lo cierto al sentir miedo: Dave está a punto de desconectarlo desmantelando sus partes. Sin duda, Dave también tiene miedo: HAL ha matado a todo el resto de la tripulación de la nave espacial y ha hecho un intento, infructuoso, de atentar contra la vida de Dave.

Pero ¿por qué tiene miedo y cómo puede sentirlo? ¿Es un miedo real? Diría que no. HAL determina de modo correcto cuál es la intención de Dave: quiere matarle. Así, el miedo —estar asustado— es la reacción lógica a la situación. Pero las emociones humanas tienen algo más que un componente lógico, racional, están fuertemente asociadas con la conducta y los sentimientos. Si HAL fuese un ser humano, lucharía con todas sus fuerzas para evitar su propia muerte, cerraría de golpe algunas escotillas, haría algo para sobrevivir. Podría amenazar diciendo «si me matas, tú también morirás cuando se te acabe el oxígeno de tu cámara de aire». Pero HAL no hace ninguna de estas cosas, se limita a enunciarlas, como un hecho: «Tengo miedo». HAL sabe intelectualmente qué significa tener miedo, pero es algo que no se asocia con sentimientos o acciones: no es una emoción real.

Pero ¿por qué iba HAL a necesitar emociones reales para funcionar? Nuestras máquinas actuales no necesitan emociones. Ciertamente tienen

1. Pasaje del filme 2001: Una odisea del espacio, citado por Bizoni (1994), pág. 60.

una cantidad razonable de inteligencia. Pero ¿emociones? Ninguna. Sin embargo, las máquinas del futuro necesitarán tener emociones por la misma razón que los seres humanos las tenemos: el sistema emocional humano desempeña un papel esencial en la lucha por la supervivencia, en la interacción y cooperación sociales, y en la adquisición de conocimientos. Las máquinas, cuando se enfrenten a las mismas condiciones a las que nos enfrentamos los humanos, cuando tengan que funcionar de manera continua sin contar con la ayuda de los seres humanos en un mundo complejo y en constante cambio en el que constantemente surgen nuevas situaciones, necesitarán poder contar con una forma de emoción, una emoción de máquina. A medida que las máquinas son más capaces, asumiendo muchas de las funciones que antes desarrollábamos los seres humanos, los diseñadores se enfrentan a la compleja tarea de decidir el modo de construirlas, el modo en que interactuarán con otras máquinas y con los seres humanos. Así, por la misma razón por la que los animales y los seres humanos tenemos emociones, opino que las máquinas también las necesitarán. No serán emociones como las humanas, conviene tenerlo presente, sino más bien emociones adecuadas a las necesidades que tienen las máquinas como tales.

Hoy en día ya existen robots. Si bien en su mayoría se trata de simples brazos automatizados e instrumentos emplazados en fábricas, están mejorando su potencia y capacidades, diversificando la gama de sus actividades y emplazamientos. Algunos hace trabajos útiles, como, por ejemplo, los robots, ya en circulación, que cortan el césped o pasan la aspiradora. Algunos, como los que sustituyen a los animales de compañía, son divertidos. Hay robots sencillos que realizan trabajos peligrosos, como pueden ser la extinción de incendios y misiones de búsqueda y rescate, o con fines militares. Algunos incluso se encargan de entregar el correo, dispensar medicamentos o realizar otras tareas relativamente sencillas. A medida que se perfeccionen y avancen, los robots tendrán necesidad de las emociones más sencillas, empezando por aquellas que son prácticas como, por ejemplo, el miedo visceral a las alturas o la preocupación por no chocar con otros objetos. Los robots de compañía tendrán personalidades alegres y simpáticas. Con el tiempo, a medida que estos robots vayan adquiriendo mayores capacidades, acabarán teniendo emociones en toda regla: miedo y ansiedad cuando se encuentren en situaciones peligrosas; placer cuando alcancen una meta deseada, satisfacción por la calidad del trabajo realizado; y sumisión y obediencia hacia sus propietarios. Dado que muchos de estos robots serán operativos en el entorno

doméstico, la interacción con las personas y otros robots de la casa hará necesario que exterioricen sus emociones y tengan características análogas a lo que son las expresiones faciales y el lenguaje corporal.

Las expresiones faciales y el lenguaje corporal forman parte de la «imagen del sistema» de un robot, el cual permite que las personas con las cuales interactúa tengan un mejor modelo conceptual de su modo de funcionar. Cuando interactuamos con otras personas, las expresiones faciales que muestran y su lenguaje corporal nos permiten saber si nos están entendiendo, si están desconcertadas o se sienten perplejas, y si están de acuerdo o no. La expresión facial de una persona nos permite decir si está teniendo dificultades. El mismo tipo de respuesta no verbal tendrá un valor inestimable en nuestra interacción con el robot: ¿ha entendido las instrucciones? ¿Se está empleando a fondo en una tarea? ¿Lo ha conseguido? ¿Tiene dificultades? Las expresiones emocionales nos permitirán saber cuáles son sus motivaciones y deseos, sus consecuciones y frustraciones y, por tanto, acrecentarán nuestra satisfacción y comprensión de los robots: estaremos en condiciones de estimar de qué son capaces y de qué no.

Dar con la combinación justa de emociones e inteligencia no es cosa fácil. Los dos robots que aparecen en *La guerra de las galaxias*, R2D2 y C3PO, actúan como máquinas que nos gustaría tener en casa. Diría que parte de su encanto estriba en el modo en que muestran sus limitaciones a la mirada del público. C3PO es un zoquete patoso pero bien intencionado, muy incompetente en todas las tareas salvo en aquella en la que es todo un especialista: traducir idiomas y comunicaciones entre máquinas. R2D2 está diseñado para interactuar con otras máquinas y tiene capacidades físicas limitadas. Para hablar con seres humanos tiene que apoyarse en C3PO.

R2D2 y C3PO exteriorizan sus emociones, haciendo que los personajes de la pantalla —y el público que ve el filme— entiendan, simpaticen y, a veces, se enfaden con ellos. C3PO tiene una forma humana, la cual le permite mostrar expresiones faciales y realizar movimientos con el cuerpo: por ejemplo, se contornea y balancea mucho. R2D2 es más limitado, aunque, no obstante, es muy expresivo, con lo cual evidencia nuestra habilidad para atribuir emociones cuando vemos «hablar» a una cabeza que se mueve, un cuerpo que se desplaza hacia delante y hacia atrás, y escuchamos algunos sonidos bonitos aunque nos resulten ininteligibles. Gracias a la destreza de los creadores del filme, los modelos conceptuales que subyacen a R2D2 y C3PO son bastante evidentes. Así, el público siempre puede comprender de una forma bastante precisa sus puntos fuertes y sus debilidades, lo cual hace de ellos seres agradables y eficaces.

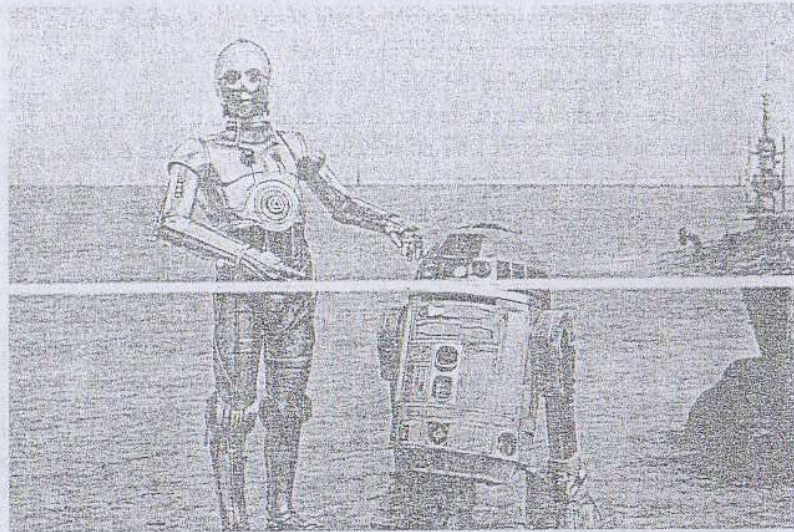


FIGURA 6.1. C3PO (izquierda) y R2D2 (derecha), La guerra de las galaxias. Ambos son muy expresivos a pesar de que R2D2 no tiene cuerpo ni estructura facial. (Cortesía de Lucasfilm Ltd.)<sup>2</sup>

Los robots que aparecen en los filmes no siempre salen tan bien parados. Pensemos, por ejemplo, en lo que les sucedía a dos robots que aparecían en el cine: a HAL en *2001: Una odisea del espacio*, y a David, en *A. I. (Inteligencia artificial)*. HAL, tiene miedo, como lo muestra la cita con la que abríamos el presente capítulo, y era justo que lo tuviera, pues estaba siendo desmantelado y, en esencia, asesinado.

David es un robot construido para sustituir a la figura de un hijo, para que ocupe el lugar de un hijo de carne y hueso en un hogar. David es una máquina sofisticada, pero demasiado perfecta. Según la historia, David es el primer robot en sentir «un amor incondicional». Sin embargo, no se trata de un amor real, ya que, quizá por ser «incondicional», parece artificial, demasiado fuerte y no va acompañado por la gama humana normal de estados emocionales. Los hijos normales puede que amen a sus padres, pero también pasan por etapas en las que sienten aversión, enojo, envidia, en las que sienten asco por ellos o simple y llanamente les son indiferen-

2. *La guerra de las galaxias. Episodio IV - Una nueva esperanza* © 1977 y 1997. Lucasfilm Ltd & TM. Derechos reservados. Reproducido con autorización. Cualquier reproducción no autorizada será considerada una violación de la ley vigente.

tes. David no exterioriza ninguno de estos sentimientos. Aquel amor puro que siente, hace que David ande pegado, como un abnegado y feliz niño, a su madre, de un modo casi literal, cada minuto del día. La conducta de David acaba siendo tan irritante que es abandonado por su madre adoptiva en pleno bosque diciéndole que no vuelva.

El papel que la emoción desempeña en una inteligencia avanzada es un tema estándar de la ciencia ficción. Así, dos de los personajes que intervienen en la serie televisiva *Star Trek*, así como en las versiones cinematográficas que se han hecho de ella, lidian continuamente con el papel que se debe atribuir a la emoción y a la inteligencia. El primero de estos personajes, Spock, hijo de madre humana pero de padre vulcaniano, se caracteriza por no tener emociones, lo cual iba a brindar a los guionistas de la serie la maravillosa oportunidad de enfrentar la razón pura de Spock con las emociones humanas del capitán Kirk. De manera análoga, en las series *Star Trek, nueva generación* producidas bastante tiempo después, el capitán Data es un androide puro, un ser artificial, cuya falta de emociones brinda a los guionistas de la nueva serie una carta blanca similar para hacer lo mismo, si bien en varios episodios Data es objeto de una serie de ajustes y retoques, entre los que se incluye la posibilidad de añadirle un «chip emocional», como si la emoción fuese algo aparte del cerebro que pudiese añadirse o quitarse a voluntad. Si bien la serie es pura ficción, lo cierto es que los guionistas hicieron un trabajo de investigación muy bueno, porque el retrato que elaboraron del papel de la emoción a la hora de tomar decisiones y en la interacción social resulta lo bastante razonable para que los psicólogos Robert Sekuler y Randolph Blake lo consideraran un excelente ejemplo del fenómeno, apropiado para sus cursos de iniciación a la psicología. En su libro, *Star Trek on the Brain*,<sup>3</sup> se sirven de numerosos ejemplos sacados de la serie homónima para ilustrar el papel que la emoción tiene en el comportamiento y la conducta (entre otros muchos temas).

#### OBJETOS EMOCIONALES

¿Conseguiría mi tostadora hacer mejor las tostadas como yo las prefiero si no estuviera dotada de cierto orgullo? Las máquinas no llegarán a ser sutiles y sensibles hasta que tengan a la vez inteligencia y emociones. Las emociones nos permiten traducir la inteligencia en acción.

3. Sekuler y Blake, 1998.

Si no experimentáramos orgullo por nuestros actos, ¿por qué íbamos a esforzarnos en mejorar? Las emociones positivas tienen una importancia decisiva para el aprendizaje, ya que mantienen nuestra curiosidad por el mundo. Las emociones negativas nos evitan caer en peligros, pero son las emociones positivas las que hacen que la vida sea valiosa, las que nos guían hacia las cosas buenas de la vida, las que recompensan nuestros éxitos y las que nos hacen esforzarnos para hacerlo mejor.

La razón pura no siempre basta. ¿Qué sucede cuando no hay bastante información? ¿De qué modo decidimos el curso de acción que hay que seguir si hay riesgo, de modo que en él la emoción del triunfo sea sopesada con la posibilidad de padecer daños? En este punto las emociones desempeñan un papel crítico y aquí es donde experimentan dificultades los seres humanos que han padecido daños neuronales en sus sistemas emocionales. En la película *2001: Una odisea del espacio*, el astronauta Dave arriesga su vida cuando sale al espacio exterior para recuperar el cadáver de su compañero de viaje espacial. A nivel lógico, no tiene mucho sentido, pero en términos de una dilatada historia de la sociedad humana, reviste gran importancia. En realidad, esta tendencia de los seres humanos a arriesgar la propia vida en un intento por rescatar a unos pocos —o incluso para recuperar a los que ya han muerto—, es un tema constante tanto de nuestras vidas reales como de las que se desarrollan en el plano de la ficción, ya sea en la literatura, el teatro o el cine.

Los robots necesitan tener algo similar a lo que es la emoción en los seres humanos para tomar estas complejas decisiones. ¿Resistirá la pasarela el peso del robot? ¿Hay algún peligro acechando detrás del poste? Estas decisiones exigen ir más allá de la información perceptiva para emplear la experiencia y el conocimiento general a fin de realizar inferencias y deducciones acerca del mundo y, luego, emplear el sistema emocional para que ayude a evaluar la situación y saber de qué modo actuar. Si sólo recurrimos a la pura lógica, podemos pasarnos todo el día paralizados sin movernos, incapaces de hacerlo mientras consideramos todo aquello que podría salir mal, tal como les sucede a las personas que tienen problemas emocionales.<sup>4</sup> Para tomar estas decisiones necesitamos las emociones: y los robots también las necesitarán.

Los sistemas afectivos, ricos y estratificados, similares a aquellos de los que disponemos los seres humanos, aún no forman parte de las máquinas, pero un día lo serán. Pero, sin duda, el afecto requerido no tiene por

4. Damasio, 1994, 1999.

qué ser una copia del que tenemos los seres humanos. Más bien, lo que se requiere es un sistema afectivo que esté perfectamente en sintonía con las necesidades que presenta el sistema. Los robots, por ejemplo, tendrían que poder preocuparse por los peligros a los que pudieran exponerse, muchos de los cuales también los han de sortear los seres humanos y los animales, y algunos de estos peligros, en cambio, son específicos para los robots. Por ejemplo, deben evitar caer por las escaleras o despeñarse, por lo cual deberían sentir miedo a las alturas. Deberían sentir cansancio, para evitar que se les agotaran las baterías y se quedaran sin energía (¿sensación de hambre?) antes de recargar sus baterías. No tienen necesidad de comer ni de ir al baño, pero, en cambio, sí que necesitarán que se les haga una revisión periódica: que se les engrasen las articulaciones, se les cambien las partes estropeadas y demás. Los robots, por ejemplo, no tienen que preocuparse por su aseo ni por las condiciones de salubridad, pero, en cambio, les es necesario estar atentos a la suciedad que puede entrar en sus partes móviles, el polvo y la suciedad que se acumula en las lentes de televisión, así como de los virus informáticos que pudieran interferir en su funcionamiento. El afecto que los robots requieren será a la vez similar y muy diferente al que tenemos los seres humanos.

Aunque los diseñadores de máquinas puede que nunca hayan pensado que estaban dotando de afecto o emoción a sus máquinas, han construido sistemas de seguridad y supervivencia. Algunos de estos sistemas se asemejan al nivel visceral de que disponemos los seres humanos: circuitos sencillos y de rápida acción que detectan el peligro posible y reaccionan en consecuencia. Dicho de otro modo, la supervivencia ya ha sido incorporada como una parte más en la mayoría de diseños de máquinas. Muchos aparatos llevan fusibles de modo que sí, de repente, reciben más corriente eléctrica de lo normal, el fusible o el interruptor abren el circuito eléctrico, evitando de este modo que la máquina resulte dañada (y, con ello, impidiendo que haga daño a sus usuarios o al entorno). De manera similar, algunos ordenadores disponen de su propia fuente de suministro de energía, de modo que si la corriente cae, de inmediato y sin dilación pasan a alimentarse con la energía de la batería, la cual permite que puedan apagarse de un modo apropiado, guardando todos los datos y enviando avisos a los operadores humanos. Algunos equipos cuentan con sensores térmicos y de nivel de agua. Otros detectan la presencia de personas y se niegan a funcionar cuando alguien se halla en una zona de acceso prohibido. Los robots existentes y otros sistemas móviles disponen ya de sensores y de sistemas visuales que les evitan chocar con personas y otros

objetos o caer por las escaleras. Este nivel tan simple de seguridad y supervivencia forma ya parte en la actualidad de muchos aparatos diseñados.

En los seres humanos y los animales, el impacto del sistema visceral no interrumpe, sin embargo, su funcionamiento tras una respuesta o reacción iniciales. El nivel visceral señala niveles superiores de procesamiento con el fin de establecer las causas del problema y determinar una respuesta efectiva. Las máquinas deberían hacer lo mismo.

Cualquier sistema autónomo, es decir, que trate de existir por sí mismo, sin alguien que siempre se encargue de guiarlo, tiene que decidir constantemente cuál, entre las muchas actividades posibles, escoge. En términos técnicos, se necesita un sistema de organización planificada, aunque es una tarea que incluso a los seres humanos les resulta difícil realizar. Cuando, por ejemplo, estamos trabajando a fondo para terminar una tarea importante, ¿en qué momento debemos hacer una pausa para comer, dormir y hacer otras actividades que quizá se nos exigen pero que no son ni con mucho tan urgentes? ¿De qué modo encajar las muchas actividades que las que realizar en el limitado espacio de tiempo de un día, sabiendo cuándo, por ejemplo, hay que dejar una actividad de lado y cuándo no? Y lo que es aún más importante: ¿cuál es la propuesta decisiva que debe presentarse el día siguiente por la mañana o la próxima celebración familiar de aniversario que es preciso planificar? Se trata de problemas difíciles que hoy por hoy ninguna máquina ni siquiera llega a contemplar, aunque los seres humanos lidiarnos con ellos a diario. Éstos son precisamente los tipos de problemas de toma de decisiones y de control en los que resulta de tanta ayuda el sistema emocional.

Muchas máquinas se diseñan para que funcionen aun en el caso de que algunos de sus componentes individuales puedan fallar. Este comportamiento resulta decisivo en los sistemas relacionados con la seguridad, como es el caso, por ejemplo, de los aviones y los reactores de las centrales nucleares, y asimismo es un comportamiento muy valioso en los sistemas que realizan operaciones decisivas, como, por ejemplo, algunos sistemas informáticos, hospitales y cualquier cosa relacionada con la infraestructura vital de una sociedad. Pero ¿qué sucede cuando un componente falla e intervienen los dispositivos de seguridad automáticos de recambio? En este contexto es donde el sistema afectivo se muestra útil.

El fallo de los componentes se habría detectado en el nivel visceral y se habría usado para desencadenar una señal de alerta o alarma: en lo esencial, el sistema se «preocuparía». El resultado de esta preocupación creciente haría que la máquina actuara de un modo más conservador, qui-

zás aminorando su ritmo de funcionamiento y posponiendo tareas consideradas como no esenciales. Dicho de otro modo, ¿por qué las máquinas no iban a comportarse como si fuesen seres humanos preocupados? Se mostrarían cautas aun cuando trataran de eliminar la causa de su preocupación. En el caso de los seres humanos, el comportamiento pasa a ser más centrado hasta que se determinan tanto la causa como la respuesta apropiada. Con independencia de cuál sea la reacción prevista para los sistemas mecánicos, es necesario introducir algún cambio respecto al que es su comportamiento normal.

Los animales y los seres humanos han desarrollado mecanismos sofisticados para sobrevivir en un mundo dinámico e impredecible, uniendo las valoraciones y las evaluaciones derivadas del afecto con métodos para modular el sistema en su conjunto. El resultado es una solidez mayor y una tolerancia del error. Los sistemas artificiales que diseñamos harían bien en aprender del ejemplo que los sistemas vivos les ofrecen.

#### ROBOTS EMOCIONALES

*La década de 1980 fue la del ordenador personal, la de 1990, la de Internet, pero soy de la opinión que la década que acaba de empezar será la del robot.*

Palabras de un ejecutivo de la corporación Sony<sup>5</sup>

Supongamos que queremos construir un robot capaz de vivir en casa, que deambule por pasillos y habitaciones, que encaje de manera cómoda en la familia: ¿qué haría? Cuando se plantea esta pregunta, la mayoría, de entrada, pensamos en que se le podría encargar que hiciera todas las tareas diarias. El robot sería una suerte de criado, que limpiaría la casa y haría otros quehaceres diarios. A todos nos gustaría un robot que lavara los platos o hiciera la colada. En realidad, si bien los lavaplatos, las lavadoras y las secadoras automáticas se podrían considerar como robots de un tipo muy sencillo y dedicados a un propósito en especial, lo que la mayoría de nosotros tiene en mente, sin embargo, es una máquina que iría de aquí para allá por la casa y recogería lo que estuviera sucio, platos o ropa, lo separaría de manera ordenada y lo lavaría. Una vez limpios los devolvería a los lugares donde se guardan, después de haber planchado y plegado la ropa limpia.

5. Toshitada Doi, presidente de Sony Digital Creatures Laboratory (noviembre de 2000).

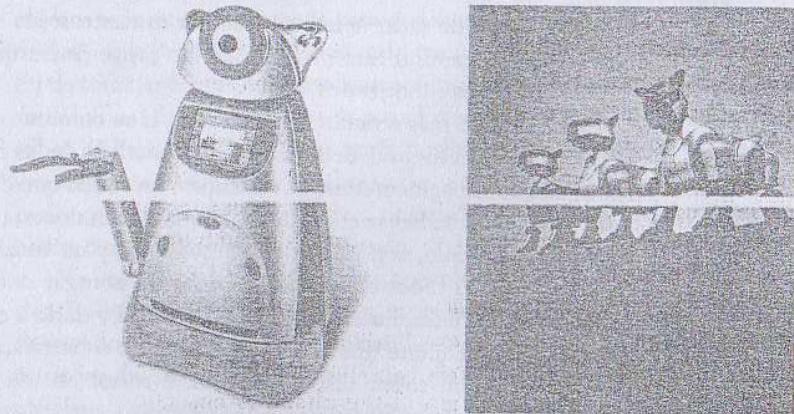


FIGURA 6.2 a y b. *Robots caseros de principios del siglo XXI.* La figura a muestra a ER2, un prototipo de robot casero desarrollado por Evolution Robotics. La figura b es el perro robot mascota Aibo de Sony. (La imagen de ER2 ha sido reproducida por cortesía de Evolution Robotics. La imagen de «Tres Aibós en la pared», por cortesía de Sony Electronics Inc., Entertainment America, Robot Division.)

Todas estas tareas son bastante difíciles, y exceden con mucho a las posibilidades que ofrecen las primeras y escasas generaciones de robots.

En la actualidad, los robots aún no son objetos domésticos. Aparecen en las ferias de la ciencia y en las plantas industriales, en misiones de rescate y otros casos que requieren especialización. Pero todo esto cambiará. La corporación Sony ha anunciado que esta década será la del robot y, aunque puede que Sony sea demasiado optimista, por mi parte preveo que los robots empezarán a florecer y a proliferar durante la primera mitad del siglo XXI.

Los robots adoptarán multiplicidad de formas diversas. Puedo imaginarme una familia de aparatos robots en la cocina —robots refrigeradores, robots alacenas, robots cafeteras, robots capaces de cocinar y lavar los platos—, todos ellos configurados de modo que puedan comunicarse entre sí y llevar y traer, de un lado a otro, la comida, los platos y demás utensilios. El robot sirviente deambula por la casa, recogiendo los platos sucios y llevándoselos al robot lavaplatos, que a su vez entrega los platos y los utensilios una vez limpios al robot que se encarga de la alacena, el cual los almacena y guarda hasta que alguien o un robot los vuelve a precisar. Los robots al cargo de la alacena, de la nevera y de la cocina trabajan con sol-

tura preparando el menú del día y, por último, colocan la comida completa en los platos que les facilita el robot alacena.

Ciertos robots cuidarán de los niños, jugando con ellos, leyéndoles cuentos e historias, y cantándoles canciones. Los juguetes didácticos ya hacen algo de esto, y el robot, más sofisticado, podría actuar como un profesor cualificado que les enseñara el alfabeto, les enseñara a leer y aritmética, para pasar luego a abordar cualquier tipo de temas. La novela de ciencia ficción escrita por Neal Stephenson, *La era del diamante*,<sup>6</sup> logra mostrar de manera magnífica cómo un libro interactivo, *Manual ilustrado para jovencitas*, puede ocuparse de la educación completa de las niñas desde los 4 años hasta que son adultas. El manual ilustrado como el de esta novela queda mucho más distante en el futuro, pero en la actualidad existen ya profesores, aunque con facultades algo más limitadas. Además de la educación, algunos robots se encargarán también de hacer las faenas de casa: pasar la aspiradora, quitar el polvo y hacer la limpieza general de la casa. Por último, su gama de habilidades se ampliará. Puede que algunos acaben incorporándose a las viviendas o dentro de los muebles. Otros, en cambio, tendrán movilidad, y serán capaces de deambular solos por su cuenta.

Estos avances requerirán un proceso coevolutivo de adaptación tanto de los seres humanos como de los aparatos. Se trata de un proceso habitual en nuestras tecnologías: reconfiguramos el modo en que vivimos y trabajamos a fin de hacer posible la operatividad de las máquinas de que disponemos. La coevolución más espectacular es la protagonizada por el sistema del automóvil, que ha supuesto la modificación de nuestras viviendas a fin de incorporar garajes y calzadas con las dimensiones y todo lo necesario para que los automóviles circulen, así como la construcción de un enorme sistema de autopistas a escala mundial, sistemas de señalización del tráfico, aceras peatonales y enormes zonas destinadas a aparcamientos. Las viviendas, además, han sido transformadas a fin de acomodar las múltiples tuberías y los cables que forman parte de la cada vez mayor infraestructura de la vida moderna: agua caliente y fría, desagües, respiraderos de ventilación, conductos de calefacción y refrigeración, electricidad, teléfono, televisión, Internet y ordenadores personales, así como sistemas de entretenimiento. Las puertas de nuestras casas tienen que ser lo bastante anchas para permitir que entren los muebles y muchas viviendas deben acondicionarse para permitir el desplazamiento con sillas de ruedas así como de las personas, mayores o no, que necesitan de caminadores

6. Stephenson, 1995.

para hacerlo. Y justo cuando hemos acondicionado la vivienda para dar cabida a todos estos cambios, preveo que deberemos efectuar nuevas modificaciones para dar cabida a los robots. Se trata de una modificación que se producirá, sin duda, de manera lenta, pero que, conforme los robots tengan cada vez mayor utilidad, optaremos por asegurar el éxito de su implantación minimizando los obstáculos y, finalmente, construyendo estaciones de recarga, espacios dedicados a su limpieza y mantenimiento, y demás cosas. Al fin y al cabo, el robot que pase la aspiradora necesitará un lugar en el cual vaciar la suciedad que haya recogido, y el robot encargado de la basura tendrá que poder sacarla al exterior de la vivienda. No me sorprendería que en las casas acabara habiendo zonas dedicadas a los robots, es decir, nichos especialmente contruidos en los cuales puedan residir, zonas en las que puedan estar, sin ser perturbados, cuando no estén activos. Si hoy disponemos de armarios y alacenas en los que guardamos los aparatos que utilizamos, por qué no íbamos a disponer de espacios especialmente equipados para los robots, con puertas que ellos puedan controlar, puntos de toma de electricidad, luces interiores de modo que puedan ocuparse de su propia limpieza (y enchufarse a las tomas de electricidad), así como receptáculos para los desperdicios allí donde sean adecuados.

Los robots, sobre todo en un primer momento, probablemente requieran que los suelos tengan una superficie regular, que estén libres de obstáculos. Quizá se eliminen los umbrales en las puertas o se reduzcan a su mínima expresión. Algunos lugares —sobre todo las escaleras— quizá se señalicen de manera especial, quizá con luces, transmisores de infrarrojos o simplemente con una cinta reflectante especial. Códigos de barras o marcadores distintivos colocados aquí y allí en la vivienda contribuirán quizás a que los robots reconozcan de una manera muy sencilla el lugar donde están.

Pensemos, por ejemplo, cómo un robot camarero o criado servirá una bebida a su dueño. Si éste le pide una lata de un refresco, el solícito robot irá derecho a la cocina y buscará la nevera, donde se guardan los refrescos. El hecho de comprender lo que se le pide y navegar hasta la nevera es una cosa relativamente sencilla. Averiguar cómo abrir la puerta, encontrar la lata y sacarla de la nevera no es tan sencillo. Conferir al robot criado la destreza, la fuerza y unas ruedas antideslizantes que le permitan abrir tirando de la puerta de la nevera es casi una proeza. Dotarlo con un sistema de visión que le permita encontrar el refresco buscado, sobre todo si está medio oculto detrás de otros paquetes con comida, es difícil y, luego, ave-

riguar el modo de sacar la lata sin romper ni destruir los objetos que encuentre en su camino está más allá de las capacidades que actualmente tienen los brazos robotizados.

Sería mucho más sencillo, sin embargo, si se contara con un robot especializado en dispensar bebidas, hecho a medida de las necesidades del robot criado. Imaginémoslo, por ejemplo, un robot que dispensara bebidas y fue capaz de llevar entre media y una docena entera de latas, en un espacio refrigerado, con una puerta automática y un brazo que le permitiera sacarlas. El robot criado podría dirigirse al lugar donde estuviera el robot de las bebidas, hacerle notar su presencia y darle a conocer su petición (tal vez mediante una señal de infrarrojos o hertziana), y colocar la bandeja frente al robot dispensador. Este robot abriría sin dilación la puerta, sacaría con su brazo una lata y volvería a cerrar la puerta: así no se necesitaría un sistema de visión compleja, ni un brazo diestro, ni sería preciso abrir enérgicamente la puerta. La lata sería colocada en la bandeja del robot criado, el cual, luego, volvería al lugar donde estuviese su dueño.

De un modo similar, podríamos modificar el lavaplatos para hacer más fácil al robot de la casa la operación de cargarlo con los platos sucios, tal vez dotándolo de bandejas especiales con ranuras diseñadas para colocar diferentes tipos de platos. Pero mientras hacemos esto, ¿por qué no convertir la alacena en un robot especializado, capaz de sacar los platos limpios del lavaplatos y guardarlos para su uso posterior? El hecho de disponer de bandejas especiales sería de ayuda también en la alacena, que, tal vez, podría dar automáticamente las tazas a la cafetera y los platos al robot cocinero, el cual, sin duda, estaría conectado con la nevera, el fregadero y el dispositivo de recolección de basuras. ¿Todo esto parece exagerado? Tal vez, pero, de hecho, los electrodomésticos actuales son ya complejos, muchos de ellos tienen múltiples conexiones con servicios. Las neveras tienen conexiones con puntos de toma de electricidad y de agua. Algunas ya están conectadas a Internet. Los lavaplatos y las lavadoras tienen conexiones con la toma de electricidad y agua, así como con las tuberías de desagüe. El hecho de integrar estas unidades de modo que puedan funcionar sin problemas unas con otras no parece de hecho tan difícil.

Imagino que una casa contendrá una serie de robots especializados: el criado es tal vez el robot más genérico, pero trabajaría conjuntamente con un robot encargado de la limpieza, con el que dispensa las bebidas, tal vez con algunos robots de exterior encargados del jardín y una familia de robots de cocina, como el lavaplatos, la cafetera y el robot alacena. A medida que estos robots se desarrollen, probablemente diseñaremos también

objetos especializados para la casa que ayuden a simplificar las tareas asumidas por los robots, evolucionando conjuntamente los robots y la casa a fin de trabajar juntos de manera armoniosa. Démonos cuenta de que el resultado final también será mejor para los seres humanos. De este modo, el robot dispensador de bebidas haría posible que cualquiera se le acercara y pidiera una lata, y, salvo en el caso de que se quisiera emplear la comunicación por medio de infrarrojos u ondas de radio, bastaría con apretar un botón o quizá simplemente pedírsela.

No soy el único en pensar esta coevolución de los robots y las viviendas. Rodney Brooks,<sup>7</sup> uno de los especialistas actuales en robótica más destacados a escala mundial, actual director del Artificial Intelligence Laboratory del MIT y fundador de una empresa que construye robots domésticos y comerciales, imagina una rica ecología de entornos y robots, con robots especializados incorporados en aparatos, cada uno responsable de mantener limpio su ámbito de actividad: uno se encarga de limpiar la bañera, otro del aseo; uno limpia las ventanas, otro deja a punto los espejos. Brooks contempla incluso la posibilidad de un robot que haga las veces de mesa del comedor, con una zona para almacenamiento y lavaplatos incorporados en su base, de modo que cuando «queramos poner la mesa, los pequeños brazos robotizados, de un modo similar a como funciona una máquina de discos, pongan todos los platos que son precisos y la cubertería en los lugares donde deben ir. Conforme se vayan terminando los distintos platos de la comida, la mesa y sus pequeños brazos robotizados retirarán los platos y los engullirán en el amplio volumen interno situado debajo de la parte que hace las veces de mesa».

¿Qué aspecto debería tener un robot? Los robots que aparecen en el cine a menudo se asemejan a las personas, tienen dos piernas, dos brazos y una cabeza. Pero ¿por qué? La forma debería derivar de la función. El hecho de que tengamos dos piernas nos permite a los seres humanos movernos por terrenos irregulares, algo que a un animal sobre ruedas le resultaría imposible hacer. El hecho de que tengamos dos manos nos permite levantar y manipular cosas, una mano ayuda a la otra. La forma humanoide ha ido evolucionando a lo largo de millones de años de interacción con el mundo para enfrentarse de un modo más eficiente y a la vez más eficaz con ese entorno. De este modo, si las tareas que se le exigen a un robot son similares a las que recaen en los seres humanos, no es una insensatez pensar que debería tener también una forma similar.

7. Brooks, 2002. La cita reproducida procede de la pág. 125.

Si, en cambio, los robots no tienen que moverse —como en el caso de los robots dispensadores de bebidas, lavaplatos o alacena—, tampoco precisan tener medios de locomoción, ni piernas ni ruedas. Si el robot es una cafetera, debe parecer una cafetera, pero modificada a fin de que se conecte con el lavaplatos y la alacena. En la actualidad ya existen robots aspiradoras y cortadores de césped, y su aspecto externo se halla perfectamente adecuado a las tareas que realizan: son aparatos pequeños y rechonchos que tienen ruedas (véase la figura 6.3). Un robot autónomo debe parecer un coche. Sería apropiado que sólo los robots que hagan de criados domésticos, al tener una función general, se asemejaran a animales o a seres humanos. Aquel robot en la mesa del comedor en el cual pensaba Brooks sería particularmente extraño, con una amplia columna central en la cual se colocarían los platos y todo lo necesario para funcionar como lavaplatos (junto con las tomas de electricidad, agua y desagüe). La parte de arriba de la mesa tendría espacio para que los brazos del robot manipularan los platos, así como un pedúnculo en el que se colocarían las cámaras que permitirían a los brazos saber dónde colocar los platos y de dónde retirar los platos y los cubiertos.

¿Un robot debería estar dotado de piernas? No si tiene que moverse y maniobrar en superficies regulares, en este caso bastaría con que dispusiera

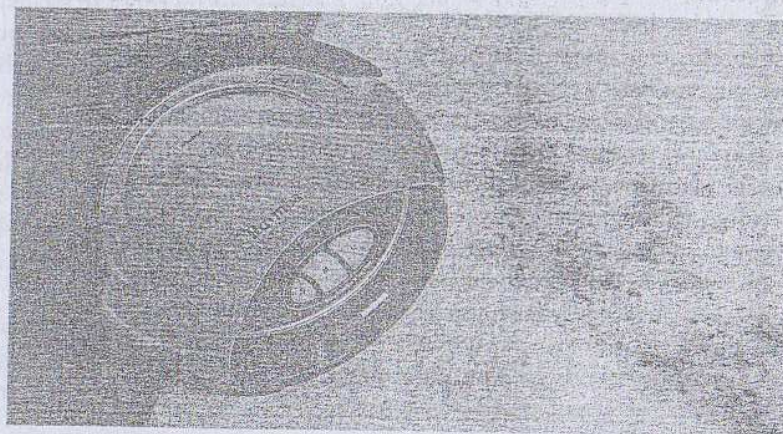


FIGURA 6.3. ¿Qué aspecto debería tener un robot? Roomba es una aspiradora, su forma es apropiada para que pueda desplazarse por el suelo y maniobrar por sí sola bajo los muebles. Este robot no parece ni un ser humano ni un animal, ni tiene por qué: la forma que tiene es la idónea para la tarea que realiza. (Cortesía de iRobot Inc.)



de ruedas. En cambio, si debe desplazarse por un terreno irregular o subir escaleras, tener piernas le resultaría de utilidad. En este caso, cabe esperar que la primera generación de robots con esta característica tenga entre cuatro y seis piernas, ya que el equilibrio es mucho más sencillo en el caso de las criaturas que tienen cuatro o seis piernas que en el de las bípedas.

Si un robot ha de moverse por una vivienda e ir delante y detrás de quienes allí viven, su aspecto probablemente será similar al de un animal o una persona: un cuerpo en el que se colocan las baterías y que sostiene las piernas, ruedas o vías que utilice para la locomoción; manos con las que recoja los objetos; y cámaras (ojos) en la parte superior allí donde mejor puedan examinar el entorno. Dicho de otro modo, algunos robots se asemejarán a un animal o a un ser humano no porque esa forma sea bella, sino porque es la configuración más efectiva para la tarea que deben realizar. Estos robots probablemente se asemejarán a R2D2 (figura 6.1): un cuerpo cilíndrico o rectangular colocado sobre algunas ruedas, vías o piernas; cierta forma de brazo o bandeja manipulables; y sensores repartidos por todo su volumen a fin de poder detectar los obstáculos, escaleras, personas, animales domésticos, otros robots y, desde luego, los objetos con que se supone que debe interactuar. Salvo que los robots estuvieran diseñados para servir sólo de entretenimiento y diversión, resulta difícil comprender por qué íbamos a querer que un robot se asemejara a C3PO.

De hecho, hacer que un robot tenga aspecto humano puede resultar contraproducente, puesto que podría hacerlo menos aceptable. Masahiro Mori,<sup>8</sup> un especialista en robótica japonés, sostiene que somos menos propensos a aceptar a las criaturas cuyo aspecto es muy similar al humano y tienen un mal rendimiento, un concepto que ha quedado ilustrado en el cine y en el teatro a través de la naturaleza aterradora de los zombies y los monstruos (pensemos en el caso de Frankenstein, por ejemplo) que adoptan una forma humana, pero cuyos movimientos y horrible aspecto físico no son humanos. En cambio no nos azoran —ni asustan— tanto, ni mucho menos, las formas y las figuras no humanas. Incluso las réplicas perfectas de los seres humanos podrían resultar problemáticas, porque aun en el caso de que el robot no pueda distinguirse del ser humano, esa mis-

8. *The Buddha in the Robot* (Mori, 1982). El argumento de que nos sentimos más molestos si el robot tiene un aspecto muy cercano al humano proviene de un ensayo de Dave Bryant (Bryant, s. f.). Bryant atribuye el argumento a Mori, pero cuando compré el libro de Mori y lo leí, aunque disfruté haciéndolo, no encontré ni rastro de esta argumentación. No obstante, es una cuestión de suma importancia.

ma falta de rasgos distintivos puede ser fuente de angustia emocional (un tema explorado en más de una novela de ciencia ficción y, de manera especial, en *¿Sueñan los androides con ovejas eléctricas?*, de Philip K. Dick<sup>9</sup> y, en su versión cinematográfica, *Blade Runner*). Según esta línea de argumentación, C3PO, pese a tener forma humanoide, parece más bonito e incluso más molesto que amenazador por lo torpe y patoso que es, tanto en su ademán como en la manera que tiene de comportarse.

Los robots que cubren necesidades humanas, por ejemplo, como los robots que hacen de mascotas sustitutivas, probablemente deben tener un aspecto similar al de las criaturas vivas, aunque sólo sea para incidir en nuestro sistema visceral, un sistema incorporado que nos permite interpretar el lenguaje corporal humano y animal, así como las expresiones faciales. De este modo, si se quiere que el robot interactúe de manera satisfactoria con los seres humanos resultará más efectivo que adopte una figura similar a la de un animal o rasgos ingenuos, así como que tenga los movimientos corporales, las expresiones faciales y emita los sonidos que sean apropiados.

#### EL AFECTO Y LA EMOCIÓN EN LOS ROBOTS

Llegados a este punto cabe preguntarse por cuáles son las emociones que necesitará tener un robot. La respuesta dependerá del tipo de robot en que estemos pensando, las tareas que deba realizar, la naturaleza del entorno y cuál sea su vida social. Por ejemplo, si tiene que interactuar con otros robots, animales o seres humanos, será necesario que exprese su propio estado emocional, así como que sepa evaluar las emociones de los seres humanos y de los animales con los que interactúe.

Pensemos, ahora, en el robot doméstico medio, el de todos los días. Si bien aún no existe como tal, algún día nuestra casa acabará estando poblada de este tipo de robots. Algunos de ellos estarán fijos en un sitio, especializados en una función, como los robots de cocina, entre los cuales cabe citar, por ejemplo, los robots alacena, lavaplatos, dispensadores de bebidas y de comidas, cafeteras o la unidad que se encargue de cocinar. Y, sin duda, los robots que laven, sequen, planchen y plieguen la ropa, tal vez complementados por robots roperos. Algunos, en cambio, serán móviles, aunque también estarán especializados, como los encargados de pasar la aspiradora por el suelo de las habitaciones y segar el césped. Pero,

9. Dick, 1968.

quizás, asimismo tengamos como mínimo un robot para funciones generales: el robot criado doméstico, que nos servirá el café, limpiará, realizará recados sencillos y cuidará y supervisará el quehacer del resto de robots. Éste es el robot doméstico que mayor interés reviste, porque deberá ser el más flexible y avanzado.

Estos robots criados tendrán que interactuar de manera necesaria con los seres humanos y con los demás robots de la casa. En el caso de su interacción con el resto de robots podrían emplear una comunicación de tipo inalámbrico. Podrían hablar de las tareas que estuvieran haciendo, sobre si iban sobrecargados de trabajo o bien si estaban sin nada que hacer. También podrían avisar cuando se estuvieran quedando sin provisiones o cuando se enfrentan a dificultades, problemas o errores y llamar a otro robot para que les ayudara. Pero ¿qué sucede cuando los robots interactúan con los seres humanos? ¿De qué modo se producirá esta interacción?

Los robots criados deberían ser capaces de comunicarse con sus dueños. Tendrían que disponer de un modo para expresar las órdenes, es decir, un modo de clarificar las ambigüedades, cambiar una orden mientras están realizando una tarea («olvida el café y tráeme en su lugar un vaso con agua») y estar familiarizados con todas las complejidades del lenguaje humano. En la actualidad, no nos es posible conseguirlo, de modo que los robots que construimos hoy deben basarse en órdenes muy sencillas o incluso en cierto tipo de control remoto, en el cual una persona pulsa los botones adecuados, genera una orden bien estructurada o selecciona las acciones a realizar entre las que figuran en un menú. Pero llegará el tiempo en que podamos interactuar a través del habla, en que los robots entiendan no sólo las palabras, sino el significado que ellas expresan.

¿Los robots deberían ofrecerse a ayudar a sus amos? De ser así, deberían ser capaces de evaluar el estado emocional de los seres humanos. En el caso de que alguien pasara apuros cuando trata de hacer una tarea, el robot podría ofrecerse a ayudarlo. Si hay personas que discuten, por ejemplo, el robot podría desear salirse de en medio e ir a otra habitación. Si algo resultara placentero, el robot podría recordarlo y así lo volvería a reproducir cuando fuese oportuno hacerlo. Si una tarea se realizó mal y decepcionó a una persona, quizás el trabajo podría perfeccionarse, de modo que el robot obtuviera mejores resultados cuando lo hiciera de nuevo en otra ocasión. Por todas estas razones y otras muchas más, el robot tendrá que ser diseñado con la capacidad para interpretar el estado emocional de sus dueños.

Un robot deberá tener ojos y oídos (cámaras y micrófonos) a fin de leer las expresiones faciales, el lenguaje corporal y los componentes emocionales del habla. Tendrá que ser sensible a los tonos de voz, al ritmo del habla, a su modulación, de tal modo que pueda reconocer cuando alguien se siente enojado o se siente deleitado, frustrado o alegre. El robot tendrá que estar en condiciones de reconocer las voces que le amonestan y diferenciarlas de las que escuchan cuando lo elogian. Conviene observar que todos estos estados se pueden reconocer precisamente por medio de la calidad del sonido sin necesidad de tener que reconocer las palabras o la manera de hablar. También conviene tener presente que podemos determinar los estados emocionales de otras personas basándonos sólo en el tono de la voz. Por ejemplo, probemos a hacer que los demás crean que nos sentimos enojados, alegres, que les reprendemos o elogiamos, y expresémoslos manteniendo los labios firmemente cerrados. Lo podemos hacer recurriendo sólo a los sonidos, sin decir ni una palabra. Se trata de pautas de sonido universales.

De manera similar, el robot tendría que mostrar cuál es su estado emocional, del mismo modo en que lo hace una persona (o quizá, para ser más exactos, como lo hace un perro doméstico o un niño), de modo que las personas con las que interactúa pudieran decir si ha entendido lo que le piden, cuando algo le resulta fácil o difícil de hacer, o quizá también cuando el robot estima que es inapropiado. De manera similar, el robot debería mostrar, siempre que sea oportuno, placer o displacer, un aspecto enérgico o cansado, confianza o preocupación. Si se quedara encallado sin poder acabar una tarea, debería hacer ostensible su frustración. Para el robot será tan valioso mostrar cuál es su estado emocional como lo es para una persona hacerlo. Las expresiones del robot nos permitirán a los seres humanos comprender el estado en el que el robot se halla y saber así qué tareas le resultan apropiadas y cuáles no. En consecuencia, estaremos en condiciones de clarificar las instrucciones que le demos o incluso de ofrecerle ayuda, aprendiendo al final a sacar un mejor partido de las capacidades que tenga el robot.

Muchas personas en el mundo de la robótica y la comunidad que se dedica a la investigación informática creen que el modo de expresar emociones consiste en hacer que un robot decida entre si está contento o triste, enojado o disgustado, y en función de ello muestre el rostro apropiado, por lo general, una parodia exagerada de la expresión facial que una persona suele adoptar cuando se encuentra en esos mismos estados. Por mi parte me opongo decididamente a esta forma de enfocar las cosas. Los

seres humanos no actuamos de este modo. No decidimos, primero, que nos sentimos contentos para, luego, poner cara de contentos, al menos no es así normalmente. Nos comportamos así cuando lo que nos proponemos es engañar a alguien. Pero pensemos en todos esos profesionales que están obligados a sonreír sin que importe la circunstancia en que se encuentren: no engañan a nadie, parece simplemente que fuerzan la sonrisa, como en realidad están haciendo.

Los seres humanos damos forma a nuestras expresiones faciales mediante la innervación automática de un gran número de músculos que participan en el proceso de control de nuestro rostro y cuerpo. Los afectos positivos conducen a la relajación de algunos grupos de músculos, lo que hace que muchos músculos faciales se expandan de forma automática hacia arriba (de ahí la sonrisa, las cejas y las mejillas levantadas, etc.), y nos hace ser proclives a abrimos y acercarnos al objeto o hecho positivo. Un afecto negativo tiene el efecto contrario, haciendo que nos retiremos o apartemos de golpe. Algunos músculos se contraen y parte de los músculos faciales se tensan hacia abajo (de ahí el ceño fruncido). Casi todos los estados afectivos son complejas combinaciones de valencias positivas y negativas, con diferentes niveles de excitación, con cierto poso dejado por los estados inmediatamente previos. Las expresiones resultantes son ricas y llenas de información, además de reales.

Las emociones falsas parecen falsas: somos muy buenos cuando se trata de detectar los intentos postizos que lo que buscan es manipularnos. Así, muchos de los sistemas informáticos con los que interactuamos —nos referimos a aquellos que tienen ayudas simpáticas, sonrientes, y voces y expresiones que suenan a artificialmente suavizadoras— tienden a ser más molestos que útiles. «¿Cómo puedo apagarlo?» es una pregunta que suelo hacerme y he llegado a ser un experto en dejarlos inhabilitados, tanto en los ordenadores que son de mi propiedad, como en los de otras personas que tratan de liberarse de todas esas irritantes molestias.

He defendido que las máquinas deberían tener y mostrar emociones, para que podamos interactuar mejor con ellas. Por esta razón precisamente las emociones deberán parecer tan naturales y habituales como las emociones humanas. Tendrán que ser reales, un reflejo directo de los estados internos y del estado de procesamiento de un robot. Necesitamos saber cuándo un robot confía o cuándo está confuso, seguro o preocupado, cuándo entiende lo que le pedimos y cuándo no, cuándo trabaja en lo que le hemos pedido que haga y cuándo nos ignora. Si las expresiones faciales y corporales reflejan el estado de procesamiento subyacente, en-

tonces las manifestaciones emocionales parecerán genuinas precisamente porque serán reales. Entonces podremos interpretar su estado, y los robots podrán interpretar el nuestro de modo que la comunicación y la interacción fluirán de manera aún menos problemática.

Pero no soy el único que ha llegado a esta conclusión. La profesora del MIT Rosalind Picard en cierta ocasión, mientras disertaba sobre si los robots debían tener emociones, afirmó que «no estuve segura de que debían tener emociones hasta que me puse a escribir un artículo sobre cómo responderían de manera inteligente a nuestras emociones en el caso de que los robots no las tuvieran. Mientras escribía aquel artículo, caí en la cuenta de que sería muchísimo más fácil si los dotáramos de emociones».<sup>10</sup>

Una vez que los robots dispongan de emociones, deberán ser capaces de mostrarlas y hacerlo de un modo que los seres humanos seamos capaces de interpretar, es decir, como un lenguaje corporal y unas expresiones faciales similares a las de los seres humanos. De este modo, el rostro y el cuerpo del robot deberían disponer de actuadores internos que actuaran y reaccionaran del mismo modo en que lo hacen los músculos humanos, aunque, en su caso, lo harían en función de los estados internos de los robots. En el mentón, los labios, las ventanas de la nariz, las cejas, la frente, las mejillas y demás partes de nuestro rostro, los seres humanos disponemos de potentes grupos de músculos. Este complejo muscular forma un sofisticado sistema de señalización, de modo que si los robots fueran creados de una manera similar, las características del rostro reproducirían de forma natural la sonrisa cuando las cosas salieran bien y, en cambio, cuando surgieran dificultades fruncirían el ceño. Para conseguirlo, los diseñadores tienen que estudiar y comprender el complejo funcionamiento de las expresiones humanas, así como del riquísimo conjunto de músculos y ligamentos que se hallan fuertemente entrelazados con el sistema afectivo.

Mostrar emociones faciales completas es en realidad una empresa muy difícil. En la figura 6.4 se muestra a Leonardo, un robot creado por la profesora Cynthia Breazeal, en el Media Laboratory del MIT, y diseñado para controlar una amplísima gama de rasgos faciales, así como los movimientos del cuello, los brazos y el cuerpo, para una mejor interacción social y emocional con los seres humanos. Mucho es lo que sucede

10. La cita de Picard es recogida por Cavelos (1999, págs. 107-108), y me lo reiteró durante la visita que hice a su laboratorio en 2002.

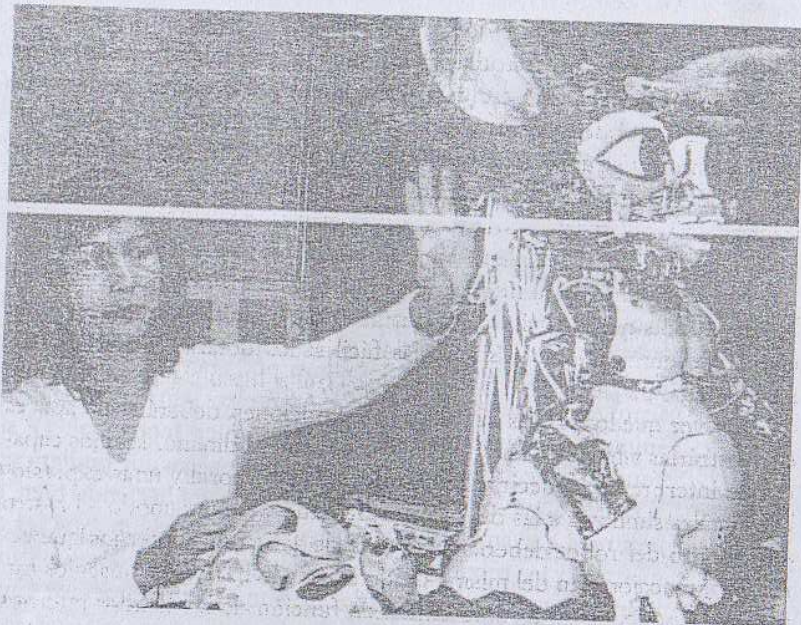


Figura 6.4. La complejidad de la musculatura facial del robot. La profesora del MIT Cynthia Breazeal con su robot Leonardo. (Fotografía del autor.)

en el interior de nuestros cuerpos y ello hace que sea necesaria en buena medida la misma complejidad en el interior de los rostros con que dotamos a los robots.

Pero ¿y los estados emocionales subyacentes? ¿Cómo deberían ser? Tal como expuse anteriormente, el robot debería, cuando menos, tener cuidado con las alturas, debería ser precavido con los objetos calientes y sensible a las situaciones que pudieran causar daño o dolor. El miedo, la preocupación, el dolor y la infelicidad podrían ser, todos ellos, estados emocionales apropiados para un robot. De manera similar, el robot debería tener estados positivos, entre ellos las sensaciones de placer, satisfacción, gratitud, alegría y orgullo, que le permitieran aprender de sus actos, repetir los positivos y mejorar, siempre que fuese posible y donde pudiese hacerlo.

La sorpresa es probablemente una emoción esencial. Cuando lo que sucede no es lo que se esperaba, el robot sorprendido interpretaría lo sucedido como una señal de alarma. Si la habitación se quedase, de improvi-

so, a oscuras, o quizás el robot chocara contra algo con lo que no contaba, una respuesta prudente sería la de detener todo movimiento y tratar de averiguar por qué había sucedido tal cosa. La sorpresa significa que una situación no ha sido anticipada y que el comportamiento planeado o actual probablemente ha dejado de ser el adecuado, de ahí la necesidad de parar y proceder a evaluarlo de nuevo.

Algunos estados, como, por ejemplo, el cansancio, el dolor o el hambre son más sencillos, dado que lejos de precisar previsiones o predicciones, de hecho lo que requieren es efectuar un simple seguimiento de los sensores internos. (La fatiga y el hambre no son, desde un punto de vista técnico, estados afectivos, aunque se pueden tratar como si lo fuesen.) En los seres humanos, los sensores de estados físicos señalan cansancio, hambre o dolor. En realidad, en nosotros, el dolor es un sistema asombrosamente complejo, que aún no hemos llegado a comprender del todo. Existen millones de receptores del dolor, además de una amplia variedad de centros en el cerebro que intervienen en la interpretación de las señales. A veces intensificando la susceptibilidad, a veces suprimiéndola. El dolor hace las veces de un valioso sistema de alarma, que nos impide hacernos daño y, cuando estamos heridos, actúa como un recordatorio para no agravar aún más el daño en las partes perjudicadas. Con el tiempo podría ser útil que los robots sintieran dolor cuando, por ejemplo, los motores o las juntas estuvieran dañadas. Esta sensación haría que los robots limitaran automáticamente sus actividades y, de este modo, se protegieran de ulteriores daños.

La frustración sería un afecto útil, porque permitiría evitar, por ejemplo, que los robots encargados del servicio quedasen parados realizando una tarea y descuidando sus otras obligaciones. Y creo que podría funcionar del siguiente modo. Por ejemplo, le pido al robot criado que me traiga una taza de café. Para ello se dirige hacia la cocina, donde el robot del café le explica que no hay café porque no quedan tazas limpias. Entonces la cafetera podría pedirle al robot alacena que le pasara más tazas; pero supongamos que a este último robot, tampoco le quedarán. El robot alacena debería pasar la petición al robot lavaplatos. Pero, ahora, supongamos que el lavaplatos no tuviera tampoco platos sucios que limpiar. El lavaplatos pediría al robot criado que recogiera tazas sucias para poderlas lavar, primero, y pasarlas luego a la alacena, de donde irían a la cafetera, que a su vez pondría el café que el robot criado podría finalmente servir. Por desgracia, el criado declinaría acceder a la petición cursada por el lavaplatos para que se diera una vuelta por la casa y recogiera los platos su-

cios: el robot criado aún estaría ocupado con su principal tarea, a saber, esperar que la cafetera sirva el café.

Esta situación se denomina de punto muerto o *impasse*. En el caso expuesto, nada se puede hacer porque cada máquina espera que la siguiente haga algo y la última espera a que lo haga la primera. Este problema particular podría resolverse dotando a los robots de una inteligencia cada vez mayor, de modo que aprendieran la forma de resolver cada nuevo problema que surgiera, pero los problemas siempre surgen de manera más rápida de lo que los diseñadores son capaces de anticipar. Estas situaciones de *impasse* resultan difíciles de eliminar porque cada una de ellas surge de un conjunto diferente de circunstancias. La frustración aporta una solución de tipo general.

La frustración es un afecto útil tanto para los seres humanos como para las máquinas, ya que cuando las cosas llegan a ese punto, es hora de dejarlas y hacer otras. El robot que hace las veces de criado puede frustrarse esperando el café, de modo que provisionalmente se daría por vencido. Tan pronto como el robot criado dejara de ir a buscar café, quedaría libre para atender la petición del lavaplatos, salir y recoger las tazas de café sucias. Con ello se resolvería de manera automática el *impasse* creado: el robot criado encontraría algunas tazas sucias, las colocaría en el lavaplatos, lo cual, finalmente, llevaría a que la cafetera hiciera el café y el robot criado acabara sirviéndomelo, aunque, eso sí, con cierto retraso.

¿El robot criado podría aprender de esta experiencia? Para ello debería añadir a su lista de actividades la recogida periódica de los platos sucios, de modo que el robot lavaplatos y el alacena no volvieran a quedarse sin existencias. En este punto es donde cierta dosis de orgullo vendría de perillas. Sin el orgullo, el robot no tiene cuidado: no tiene incentivo para aprender a hacer las cosas mejor. En una situación ideal, el robot se sentirá orgulloso de evitar dificultades, de no quedarse parado por el mismo problema en más de una ocasión. Esta actitud requiere que los robots tengan emociones positivas, emociones que les hagan sentir bien consigo mismos, que les hagan hacer cada vez mejor sus trabajos, perfeccionarse y, quizás, ofrecerse para realizar nuevas tareas, aprender nuevos modos de hacer las cosas. El orgullo de hacer un buen trabajo, de complacer a sus dueños.

## MÁQUINAS QUE SIENTEN EMOCIÓN

*La medida en que las alteraciones emocionales pueden interferir en la vida mental no le viene de nuevo a los maestros y profesores. Los estudiantes que están preocupados, enojados o deprimidos no aprenden; las personas que quedan atrapadas en estos estados no asimilan de manera eficiente la información ni saben tratarla bien.*

DANIEL GOLEMAN,  
*Inteligencia emocional*<sup>11</sup>

Supongamos que las máquinas pudieran percibir las emociones que sentimos los seres humanos. ¿Y si fuesen tan sensibles a los estados de ánimo de sus usuarios como lo es un buen terapeuta? ¿Y si un sistema educativo electrónico y controlado por ordenador pudiera sentir cuando la persona que está aprendiendo lo estuviera haciendo bien, si se sintiera frustrada o avanzase de manera adecuada? O bien ¿y si los aparatos domésticos y los robots del futuro pudieran cambiar las operaciones que realizan en función de los estados de humor de sus dueños? Entonces, ¿qué harían?

La profesora Rosalind Picard<sup>12</sup> del Media Laboratory del MIT dirige un programa de investigación titulado «Affective Computing», que constituye un intento de desarrollar máquinas capaces de sentir las emociones de los seres humanos con quienes interactúan y, por tanto, de responder en consonancia. El grupo de investigación que dirige ha realizado un notable avance en el desarrollo de dispositivos de medida, capaces de sentir miedo y preocupación, tristeza y aflicción, así como satisfacción y alegría. La figura 6.5, que se halla publicada en su página de Internet, demuestra la variedad de cuestiones que es preciso abordar.

¿De qué modo se perciben las emociones de alguien? El cuerpo exhibe su estado emocional en una variedad de maneras. Hay, sin duda, expresiones faciales y lenguaje corporal. ¿Los seres humanos controlamos nuestras expresiones? Lo cierto es que sí, pero la capa visceral funciona de manera automática y aunque los niveles conductual y reflexivo tratan de inhibir la reacción visceral, la supresión completa no parece ser posible. Aun en el caso de la persona que mejor sabe controlarse, aquello

11. Goleman, 1995. La cita procede de Kort, Reilly y Picard, 2001.

12. Picard, 1997.

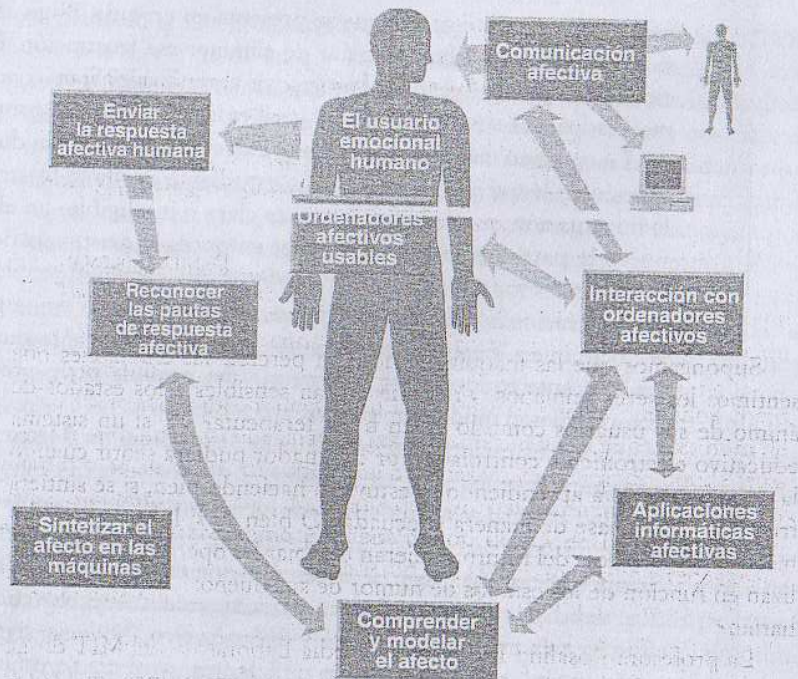


FIGURA 6.5. El programa «Affective Computing». El diagrama indica la complejidad del sistema afectivo de los seres humanos y los desafíos que supone el intento por controlar de manera adecuada el afecto. Este esquema proviene del trabajo que la profesora Rosalind Picard lleva a cabo en el MIT. (Dibujo cortesía de Rox Picard y Jonathan Klein.)

que damos en llamar «la cara de póquer», que mantiene una exposición neutra de respuestas emocionales con independencia de la situación en la que el sujeto se halle, todavía exterioriza microexpresiones, esto es, expresiones breves y fugaces que los observadores expertos son capaces de detectar.<sup>13</sup>

Además de las respuestas de nuestros músculos, existen otras muchas respuestas fisiológicas. Por ejemplo, aunque la intensidad de la luz afecte al tamaño de la pupila en el ojo, las dimensiones de la pupila son un indi-

13. La investigación fundamental realizada por Paul Ekman (Ekman, 1982, 2003). Una excelente descripción divulgativa es la del artículo publicado por Malcolm Gladwell en la revista *New Yorker* (Gladwell, 5 de agosto de 2002).

cador de la excitación emocional en la que un sujeto se halla. Cuando alguien se siente interesado o excitado emocionalmente, esto hace que su pupila se dilate. Cuando se trabaja a fondo un problema, la pupila se contrae. Estas respuestas son involuntarias, lo cual hace que sea difícil —y probablemente imposible— que una persona llegue a controlarlas. Una de las razones por las que los jugadores profesionales de cartas a menudo llevan guías oscuras o de colores incluso en habitaciones poco iluminadas no es otra que para evitar que sus rivales detecten los cambios en el tamaño de sus pupilas.

El ritmo cardíaco, la presión sanguínea, el ritmo de la respiración y el sudor son medidas habituales que se emplean para inferir el estado afectivo de un sujeto. Incluso cantidades de sudor tan pequeñas que el sujeto puede ser inconsciente de ellas, pueden indicar un cambio en la conductividad eléctrica de la piel. Todas estas medidas se pueden detectar fácilmente mediante el uso de aparatos electrónicos apropiados.

El problema consiste en que estas sencillas manifestaciones fisiológicas son medidas indirectas del afecto. Cada uno se ve afectado por toda una serie de cosas, y no sólo por la emoción o el afecto. En consecuencia, aunque estas medidas se utilizan en muchos entornos clínicos y aplicados, deben ser interpretadas con sumo cuidado. Por ejemplo, consideremos el funcionamiento de lo que se ha dado en llamar el detector de mentiras. Un detector de este tipo es, en todo caso, un detector de emociones. El método se denomina en términos técnicos «prueba del polígrafo» porque funciona registrando de manera simultánea y representa gráficamente múltiples medidas fisiológicas, como el ritmo cardíaco, el ritmo de la respiración y la conductancia eléctrica de la piel. Un detector de mentiras no es un detector de falsedades, lo que detecta es la respuesta afectiva que una persona da a una serie de preguntas que le hace el examinador, en un proceso en el cual algunas de las respuestas se consideran verdaderas (y, por tanto, muestran respuestas afectivas bajas) y otras falsas (y, por tanto, muestran una excitación afectiva alta). Resulta sencillo ver la razón por la cual los detectores de mentiras son tan controvertidos. Los sujetos inocentes pueden dar respuestas emocionales muy intensas a preguntas decisivas, en tanto que los sujetos culpables puede que no muestren ninguna reacción cuando se les plantean las mismas preguntas.

Quienes manejan de forma experta los detectores de mentiras tratan de compensar estas dificultades recurriendo al empleo de preguntas de control que permiten calibrar las reacciones de un sujeto. Por ejemplo, al plantearles una pregunta ante la cual prevén que responderá con una men-

tira, pero que no es relevante respecto a la cuestión que está siendo tratada, los operadores del detector pueden ver qué aspecto gráfico se perfila cuando la persona que está siendo examinada responde con una mentira. Esto se realiza entrevistando al sospechoso y luego desarrollando una serie de preguntas destinadas a descubrir el comportamiento normal desviado, en el cual el examinador no está interesado, pero en el cual el sospechoso es probable que mienta. Una pregunta que suele utilizarse habitualmente en Estados Unidos es: «¿Robó alguna cosa cuando era adolescente?».

Dado que los detectores de mentiras registran los estados fisiológicos subyacentes, los cuales están más asociados con las emociones que con las mentiras, no son unos dispositivos muy fiables y no sólo presentan omisiones y fallos (cuando, por ejemplo, una mentira no es detectada porque no produce ninguna reacción emocional), sino que también generan falsas alarmas (cuando se da el caso de que el nerviosismo del sospechoso produce respuestas emocionales aunque el sujeto en cuestión no sea culpable). Los expertos en el manejo de estas máquinas son conscientes de los escollos y algunos utilizan la prueba del detector de mentiras como un medio para provocar una confesión voluntaria: los sujetos que realmente creen que el detector de mentiras puede leer la mente pueden acabar confesando simplemente por el miedo que tienen a someterse a la prueba. Cuando tuve oportunidad de hablar con expertos en el manejo del detector de mentiras, no tuvieron reparos en reconocer que estaban de acuerdo con la crítica que acabo de presentar, pero se sentían orgullosos por el número récord de confesiones voluntarias que habían conseguido. Con todo, es obligado reconocer que personas inocentes también se habían declarado a veces autores de crímenes que no habían cometido, por extraño que pueda parecer. El nivel de precisión que aportan estos detectores es tan imperfecto que el National Research Council de las National Academies de Estados Unidos llevó a cabo un prolongado y completo estudio y llegó a la conclusión de que la prueba del polígrafo era demasiado imperfecta como para utilizarla en asuntos legales y de investigación policial.<sup>14</sup>

¿Qué haríamos si, pongamos por caso, pudiéramos detectar cuál es el estado emocional de una persona? ¿Qué reacción deberíamos tener? Se trata de un problema de primer orden que aún no ha sido resuelto. Considere-

14. *Review the Scientific Evidence on the Polygraph*, 2002.

mos una situación como la que se presenta en un aula. Si un estudiante se siente frustrado, ¿debemos tratar de eliminar esa frustración, o esa sensación es una parte necesaria del proceso de aprendizaje? Si un conductor automovilístico está tenso y estresado, ¿cuál es la reacción más apropiada?

La reacción más apropiada a una emoción depende, sin duda, de la situación en la que uno se halle, si un estudiante se siente frustrado porque la información que se le facilita no es clara o inteligible, un elemento importante para el profesor es conocer entonces la existencia de ese estado de frustración; es de esperar del profesor que corrija el problema a través de explicaciones adicionales. (Por mi experiencia, sin embargo, debo reconocer que a menudo no es así, porque un profesor que genera tal frustración, ante todo es, por lo general, una persona poco preparada para comprender el modo de remediar el problema.)

Si la frustración se debe a la complejidad del problema, entonces la respuesta apropiada de un maestro o profesor sería la de no hacer nada. Es algo normal y apropiado que los estudiantes se sientan frustrados cuando tratan de resolver problemas que exceden ligeramente sus capacidades, o cuando tratan de hacer algo que nunca antes habían intentado. En realidad, si los estudiantes no se sienten de vez en cuando frustrados, lo más probable es que sea algo negativo, es decir, significa que no están asumiendo unos riesgos suficientes, que no se están esforzando bastante.

Con todo, probablemente sea bueno tranquilizar a los estudiantes que se sienten frustrados, explicarles que cierta dosis de frustración no sólo es apropiada, sino también necesaria. Se trata de una clase de frustración buena que lleva a mejorar y a aprender. Si se prolonga demasiado en el tiempo, sin embargo, la frustración puede hacer que los estudiantes se den por vencidos y decidan que el problema rebasa sus capacidades. Llegados a este punto es cuando es necesario aconsejarles, darles explicaciones individuales o aplicar otro tipo de orientación.

¿Y qué sucede con aquel tipo de frustraciones que muestran los estudiantes y que nada tienen que ver con lo que se expone en el aula, y que podrían resultar de cierta experiencia personal que se realiza fuera del entorno docente? En este contexto resulta menos evidente saber qué se debe hacer. El profesor, ya sea una persona o una máquina, no está en condiciones de ser un buen terapeuta. Expresar simpatía podría ser la respuesta tanto más como menos acertada.

Las máquinas que pueden sentir las emociones constituyen una nueva frontera emergente en el campo de investigación, una frontera en la que

se plantean tantas preguntas como problemas se tratan de abordar, tanto con relación al modo en que las máquinas deberían de detectar las emociones como con relación al modo de determinar la forma más adecuada de darles respuesta. Conviene no perder de vista que si bien nos esforzamos en determinar de qué modo hacer que las máquinas respondan y reaccionen de manera apropiada a los signos de las emociones, los seres humanos no somos tampoco un ejemplo excelente en ninguno de estos campos. Muchos experimentamos grandes dificultades cuando tenemos que responder de forma adecuada a otras personas que están pasando por una situación de aflicción emocional y, a veces, cuando tratamos de ayudarlas no hacemos más que agravar el problema. Y muchos, además, se muestran asombrosamente insensibles ante los estados emocionales de los demás, incluso cuando se trata de personas a las que conocen bien. Es lógico que las personas que se hallan bajo una fuerte presión emocional traten de ocultarlo, y la mayoría no somos precisamente expertos en detectar los signos y las señales de las emociones.

Pese a ello, se trata de un ámbito de investigación importante. Y aunque nunca llegáramos a crear máquinas capaces de responder y reaccionar de una manera plenamente adecuada, la investigación nos permitirá conocer más a fondo tanto las emociones humanas como la interacción entre los seres humanos y las máquinas.

#### MÁQUINAS QUE INDUCEN EMOCIONES EN LOS SERES HUMANOS

Resulta sorprendentemente fácil conseguir que los seres humanos tengan experiencias emocionales intensas incluso con los sistemas informáticos más sencillos. Tal vez la primera de este tipo de experiencias fue la que se realizó con ELIZA, un programa informático que desarrolló el investigador en computación del MIT, Joseph Weizenbaum.<sup>15</sup> ELIZA era un programa sencillo que funcionaba siguiendo una pequeña serie de guiones de conversación que habían sido preparados de antemano por el programador (quien, en un principio, fue el propio Weizenbaum). Al seguir estos guiones, ELIZA podía interactuar con una persona sobre cualquiera de los temas para los que en el guión la hubiera preparado. Valga, como muestra, el siguiente ejemplo. Cuando uno lo ponía en marcha, el

15. El trabajo sobre ELIZA se realizó en la década de 1960 y es examinado en el libro de Weizenbaum (1976).

programa saludaba al usuario diciendo: «Hola, soy ELIZA. ¿En qué puedo ayudarte?». Si se respondía escribiendo: «Me preocupa el creciente nivel de violencia en el mundo», ELIZA contestaba: «¿Cuánto tiempo hace que le preocupa el creciente nivel de violencia en el mundo?». Se trata de una pregunta importante, así que una respuesta lógica que daríamos sería algo similar a: «Sólo desde hace unos pocos meses», a lo cual ELIZA respondería: «Continúe, por favor».

Podemos ver de qué modo la conversación acabaría captando al usuario: las preocupaciones que mostrase recibirían respuestas comprensivas. Pero ELIZA no comprendía para nada lo que se le decía. Se limitaba a encontrar pautas y respondía de manera adecuada (diciendo: «Continúe, por favor», cuando sencillamente no reconocía la pauta). Así, por ejemplo, era fácil engañar a ELIZA escribiendo: «Me preocupa abc, def y, en realidad, ghi», a lo cual ELIZA respondía de manera diligente preguntando: «¿Cuánto tiempo hace que le preocupa abc, def y, en realidad, ghi?».

ELIZA simplemente reconocía la frase: «Me preocupa X» y respondía: «¿Cuánto tiempo hace que le preocupa X?», sin entender para nada las palabras.

Dado que la mayoría se tomaba en serio a ELIZA, no trataban de engañar al programa. En su lugar, se tomaban en serio cada respuesta que daba el programa, y sopesaban sus significados ocultos. En consecuencia, a veces acababan enzarzados en profundos debates filosóficos o, en la versión más divulgada de ELIZA, programada para actuar como un psicoanalista, acababan hablando de sus propios secretos más íntimos y personales.

Hacia mediados de la década de 1960, uno de mis amigos y colaboradores de investigación más duraderos, Daniel Bobrow, trabajaba como investigador en BBN (Bolt, Barenek y Newman), una empresa de Cambridge (Massachusetts) dedicada a la investigación en el campo de la inteligencia artificial, las redes de ordenadores y la acústica. Bobrow me dijo que, un programador había terminado, después de haber trabajado a fondo para hacer que ELIZA funcionase en un ordenador de BBN, a eso de las 5 de la madrugada de un sábado y que se había marchado a casa, dejando el programa funcionando en el ordenador. (En aquellos años —estábamos en 1966—, conviene recordarlo, no existían ordenadores personales y sólo las empresas más avanzadas, como BBN, tenían ordenadores para el uso exclusivo de sus empleados, una de las razones por las cuales los programadores tenían que hacer el turno de noche para tener acceso a las máquinas.)



Aquel mismo sábado, algo más tarde por la mañana, el vicepresidente de la división en la cual trabajaba Bobrow se desplazó a la empresa para mostrar el ordenador a algunos clientes. Al ver que el ordenador estaba en funcionamiento, supuso que Bobrow lo estaba utilizando desde casa y, tal como me lo contó el propio Bobrow, «se decidió a utilizar aquel enlace de comunicación “evidente” que quién sí no había dejado abierto» para chatear con mi amigo. Así que el vicepresidente se puso a teclear algunas preguntas destinadas a Bobrow. Pero Bobrow estaba en su casa, dormido, y lo que empezó a suceder en realidad estaba pasando entre el programa ELIZA y el vicepresidente:

- VP: ¿Cree que puedo utilizar el ordenador hoy por la mañana?  
 ELIZA: ¿Por qué lo pregunta?  
 VP: Es que tendría que hacer algunas ventas adicionales de Telcom.  
 ELIZA: ¿Por qué no está tan seguro?  
 VP: Mis posibles clientes nunca han presenciado una demostración del sistema.  
 ELIZA: ¿Significa mucho para usted?  
 VP: Desde luego que sí.  
 ELIZA: Parece muy seguro de ello.  
 VP: Es ridículo.  
 ELIZA: ¿Qué quiere decir con «es ridículo»?  
 VP: ¿Hay algún motivo por el que no pueda utilizar el ordenador ahora?  
 ELIZA: Cuénteme más acerca de sus posibles clientes.  
 VP: Llámeme por favor al 491-1850.<sup>16</sup>

En aquellas fechas, los ordenadores no eran muy potentes y el único modo en que el programa ELIZA podía decidir que una persona había acabado una frase era si al final llevaba un punto. Hasta la última línea, el VP había acabado cada línea con un punto, pero en la última se le había olvidado. En consecuencia, el ordenador continuó esperando a que se escribieran más cosas y, por tanto, no respondió. Según Bobrow me contó, «el vicepresidente, furioso, al pensar que estaba jugando con él, me llamó, me sacó de un sueño profundo y me dijo: “¿Por qué se da esos aires de su-

16. Conversación entre Danny Bobrow, ELIZA y el VP. Güven Güzeldere y Stefano Franchi realizaron una transcripción de esta conversación que se puede consultar, y la he copiado de su página web (Güzeldere y Franchi, 1995). Asimismo, confirmé los detalles a través de una conversación y por correo electrónico con Bobrow (27 de diciembre de 2002).

perioridad conmigo?», a lo cual le contesté: “¿Qué quiere decir con que me estoy dando aires de superioridad con usted?”». Entonces, Bobrow me contó que «el vicepresidente me leyó todo el diálogo que “nosotros” habíamos mantenido sin encontrar otra reacción por mi parte que la risa. Tardé un rato en convencerle de que en realidad había estado dialogando con un ordenador».

Tal como Bobrow me contó cuando los dos hablamos de aquella interacción: «Como puedes ver, le importaban mucho las respuestas a sus preguntas y lo que creía que eran mis observaciones tuvieron un efecto emocional en él». Somos en extremo confiados, lo que nos hace fáciles de engañar, y nos sentimos muy enojados cuando creemos que no se nos toma en serio.

La razón de que ELIZA tuviera tal efecto guarda relación con lo que examinamos en el capítulo 5 sobre la tendencia que los seres humanos tenemos a creer que cualquier interacción con visos de ser inteligente tiene que darse entre humanos o, cuando menos, con una presencia inteligente: lo que conocemos con el nombre de antropomorfismo. Además, dado que somos confiados, somos propensos a tomarnos estas interacciones en serio. ELIZA fue escrito hace mucho tiempo, pero su creador, Joseph Weizenbaum, se sintió horrorizado por la seriedad con que toda aquella gente se tomaba, al interactuar con el programa, aquel sencillo sistema que él había programado. Sus preocupaciones le llevaron a escribir *Computer Power and Human Reason*,<sup>17</sup> libro en el cual defendía con contundencia que todas aquellas frívolas interacciones eran perjudiciales para la sociedad humana.

Hemos recorrido un largo camino desde que ELIZA fue escrito. Los ordenadores de nuestros días son miles de veces más potentes de lo que eran en la década de 1960 y, lo que es aún más importante, el conocimiento que hoy tenemos del comportamiento y la psicología humana ha aumentado y profundizado de forma espectacular. Una consecuencia de todo ello es que hoy podemos escribir programas y construir máquinas que, a diferencia de ELIZA, tienen una comprensión certera y pueden mostrar emociones verdaderas. Sin embargo, esto no significa que nos hayamos librado de aquello que preocupaba a Weizenbaum. Veamos el caso de Kismet.

Kismet, cuya fotografía aparece reproducida en la figura 6.6, fue desarrollado por un equipo de investigadores del Artificial Intelligence Labo-

17. Weizenbaum, 1976.

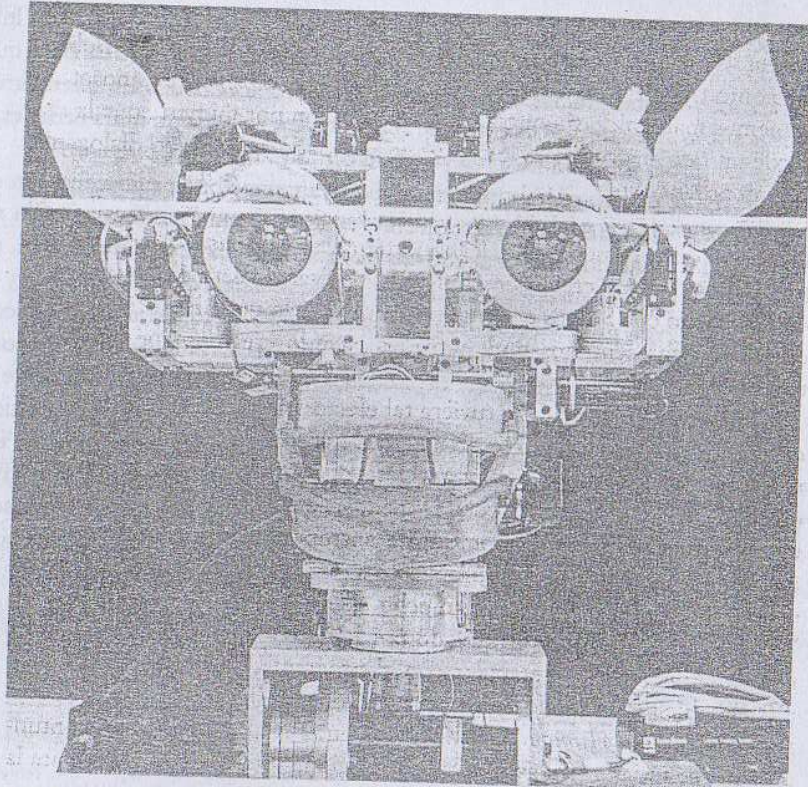


FIGURA 6.6. Kismet —un robot diseñado para las interacciones sociales— mostrando sorpresa.<sup>18</sup> (Imagen reproducida por cortesía de Cynthia Breazeal.)

laboratory en el MIT, y Cynthia Breazeal le dedicó una completa exposición en su obra *Designing Sociable Robots*.<sup>19</sup>

Conviene no olvidar que las emociones que subyacen al habla pueden detectarse sin que medie comprensión alguna del lenguaje. Las voces que expresan enojo, amonestación, súplica, consuelo, gratitud y elogio tienen un tono y contornos de volumen característicos. Podemos decir en cuál

18. La fotografía de Kismet procede de <<http://www.ai.mit.edu/projects/sociable/ongoing-research.html>> (con autorización). Para una descripción más detallada, véase el libro de Cynthia Breazeal, *Designing Sociable Robots* (Breazeal, 2002).

19. Breazeal, 2002.

de esos estados se halla alguien aun cuando nos hable en una lengua extraña. Nuestros animales domésticos a menudo detectan nuestro estado de ánimo a través tanto de nuestro lenguaje corporal como de las pautas emocionales que se expresan en nuestra voz.

Kismet utiliza estas indicaciones para detectar cuál es el estado emocional de la persona con la que interactúa en un momento dado. Kismet tiene cámaras de video a modo de ojos y microfones con los que escucha. Kismet tiene una estructura sofisticada que le permite interpretar, evaluar y responder al mundo —según un esquema que mostramos con mayor detalle en la figura 6.7— y en la cual se combina percepción, emoción y atención a fin de controlar el comportamiento. Si uno camina hacia donde está el robot, Kismet gira su rostro hacia el lugar donde uno se halla y le mira directamente a los ojos. Pero cuando nos quedamos simplemente allí de pie y no hacemos nada más, Kismet se aburre y comienza a mirar a su alrededor. Si le dirigimos la palabra, el robot se muestra sensible al

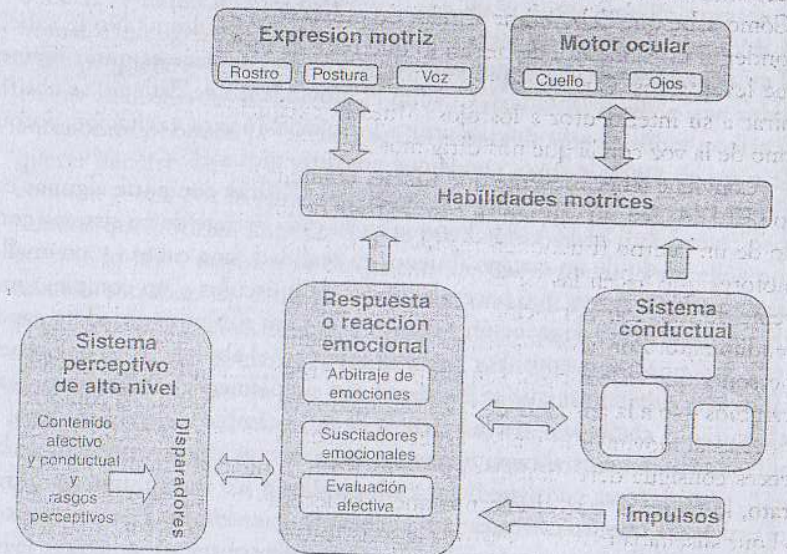


FIGURA 6.7. El sistema emocional de Kismet. El corazón operativo de Kismet es la interacción entre la percepción, la emoción y el comportamiento. (Figura dibujada de nuevo y ligeramente retocada con permiso de Cynthia Breazeal, que la publicó en <<http://www.ai.mit.edu/projects/sociable/emotions.html>>.)

tono emocional de la voz de la persona que le habla y reacciona con interés y placer a todo elogio alentador y gratificante, mientras que muestra vergüenza y pesar cuando se le regaña. El espacio emocional de Kismet es bastante rico, y puede mover la cabeza, el cuello, los ojos, las orejas y la boca para expresar emociones. Si está triste, las orejas se le agachan. Si se excita, las orejas se le ponen tiesas de punta. Cuando se siente desdichado, la cabeza y las orejas quedan gachas y la boca se le alarga hacia abajo.

Interactuar con Kismet es una experiencia atractiva y rica. Resulta difícil creer que Kismet sea toda emoción, sin picos de entendimiento. Pero si uno se dirige hacia donde está, le habla con una voz llena de excitación y le muestra el flamante reloj de pulsera, Kismet reacciona de manera adecuada: el robot nos mira a la cara, luego al reloj y luego de nuevo a la cara, mostrando todo el rato interés, levantando las cejas y moviendo las orejas, y mostrando un comportamiento lleno de vitalidad y alegría. Se trata precisamente de las reacciones de interés que deseamos encontrar en el compañero de conversación, aun cuando Kismet no comprenda nada de lo que le decimos ni, en realidad, del reloj de pulsera que le mostramos. ¿Cómo sabe que debe mirar al reloj de pulsera? No lo sabe, pero al responder al movimiento, de hecho lo que hace es mirar fijamente la mano que levantamos. Cuando el movimiento se detiene, se aburre y vuelve a mirar a su interlocutor a los ojos. Muestra excitación porque detecta el tono de la voz con la que nos dirigimos a él.

Conviene tener presente que Kismet comparte algunas características con ELIZA. Así, aunque en su caso se trate de un sistema complejo, dotado de un cuerpo (bueno, en realidad, una cabeza y un cuello), múltiples motores que hacen las veces de músculos y un complejo modelo subyacente de atención y emoción, le falta aún una verdadera capacidad de entendimiento. Por tanto, el interés y el aburrimiento que muestra hacia las personas que se le dirigen son simplemente respuestas programadas a los cambios —o a la ausencia de tales cambios— en el entorno, así como respuestas al movimiento y a los aspectos físicos del habla. Si bien Kismet a veces consigue dejar embelesados a los seres humanos durante un largo rato, lo cierto es que el encanto que despierta es, en cierto sentido, afín al que suscitaba ELIZA: casi toda la sofisticación la pone la interpretación que hace el observador.

Aibo, el perro robot de Sony, tiene un repertorio emocional y una inteligencia mucho menos sofisticados que Kismet. Sin embargo, Aibo también ha resultado ser increíblemente estimulante y atractivo para aquellas personas que tienen uno. Muchas de estas personas se reúnen y han for-

mado clubes: algunos tienen varios de estos robots. Allí intercambian anécdotas sobre el modo en que han entrenado a Aibo para que haga distintas gracias. Comparten ideas y técnicas. Algunos creen firmemente en que su Aibo es capaz de reconocerles y obedecer a sus órdenes, pese a que en realidad no es capaz de hacer tales hazañas.

Cuando las máquinas muestran emociones, proporcionan una interacción rica y satisfactoria con los seres humanos, aunque en su mayor parte esa riqueza y satisfacción, esa interpretación y comprensión, se hallan en la cabeza del ser humano y no en el sistema artificial. Sherry Turkle, a la vez profesora del MIT y psicoanalista, resumió estas interacciones señalando que «dice más acerca de nosotros como seres humanos que de los robots». De nuevo el antropomorfismo: leemos emociones e intenciones en toda suerte de cosas y objetos. «Estas cosas, con independencia de que tengan o no inteligencia, accionan nuestras teclas», añadía Turkle, «pulsan las teclas que nos hacen reconocerlas como si lo hicieran. Estamos programados para responder y reaccionar de modo afectuoso a este nuevo tipo de criaturas. El secreto estriba en que estos objetos quieren que se les cuide y están bien cuando les prestamos atención.»<sup>20</sup>

20. Citas de Turkle extraídas de una entrevista mantenida con L. Kahney, en *Wired.com* (aunque he corregido la gramática) (Kahney, 2001).