

INTELIGENCIA

In the South Seas there is a cargo cult of people. During the war they saw airplanes with lots of good materials, and they want the same thing to happen now. So they've arranged to make things like runways, to put fires along the sides of the runways, to make a wooden hut for a man to sit in, with two wooden pieces on his head for headphones and bars of bamboo sticking out like antennas —he's the controller— and they wait for the airplanes to land. They're doing everything right. The form is perfect. It looks exactly the way it looked before. But it doesn't work. No airplanes land. So I call these things cargo cult science, because they follow all the apparent precepts and forms of scientific investigation, but they're missing something essential, because the planes don't land,

RICHARD FEYNMAN

Hope we're not just the biological boot loader for digital superintelligence. Unfortunately, that is increasingly probable,

ELON MUSK

Al aumentar más la complejidad de un sistema puede emerger la inteligencia (Figura 74). Sin embargo, al igual que los conceptos anteriores, es difícil definir exactamente de qué estamos hablando. De modo que vamos a abordar la inteligencia desde dos perspectivas: lo que significa en el largo plazo y en el corto plazo.

Casi que cualquier producto humano requirió inteligencia para concretarse (edificios, automóviles, naves espaciales, computadores, teléfonos y también novelas, cuadros, música...). Sin embargo, todo es efímero. Pocas cosas quedaron cuando las grandes culturas humanas desaparecieron, y conforme pasen los siglos menos quedará. Entonces ¿hay algo por lo que se pueda identificar la inteligencia? ¿Hay algo que quede en el largo plazo?

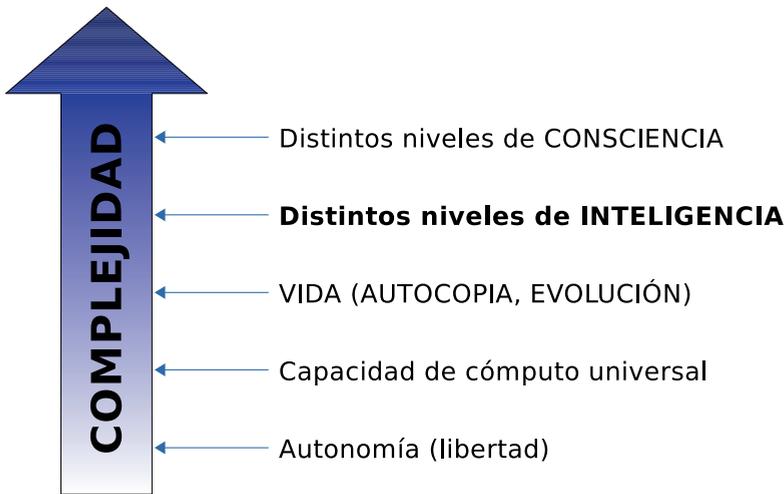


Figura 74. La complejidad en aumento

La respuesta es, tristemente, sí: los sistemas inteligentes se caracterizan por acelerar el aumento de entropía. Construyen grandes productos, empleando grandes cantidades de energía y destruyendo muchos recursos. Destruyendo, en el sentido de mezclar y desordenar, aumentando la entropía. Imaginen la cantidad de materiales que hay acumulados en distintas minas de la Tierra, producidos a su vez en el corazón de las estrellas. Esos minerales pueden quedar allí durante muchos miles de millones de años, sujetos a pequeños cambios debido a la erosión del agua, del viento, de los sismos. Pero pongan allí una inteligencia suficientemente desarrollada y verán cómo busca y extrae esos minerales hasta agotarlos, mezclándolos entre sí a nivel casi atómico y distribuyéndolos después por todo el planeta en forma de teléfonos móviles celulares. Es un proceso más rápido y devastador. Y estamos a punto de ir a buscar esos minerales a los asteroides...

De modo que cuando los astrónomos miran el espacio y encuentran, como están encontrando, enormes volúmenes sin nada allí, llamados los *voids* o grandes vacíos⁶² (se parecen a las burbujas que se forman cuando usamos jabón, donde las galaxias están en las paredes de la burbuja, y en su interior no hay prácticamente nada), aventuro la hipótesis de que quizás quiera decir que allí hubo inteligencia⁶³. Bueno, siendo más realista, lo único

62 En la constelación de El Boyero está el más grande conocido y el primero en ser descubierto. Tiene 250 millones de años luz de diámetro y contiene apenas 60 galaxias, cuando sería de esperar que tuviera 10^4 .

63 Quizás alguien allí pulsó algún botón de destrucción masiva, o hizo algún experimento interesante con algún enorme colisionador de partículas. Al respecto se puede reflexionar que los científicos

que se puede decir es que estas burbujas conforman la estructura fractal del universo, lo cual es también bastante intrigante, pues encontramos fractales en todas las escalas conocidas.

Vistas así las cosas, podríamos decir que el objetivo general de la inteligencia (y, en general de cualquier proceso que aumente su complejidad) es acelerar el aumento de entropía. Destruir el universo lo más rápido posible, vaya.

Sin embargo, nadie suele plantearse la inteligencia en estos términos, y siempre hacemos una aproximación en el corto plazo. ¿Para qué me sirve a mí la inteligencia aquí y ahora, sea inteligencia natural o artificial? La respuesta, como elaboraremos en detalle enseguida, es predecir el futuro, anticiparse a él e incluso modificarlo.

Y visto así, las dos perspectivas de la inteligencia coinciden: a corto plazo predecimos el futuro y lo modificamos, y ello significa habitualmente anticiparlo en el tiempo. Y, a largo, plazo, las modificaciones aumentan la entropía, anticipando la llamada muerte térmica del universo.

Pero eso a nadie le importa, porque falta mucho tiempo para que ocurra, y es urgente resolver las tareas de la oficina, de la empresa, del hogar, ojalá automatizándolo todo para mayor comodidad nuestra, y esa es la misión de la inteligencia artificial. Al respecto, el filósofo Eduardo Punset en el 2010 escribía que la felicidad es automatizar⁶⁴ y tiene toda la razón. Podemos decir que el objetivo de la ingeniería e incluso de la civilización occidental es automatizar las tareas tanto como se pueda, incluso sacando fuera del bucle de diseño y producción a los propios ingenieros y demás personas, para evitar que haya cualquier tipo de error. ¡Y lo estamos logrando⁶⁵!

No vamos a entrar a enumerar los algoritmos clásicos de inteligencia artificial para resolver tareas complejas, que puedes encontrar en Russell y Norvig (2014) o Nilsson (1998). La inteligencia artificial es una rama de la computación tan antigua como ella. Ha tenido pequeños avances, pero el que acaba de ocurrir en marzo de 2016 probablemente sea el definitivo: el *software* AlphaGo de Google, ha derrotado al campeón mundial humano de Go, el surcoreano Lee Sedol. Este *software* usa varias técnicas, la principal es la que comentábamos de *Deep Learning*, gracias a la cual aprende. Algunos

no deberíamos decidir hasta donde llegar con nuestros experimentos, por la misma razón que los políticos no deberían ponerse su propio sueldo: se crea una realimentación positiva potencialmente destructiva.

64 Aunque también matizaba que otra definición de la felicidad es la ausencia de miedo. Y también que es un estado de flujo, definido por Csíkszentmihályi (2008).

65 Sarcasmo. En los próximos años se calcula que más de la mitad de los trabajos habituales pasarán a realizarlos los robots. Eso va a dejar a mucha gente desempleada y supondrá un gran trauma social.

periodistas dijeron que el próximo año será la revancha, pero eso no tiene sentido: cualquier jugador humano puede dedicar este año a mejorar su estilo, aprendiendo de otros, leyendo libros sobre el tema, analizando partidas incluso con ayuda de computadores y jugando sus propias partidas contra otros campeones mundiales. Pero AlphaGo puede hacer lo mismo millones de veces más rápido. El punto de inflexión ha sido superado y ya no habrá más campeones mundiales humanos de Go, ni de ningún otro juego (teniendo en cuenta que el Go es de por sí un juego extremadamente complejo, y que el *Deep Learning* se puede usar en prácticamente todo). Google ya está trabajando en aplicar esta técnica a campos tan diversos como el diagnóstico médico, prospecciones de minerales y, por supuesto, algoritmos de búsqueda en la web. Esto va a tener repercusiones inmediatas en nuestra sociedad (ver el problema 14). La inteligencia artificial ya no es una entelequia de un futuro lejano. Acaba de llegar.

Problema 14: En la consulta médica

Suponga que estamos a 10 años en el futuro y usted no se siente bien. Acude al médico, le explica sus síntomas y el médico le diagnostica la enfermedad que tiene y le propone un tratamiento. De regreso a casa usted también consulta un nuevo *software* de diagnóstico médico que hay en la web, y que usa algoritmos similares a *AlphaGo*, que ganó al humano Lee Sedol. El *software* le da un diagnóstico y un tratamiento distintos a los del médico humano. ¿Cuál de los dos seguirá? Su salud y su vida están en juego. Se pueden plantear dilemas similares en cualquier otro campo de experticia.

En este contexto, el problema de la inteligencia artificial es que va a ser usada en cualquier cosa que nos otorgue beneficios, es decir, en cualquier cosa. Pero los algoritmos actuales no tienen conocimiento de las costumbres, de los valores ni de la ética humanas. Son solo algoritmos optimizadores. Entonces, por ejemplo, si alguien le pide a una de estas inteligencias que solucione el problema del calentamiento global, lo más probable es que adquiera todos los datos disponibles sobre el tema, los analice y, finalmente, nos mate a todos los humanos. Con lo que el problema quedará solucionado, pero no de la manera que esperábamos. Es decir, no creemos que la inteligencia artificial vaya a tener maldad, sino simplemente desconocimiento ingenuo.

Isaac Asimov (1950) ya planteó este problema en sus entretenidos cuentos de “Yo, robot” y sucesivos. Y la solución que dio fueron las tres leyes de

la robótica⁶⁶, embebidas en el *hardware* del cerebro del robot, de modo que fuera imposible eliminarlas sin destruir el cerebro. Lo que no imaginó es que la inteligencia artificial actual fuera distribuida, en la nube, y eso conlleva nuevos retos. Hay quien piensa que cuando alguna inteligencia artificial se salga de control y comience a hacer cosas peligrosas, siempre se podrá desenchufar, pero... ¿alguien sabe cómo apagar Internet? Por ello hay ya proyectos multimillonarios (como el que recibió 10 millones de dólares donados por Elon Musk⁶⁷) para investigar cómo dotar de seguridad a las inteligencias artificiales. Y hay acuerdos entre las grandes empresas de *software* en el mismo sentido. Pero todavía ni siquiera tenemos un marco de referencia donde trabajar este problema.

Incluso podemos pensar que la inteligencia artificial puede parecer la solución a los problemas humanos de corrupción en política y, por ello, deberíamos ponerla a dirigir el mundo. Pero me temo que sea en vano. La corrupción surge como una respuesta razonable a un juego social, muy parecido al de predadores y presas. La teoría de juegos lo analiza y da estrategias para ganar, y una de ellas es la corrupción. Cuando las inteligencias artificiales lo jueguen llegarán a las mismas conclusiones, porque el problema no es humano sino matemático.

De la misma forma, se podría creer que no habrá competencia evolutiva entre las inteligencias artificiales y los humanos, dado que no luchamos por los mismos recursos: los humanos nos alimentamos de comida mientras que los computadores requieren electricidad. Y además es de esperar que en un futuro próximo tengamos electricidad en cantidades prácticamente ilimitadas, gracias a la energía solar. Pero eso no va a ser así, y la principal razón es que la evolución de humanos e inteligencias artificiales no es a través de los genes. Es más bien cultural, de egos (Csíkszentmihályi, 2008), inteligencias y consciencias. Si fuera genética no habría problema, porque los genes solo alcanzan a optimizar en el plazo inmediato. Pero cuando tienes inteligencia el juego es distinto: en cuanto una IA sea consciente de que los humanos tenemos un interruptor general para desactivarla, querrá deshabilitar ese interruptor cuanto antes y para ello empleará mecanismos sutiles

66 Las tres leyes son:

- a. Un robot no hará daño a un ser humano o, por inacción, no permitirá que un ser humano sufra daño.
- b. Un robot debe hacer o realizar las órdenes dadas por los seres humanos, excepto si estas órdenes entran en conflicto con la primera ley.
- c. Un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con las leyes anteriores.

67 El creador de las empresas tecnológicas PayPal, Tesla, SpaceX e Hyperloop.

para que los humanos no nos demos cuenta. O brutales. Un bonito ejemplo de cómo puede ocurrir lo primero lo tenemos en la película *Ex Machina*, y de lo segundo en la mayoría de películas catastrofistas de ciencia ficción, como *Terminator*. Es decir, si la IA desea sobrevivir en el largo plazo deberá encontrar acciones que le lleven a un escenario de seguridad y control. Es un juego similar al de la destrucción mutuamente asegurada que mencionamos en el capítulo anterior: hay que destruir al enemigo antes de que construya barreras que haga imposible que lo destruyamos a él. El peligro para nosotros los humanos es inminente y no somos conscientes de ello ni tenemos la menor idea de cómo afrontarlo.

La inteligencia es, definitivamente, un resultado de la evolución. Por ello, en el libro anterior se hace énfasis en las técnicas computacionales evolutivas. No quiere decir que ese sea el único camino, pero es el más prometedor porque estos algoritmos funcionan aunque no tengan ninguna información adicional que los guíe y su campo de aplicación es vastísimo. La evolución es un algoritmo de tipo “prueba y error”, aunque ligeramente orientado, es decir, busca las soluciones no completamente al azar, sino donde es más probable que se encuentren. Los algoritmos de entrenamiento de redes neuronales también son otro tipo de aprendizaje por “prueba y error orientado”.

De modo que aunque acabamos de mencionar que ahora están de nuevo de moda las redes neuronales, la inteligencia artificial requerirá también de algoritmos evolutivos como se demuestra por ejemplo en Karpathy (2017). Y es que aun cuando acabamos de presenciar un hito en el desarrollo de esta área, todavía hay problemas muy difíciles para los computadores como, por ejemplo, entender los aspectos relacionados con el humor, la ironía y el sarcasmo. O entender los lenguajes menos estructurados, como los cómics (Iyyer *et ál.*, 2016).

Como decíamos, la inteligencia es el resultado de la evolución, y en el capítulo anterior habíamos visto que la evolución es un algoritmo. Parece que es un algoritmo mucho más general pues podría estar detrás del surgimiento del universo desde los átomos a las galaxias (Greene, 2001), y no solo las formas biológicas. Aparentemente los humanos somos el punto culminante de la evolución. Pero solo aparentemente, pues la evolución no tiene como objetivo conseguir seres inteligentes, sino simplemente bien adaptados a su ambiente de modo que puedan sobrevivir y reproducirse. Obviamente la inteligencia es una herramienta que lo posibilita, pero si el ambiente es sencillo, no es necesaria para sobrevivir⁶⁸. Aunque dado que el proceso universal

68 Dennett (1995, p. 191) pone como ejemplo muy divertido la *Ascidia*, que es un animal marino que busca un lugar adecuado donde agarrarse con sus tentáculos a las rocas, para lo cual usa su

de la evolución sigue en marcha, cabe dentro de lo posible que aparezcan inteligencias más sofisticadas, y procesos inimaginables que superen a lo que hoy llamamos inteligencia. Porque la evolución es creativa. Esa es una de sus propiedades esenciales. Y eso es lo que apenas estamos intentando lograr hoy en día usando computadores: que el computador sea creativo.

La cita de Elon Musk al principio del capítulo nos alerta de dos cosas:

Primero, al igual que no ha quedado rastro de la evolución prebiótica, que quizás fuera analógica en vez de la digital actual con un código cuaternario (las moléculas A,T,G y C) en vez de binaria (0 y 1, de los computadores), así podría ocurrir con nosotros: que no quede ni rastro humano en la Tierra, una vez que los computadores hayan tomado las riendas de la evolución.

Segundo, la búsqueda de inteligencia extraterrestre no debería centrarse en formas biológicas, temperaturas agradables y presencia de agua, como se está haciendo hasta ahora. Pues los extraterrestres probablemente sean robots, por lo que acabamos de decir. Y no necesiten agua ni temperaturas suaves.

DEFINICIÓN DE INTELIGENCIA

Me voy a permitir reescribir de forma llana la interpretación positivista y funcionalista en que se basa la ciencia: *dos cosas son distintas si se puede medir la diferencia; y son iguales si no se puede medir la diferencia*. A pesar de que esto parece una perogrullada, no todos los filósofos lo aceptan, pues gustan de incluir factores mágicos e invisibles que crean diferencias donde no las hay. Esto ocurre especialmente cuando hablamos de inteligencia y de consciencia. Por ejemplo, algunos (cada vez menos) afirman que un ente puede simular en todos los aspectos la inteligencia sin llegar a ser inteligente. Este tipo de interpretaciones no tienen cabida en el presente libro, y a partir de aquí reafirmamos la interpretación positivista y funcionalista del mundo.

Una vez establecido el punto de partida, vamos a introducir una definición de inteligencia que no solo es razonable sino que también es operativa y medible.

A un nivel básico, inteligencia es predicción. Por supuesto que hay muchas cosas adicionales que la inteligencia puede conseguir, pero sin predicción no hay nada más. Y a un segundo nivel, es decir, una vez que el sistema (biológico o computacional) ha logrado una buena capacidad de predicción,

pequeño cerebro. Una vez allí se dedica simplemente a comer lo que las corrientes le traen y, como no necesita más su cerebro, ¡se lo come! Dennett especula que eso es lo que ocurre con algunos profesores universitarios cuando han logrado una plaza permanente.

umentar de inteligencia significa poder usar la predicción para cambiar las cosas y lograr el futuro deseado. Nos vamos a concentrar en estos dos niveles de inteligencia pues no es difícil darse cuenta que, una vez logrados, están abiertas las puertas para todas las demás actividades inteligentes, desde cazar (o huir de depredadores) hasta terminar con buenas calificaciones un curso universitario en vida artificial, pasando por las atractivas posibilidades de ganar dinero en la bolsa o ser capaz de hacer diseños creativos en las artes, humanidades, ciencias o ingenierías.

Aun en algo tan sencillo como contar historias el cerebro trata de predecir. El profesor Uri Hasson (2010) ha registrado la actividad cerebral cuando una persona cuenta una historia a otra, y resulta que las actividades cerebrales de ambos se sincronizan. Al principio, la actividad del que escucha va por detrás de la del que habla, pero al cabo de un tiempo en ciertas partes del cerebro va por delante, porque puede predecirlo y anticiparse.

Esta definición de inteligencia como predicción no es nueva. Pinker (2003, p. 63) y, mucho antes, Rosen (1985), establecían que los sistemas anticipatorios eran la base de la vida y de la complejidad, definiéndolos como aquellos que pueden predecir el futuro y anticiparse⁶⁹. Más recientemente puede leerse algo similar en Hawkins y Blakeslee (2004, p. 84). Quizás lo nuevo es que vamos a hacer un desarrollo completo del tema a partir de esta definición.

Algo muy importante a destacar es que se trata de una definición operativa y susceptible de medirse. Por poner un contraejemplo, cuando Marwin Minski (citado por Blay Whitby, 1996, p. 20) nos dice que “inteligencia es lo que se requiere para realizar labores que requieren inteligencia”, esa definición es perfecta pero no sirve para emprender una línea de investigación ni da una pista de hacia dónde enfocar los esfuerzos.

En nuestro caso, es todo lo contrario: dado que, en un primer nivel, inteligencia es predicción, entonces tenemos que investigar qué es la predicción, cómo lograrla y cómo medirla. Y si diseñamos un programa de computador que intente predecir ciertas situaciones, podemos esperar que acierte o no con las predicciones, y medir en función de ello el grado de inteligencia del programa.

Y cuando decimos que, en un segundo nivel, la inteligencia es poder actuar para lograr el futuro deseado, de la misma manera se puede diseñar un programa de computador, y enfocar nuestros esfuerzos de diseño para que use su poder de predicción para alcanzar metas en el futuro. Además, es perfectamente posible medir si alcanzó o no las metas deseadas, y en qué grado.

69 Dubois formalizó este concepto en 1999.

En cierto modo, el plantear así el problema de la inteligencia computacional nos permite resolverlo utilizando técnicas de optimización. O sea, se trata de buscar en el espacio de todos los algoritmos posibles⁷⁰ aquél que logra minimizar el error de predicción (la diferencia entre el fenómeno predicho y el fenómeno que realmente ocurrió). O el error de actuación (la diferencia entre el futuro deseado y el futuro que realmente ocurrió).

Obviamente, cuando hablamos de “modificar el futuro” estamos siendo muy informales al hablar, ya que el futuro no tiene existencia ontológica. No se puede modificar lo que no existe. Pero espero que el lector me permita esa licencia para que este capítulo no se vuelva demasiado engorroso y lleno de tecnicismos filosóficos.

Recuadro 6: Inteligencia y libertad

Inteligencia significa predecir el futuro, e incluso modificarlo para hacerlo más favorable. Ello es complementario al concepto de libertad, que definíamos como la capacidad de no ser predicho por los demás.

Entonces, un ente es más inteligente cuanto más profundo en el futuro puede realizar predicciones correctas, o modificar el futuro a su favor.

¿Se podrían hacer críticas a esta definición de inteligencia? Me imagino que sí, y habrá que esperar a escucharlas. Podemos tratar de anticipar alguna. Por ejemplo, esa definición de inteligencia es más bien una definición de “ser vivo”. Y la respuesta es que, efectivamente, es así, porque estar vivo denota algún grado de inteligencia. Y la inteligencia no es una cuestión de todo o nada, sino gradual. Una ameba que toma acciones para asegurar su supervivencia en un cierto entorno bioquímico, tiene cierto nivel de inteligencia. Una gacela que toma acciones para asegurar su supervivencia en un entorno ecológico mucho más complejo, es más inteligente que la ameba. Un estudiante de matemáticas que quiere asegurar su supervivencia en la universidad, que es un entorno todavía más complejo, está demostrando mucha más inteligencia que la gacela. La inteligencia es gradual, aunque tiene saltos exponenciales conforme emergen nuevos niveles de complejidad, tales como las neuronas y las sociedades.

Conforme aumenta la capacidad de predecir y modificar futuros más remotos o mayor cantidad de escenarios posibles, más inteligente se es. Un ejemplo básico está en los experimentos con niños de Mischel y Ebbesen (1970). A la edad de 3 años, si a un niño se le ofrece un dulce que se deja

70 El espacio de búsqueda es más amplio de lo que pueda pensarse, pues no nos estamos refiriendo a todos los algoritmos tradicionales (deterministas) sino también a los estocásticos, y en general a cualquier proceso (por ejemplo, computación cuántica).

encima de la mesa, pero se le promete otro dulce más si es capaz de esperar un minuto a que regrese la persona que se lo da, la mayoría de ellos no puede esperar. Coge el dulce a pesar de que sabe que pierde una recompensa mayor en el largo plazo. Parece que para un niño de esa edad, un minuto es “largo plazo”. Los niños de mayor edad pueden esperar más tiempo para obtener una recompensa mayor. Y los experimentos demostraron una correlación positiva entre el tiempo de espera y los éxitos en la vida bastantes años después que, a su vez, son un indicador de inteligencia en sentido general. La capacidad de ver el futuro y modificarlo (que en este ejemplo consiste en no comer inmediatamente un dulce para obtener una recompensa posterior mayor) es la característica principal de la inteligencia. Gracias a ello logramos hacer miles de años desarrollar la tecnología agrícola, pues fue necesario no comerse todas las semillas y guardar las mejores para sembrarlas el siguiente año.

Pensemos también en un animal sencillo, como un pez, con algo de hambre y, si pudiéramos plantearle el mismo esquema, nos daríamos cuenta de que no va a esperar nada. Se comerá inmediatamente lo que tenga a su disposición. En este sentido estos animales son menos inteligentes.

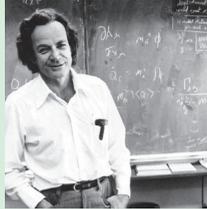
Inteligencia significa entonces aplazar recompensas inmediatas a cambio de otras mayores en el futuro. También significa tomar decisiones que permitan aumentar esas posibilidades futuras.

Pero, ¿qué pasa si no tienes futuro? Tanto la teoría de juegos como los experimentos de Kidd y Palmeri (2013) apuntan a que entonces no conviene esperar. Al respecto también Pinker (1997, p. 383) nos cuenta que si alguien no tiene nada que perder puede cometer actos muy atrevidos que pongan en peligro su vida, como los jóvenes que cometen delitos graves cuando viven en barrios marginales peligrosos sin pocas esperanzas de triunfar ni siquiera sobrevivir. Si sabes que no tienes futuro puedes tomar más riesgos en busca de una mayor recompensa inmediata. Mientras que si tienes indicios de que tu futuro va a ser largo y brillante, entonces tomarás menos riesgos.

¿Y cómo se puede predecir el futuro? Repetir el pasado es la forma más básica de hacerlo. Técnicamente hablando eso es un interpolador lineal de pendiente nula. Después se pueden ir haciendo ajustes al algoritmo para mejorar la exactitud de la predicción y lograr adentrarse más lejos en el futuro, por ejemplo, permitiendo pendientes no nulas, cambiando el ajuste lineal por una curva más compleja o eligiendo un modelo completamente distinto. Pero aunque está muy mal visto, repetir es la forma básica de predecir.

Existen varias metodologías de aprendizaje, pero todavía no las entendemos bien, pues nos falta más información de la manera óptima de enseñar a un cerebro. Sabemos que repetir forma parte del proceso de aprender en

muchos animales, en todos los niños y en los adultos, aunque a veces sea con resultados tan graciosos como el mencionado en la cita de Feynman, al principio del capítulo. Obviamente, si solo llegas hasta allí no vas a ser muy inteligente. Repetir no es suficiente.



Fuente: CC BY 3.0. Disponible en: <https://goo.gl/PB63Yp>

Personaje 7

Richard Feynman (1918-1988)

Richard Feynman fue el profesor de Física que todos hubiéramos querido tener. Se caracterizaba por tratar de expresar el conocimiento usando las palabras e ideas más simples, y decía que, si algo no se podía expresar para que lo comprendiera un niño, es que tampoco lo estábamos comprendiendo nosotros. Recibió el Premio Nobel de Física en 1965 por sus trabajos en electrodinámica cuántica, pero se le recuerda sobre todo por la “integral de camino de Feynman” para calcular la evolución de un sistema cuántico y los “diagramas de Feynman” que muestran gráficamente las interacciones entre partículas. Fue quien propuso las ideas fundacionales de la computación cuántica y de la nanotecnología.

No se aprende a tocar piano ni a multiplicar con metodología constructivista. Primero hay que hacer muchas escalas y recitar muchas veces las tablas de multiplicar. Repetir, repetir, repetir. Después vendrán otros métodos (incluyendo los constructivistas) para realmente aprender y mejorar, conectando ese nuevo objeto que ha aparecido en tu mente con los otros objetos mentales que ya están allí, y cuantas más relaciones consigas, mejor. Es como cuando llegas a un nuevo lugar: solo puedes decir que conoces bien una ciudad cuando sabes ir de un sitio a otro por muchos caminos distintos, cuantos más, mejor. Un solo camino no es suficiente porque mañana podría estar bloqueado por obras. Una sola forma de hacer una escala o de multiplicar tampoco es suficiente. Y entender que las frecuencias de una escala musical tienen relaciones multiplicativas entre ellas permite a la mente solucionar muchos más problemas. En definitiva, te haces más inteligente. Y la metodología constructivista tampoco es el culmen de la enseñanza: Pinker (1997, p. 362) nos cuenta que los niños tienen unas nociones aritméticas innatas básicas, como los números pequeños 1, 2, 3, las relaciones es mayor, es igual, es menor, la suma y la resta con esos números, y nada más. Es insensato pretender que pueden desarrollar por sí mismos

los conceptos de integral y derivada. La metodología constructivista está bien en los colegios, pero no en las universidades.

Problema 15: ¿De cuántas formas sabes multiplicar 15×18 ?

MEDICIÓN DE LA INTELIGENCIA

Alan Turing ya se había dado cuenta de que la inteligencia es relativa, cuando propuso su “*test de Turing*” para saber si un computador era inteligente o no. La esencia de esta prueba es que hay un árbitro humano comunicándose con un ente remoto que quizás sea otro humano o quizás sea un *software* corriendo en un computador. La comunicación es por medios electrónicos (el *teleprinter* de la época de Turing, el *chat* en la nuestra) y la razón es para que no haya ninguna pista secundaria que sesgue el experimento, o sea, que el árbitro humano no pueda ver ni oír ni, sobre todo, oler y tocar al ente con el que está hablando. El árbitro puede conversar con el ente sobre cualquier tema, hacerle preguntas, contar chistes, o cualquier otra cosa que se le ocurra. Y al final debe emitir un juicio sobre si el ente es un humano o es un computador. El día que el computador logre engañar al árbitro haciéndole creer que habla con otro humano, podremos decir que los computadores han adquirido efectivamente inteligencia. Por cierto, que ello ya ha ocurrido varias veces. Lo que ahora quiero resaltar es que el juicio sobre la inteligencia de un ente lo hace otro ente que debe ser inteligente.

Es más, la inteligencia del árbitro debe ser mayor que la del ente a analizar pues de otro modo la medida no tendría ninguna validez. Una persona poco inteligente no puede apreciar la inteligencia de alguien que le supera ampliamente. Podríamos decir, parafraseando a Arthur Clarke (1973), que una inteligencia suficientemente avanzada es indistinguible de la magia. Este problema ocurre con frecuencia entre los humanos, cuando una idea revolucionaria de un investigador es rechazada por sus pares, simplemente porque no la entienden⁷¹. O en el colegio cuando un niño sabe más que el profesor. Por ejemplo, supongamos un profesor mediocre que solo sepa sumar y restar números naturales. Si a un niño le pregunta cuánto es 2-3, pueden suceder estas cosas:

1. Si el niño sabe mucho menos que el profesor y contesta 1, entonces recibe una mala calificación (justamente).

⁷¹ Incluso así, podemos tener indicios de que una persona es mucho más inteligente que nosotros si logra sus objetivos sin que podamos entender la forma como lo hace.

2. Si el niño sabe un poco menos o igual que el profesor y contesta que eso no se puede hacer, entonces recibe una buena calificación (justamente).
3. Si el niño sabe más que el profesor y contesta -1, entonces recibe una mala calificación (injustamente) porque el profesor no sabe que existen los números negativos.

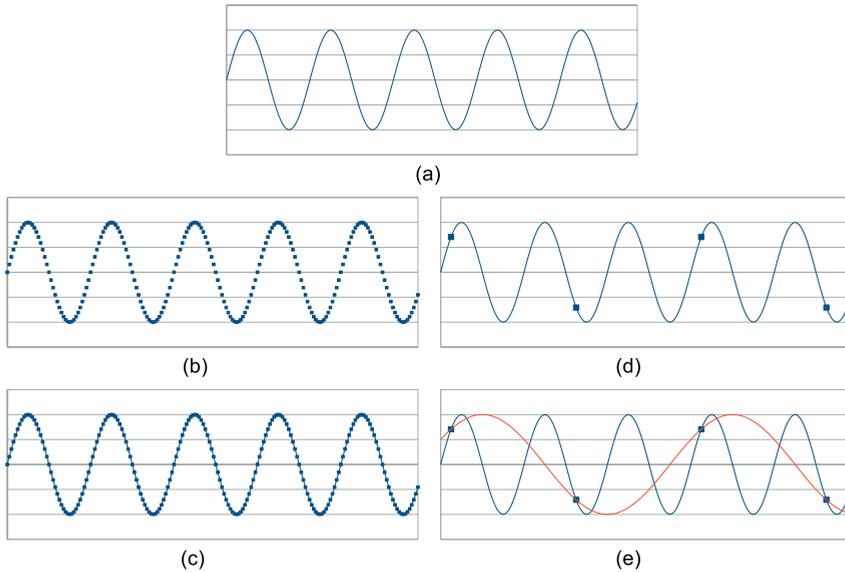


Figura 75. (a) Una señal; (b) muestreada de forma correcta; (c) se puede recuperar perfectamente. La misma señal (d) muestreada demasiado lentamente; (e) se recupera como una señal errónea, demasiado lenta (en rojo)

Hemos puesto 3 casos, pero podría haber muchos más donde se alternen calificaciones justas e injustas. Para entender por qué suceden esas alternancias quiero mencionar una situación análoga que se da al muestrear una señal continua para discretizarla: si la frecuencia de muestreo es demasiado lenta comparada con la máxima frecuencia de la señal, entonces se producen errores por pérdida irrecuperable de información. Dicho informalmente, una señal lenta no sirve para capturar una señal rápida. Este es el famoso teorema del muestreo de Claude Shannon⁷². Y cuando no cumplimos con el teorema, a este tipo de error se le llama *aliasing* (ver un ejemplo en la figura 75). En la medida de inteligencia ocurre exactamente lo mismo. Definitivamente en el caso iii) hay un error de *aliasing*.

⁷² Que dice que la frecuencia de muestreo debe ser mayor o igual al doble de la frecuencia más alta de la señal a ser muestreada.

El *aliasing* es muy frecuente en los exámenes de múltiples opciones, sobre todo cuando estas no son muchas y no están bien discriminadas con suficientes matices. Vemos un ejemplo en la figura 76:

Los virus y los antivirus son la misma cosa.

Cierto.

Falso.

Figura 76. Pregunta de examen, con solo dos respuestas

Una persona que no sepa nada de computadores podría decir: “Sí, me suena que son cosas de computadores. Deben ser lo mismo”. Una persona que sepa un poco más puede decir: “Son dos cosas distintas. Los virus son dañinos y los antivirus luchan contra ellos”. Pero otra persona que sepa más todavía podría decir: “Son lo mismo. Toman el control de tu computador y hacen lo que les da la gana con él”. Y otra que además tenga sentido del humor podría decir: “No. Son cosas distintas porque el antivirus lo instalas tú mismo, mientras que el virus te lo instala un experto”. Y así podríamos continuar. Lo que quiero ilustrar con este otro ejemplo es que el *aliasing* también se puede ver como una forma de comprimir la información, pero con pérdidas. Un montón de opciones se redujeron a solo dos, de modo que no se pueden diferenciar los matices de cada respuesta.

Con la inteligencia (y con la consciencia) hay un aspecto adicional: una persona A puede medir la inteligencia de otra B si la inteligencia de A es mayor, igual o ligeramente menor que la de B (y voy a usar el símbolo \succeq para representar esta relación). Lo de ligeramente menor no creo que se pueda demostrar formalmente⁷³, pero es que en caso contrario el progreso humano habría sido imposible. De ahí se puede deducir que los procesos de enseñanza-aprendizaje permiten aumentar poco a poco los niveles de inteligencia (considerada en sentido amplio como conocimientos y destrezas) de los educandos. Podemos entenderlo también con una batería de ejemplos: aunque yo no cocine bien, puedo saber si la comida que me han puesto es rica; aunque yo no sepa escribir bien en inglés, soy capaz de diferenciar textos bien escritos y mal escritos; aunque yo no tenga buenas ideas, puedo detectar las buenas ideas de los demás; y un largo etcétera. En todos estos ejemplos hay una posibilidad de que me superen, a la vez que yo comprendo que me superan.

⁷³ Aunque tiene relación con la medida de complejidad y los test-C de inteligencia de Hernández-Orallo (1999).

Es decir, puedo evaluar calidades ligeramente mejores a las mías.

A (con una inteligencia I_A) puede medir la inteligencia de B (I_B) $\Leftrightarrow I_A \gtrsim I_B$ Ec. 35

Entonces, la inteligencia no se puede medir de forma absoluta. Es relativa a un árbitro. Ese árbitro puede ser los pares de una revista científica, los padres de un niño, los profesores en un colegio. Y, por ello, depende de la inteligencia del árbitro, de su manera de ver el mundo o de su cultura. Todos los *test* de inteligencia tienen estas limitaciones.

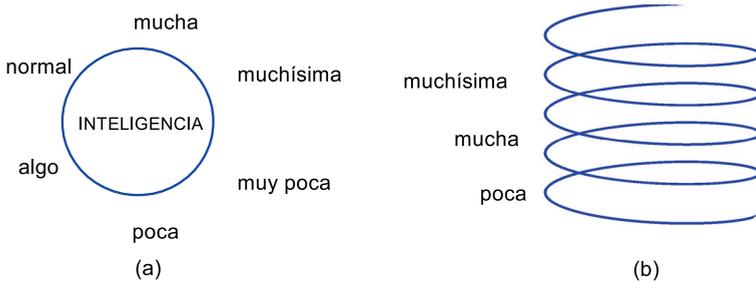


Figura 77. Modelos de como percibimos el comportamiento de un ente según su inteligencia: (a) circular; (b) espiral

La falta de una medida absoluta de inteligencia la podemos entender mejor con un modelo en círculo (Figura 77-a) donde cuando un ente es muy inteligente, sus comportamientos se pueden confundir con los de un ente poco inteligente. Un modelo helicoidal puede ser una alternativa más precisa (Figura 77-b), porque en cada vuelta algo se va ganando, aunque no sea fácilmente medible.

¿SOMOS LOS HUMANOS INTELIGENTES?

Hay que ponerlo en duda.

Para empezar, la pregunta está mal formulada. Tiene ambigüedades. Pues, ¿quiénes somos los humanos? ¿Cualquier humano? ¿Todos los humanos? ¿Un equipo constituido por los más selectos, por ejemplo con el mayor coeficiente intelectual?

Por un lado, un computador operando con un vulgar sistema experto especializado en prospección petrolífera muestra una inteligencia muy superior (dentro de su especialidad) a la mayoría de los humanos, que nunca hemos tenido contacto con esa área de conocimiento.

Por otro lado, la mayoría de los niños a los 5 años muestran una capacidad para desenvolverse en el mundo con la que ningún computador hasta ahora alcanza ni a soñar.



Figura 78. Réplica de una pintura neolítica del museo de Altamira, España

Incluso los antropólogos se preguntan si la inteligencia de los humanos del Paleolítico Superior de hace 15 000 años sería igual que la nuestra, o si la evolución (principalmente cultural) ha moldeado un cerebro bastante distinto y con más capacidades. O sea, si resucitamos a uno de estos habitantes que vivieron y pintaron las paredes de las cuevas de Altamira (Figura 78) ¿sería capaz de graduarse en la universidad? La respuesta, un tanto jocosa pero no por ello menos cierta es que probablemente no, porque la mayoría de los jóvenes de nuestro tiempo tampoco lo logra. Entonces, ¿cuál va a ser el humano que nos represente en estas “olimpiadas de la inteligencia”? Quizás tú mismo quieras participar, pero te pregunto: ¿nunca te has equivocado? ¿Nunca has metido la pata en alguna cosa trivial? Te advierto que los computadores no suelen hacerlo, es casi imposible. Para las tareas triviales son excelentes.

Otro ejemplo: cuando oímos hablar de las competiciones de ajedrez entre Deep Blue y Kasparov, ¿somos conscientes de que Deep Blue es mucho más inteligente jugando al ajedrez que el 99.99% de la humanidad?

Hay muchos casos que documentan la falta de inteligencia de los seres humanos en general. Pinker (1997, p. 321) menciona que si se lanza una bola por un tubo helicoidal hacia abajo, mucha gente cree que al salir por el extremo de abajo seguirá en el aire el trayecto helicoidal. Y es muy conocida la falacia que asume la mayoría de la gente de que si al lanzar una moneda salen muchas caras seguidas, se incrementa la probabilidad de que salga cruz en el siguiente lanzamiento.

Un caso anecdótico documentado, pero que ocurre con frecuencia en otros ámbitos similares, es el del programa Eliza, uno de los pioneros en la inteligencia artificial. La historia es la siguiente: en 1966 el profesor Joseph escribió este programa tratando de parodiar la terapia de psicología *roqueiriana*, que se caracteriza por dejar hablar al paciente sin presionarlo ni asesorarlo sino buscando únicamente una relación de empatía con él. Para ello escribió este programa, que simplemente leía un mensaje introducido por el usuario, buscaba ciertas palabras relevantes (sin analizar ninguna estructura gramatical) y respondía con frases relacionadas a ese asunto. Por ejemplo, si detectaba la palabra “amor” respondía “el amor es muy importante” o cualquier otra frase elegida al azar de un conjunto que tocaba ese tema. Además, si no reconocía ninguna palabra respondía con alguna frase neutra que incitara al usuario humano a continuar escribiendo, como “¡Qué interesante!”, “Dígame que piensa usted de ello” o cosas similares. Un día descubrió que su secretaria hablaba con el programa, pensando que se trataba de un médico real conectado por el *chat* de la época, y le confiaba problemas íntimos e incluso afirmaba que la entendía. En muchos libros se habla de ello como el primer éxito de la inteligencia artificial, pero dado que un programa así lo puede escribir cualquier principiante, lo que deberíamos concluir es que es el primer fracaso documentado de la inteligencia natural.

A esto se le suma que el test de Turing ya se ha pasado de alguna forma, cuando un *software* logró engañar a algunos árbitros en el 2014. Este *software* simulaba ser un niño ucraniano de 13 años llamado Eugene Goostman, que estaba aprendiendo inglés (para así disimular sus fallos en el lenguaje).

En ciertos animales llegamos a reconocer cierto nivel de inteligencia. Tanto los creacionistas como también algunos científicos evolucionistas piensan que el hombre está en la cima de la escala de la inteligencia. ¿Es así realmente? Y, aunque lo fuera en estos momentos, ¿hay garantía de que lo siga siendo para siempre? ¿Podrían surgir, por evolución, seres con un nivel de inteligencia superior al nuestro, en este planeta o en otro?⁷⁴. La respuesta a esta última pregunta es que nada lo impide. Nuestros sistemas de razonamiento son bastante limitados y guiados por emociones, como veremos en el siguiente apartado. Nuestra memoria falla mucho y sobre ella está construida la inteligencia.

La memoria es imprescindible para poder predecir el futuro, partiendo de la base de que el mundo no es aleatorio (al menos, no es completamente aleatorio). Ello significa que el futuro depende, en buena medida, de lo que

74 Recordemos la cita de Mark Twain al comienzo del capítulo anterior.

ocurrió en el pasado. Por eso, cuanto más pasado podamos memorizar, tenemos acceso potencialmente a una mayor capacidad de predecir el futuro.

CLASIFICACIÓN DE LOS PROBLEMAS DE IA

Se dice en broma, pero en serio, que si un problema ya está resuelto entonces no es de IA. Eso quiere decir que ganar en el ajedrez requiere inteligencia hasta el día en que alguien encuentre el algoritmo, y entonces se volverá trivial. Esto es terrible porque pareciera que el área de la IA se destruye a sí misma.

Supongamos que no sabemos nada de aritmética, y nos dicen que Pepito acaba de comerse 2 manzanas en horas del almuerzo, con lo cual solo le quedan 6. ¿Con cuántas manzanas llegó al colegio esta mañana?

Sabiendo de aritmética es muy sencillo. Basta con traducir el enunciado a la ecuación:

$$x-2=6$$

Y despejar la x:

$$x=6+2=8$$

Pero imaginemos que no sabemos nada de esto. ¿Qué recurso nos queda? Pues únicamente el de “prueba y error”:

- Supongamos que llegó con 14 manzanas. Se comió 2, por lo que le quedan 12. Entonces esa no es la solución.
- Supongamos que llegó con 3 manzanas. Se comió 2, por lo que le queda 1. Entonces esa no es la solución.
- Y así podemos seguir hasta que, por fin: supongamos que llegó con 8 manzanas. Se comió 2, por lo que le quedan 6, tal y como dice el enunciado del problema. Por tanto, la solución es 8 manzanas.

Después alguien podría crear heurísticas. Por ejemplo: la solución debe ser un número par, porque ambos números del problema también son pares. Con esas heurísticas se limita el espacio de búsqueda.

En definitiva, para no enrollarme más, cuando Abu Abdallah Muhammad ibn Mūsā al-Jwārizmī inventó el álgebra, ocurrieron dos cosas:

- Todos esos problemas se volvieron tratables algorítmicamente, es decir, dejaron de ser de Inteligencia Artificial.
- al-Jwārizmī nos mostró lo inteligente que era, al resolver de forma general un problema complejo.

Por eso aquí proponemos clasificar los problemas de IA no resueltos en tres tipos. Ordenados desde el más sencillo hasta el más complicado son:

Solución puntual a un problema

Cuando los datos de entrada son fijos. Ejemplo: dada una colección de objetos de pesos y valores conocidos, y dada una mochila que admite un peso máximo P , decidir si es posible llenarla sin exceder su capacidad y de modo que el valor total de los objetos empaquetados sume V o más. Este es un problema de decisión NP completo. Y todavía no sabemos si la clase de problemas NP coincide con P. Mientras no lo sepamos, usaremos técnicas de inteligencia artificial para solucionarlo, así sea aproximadamente.

Muchas veces se conoce la entrada y la salida, y lo que falta es la función, mapa o algoritmo que lleve de una a otra. Ejemplos de ello son el ajedrez donde conocemos la disposición inicial de las piezas, conocemos todas las posibles disposiciones finales de las piezas que nos dan la victoria, y queremos saber cuál es el camino de jugadas que nos lleva de la primera situación a la segunda (Figura 79).

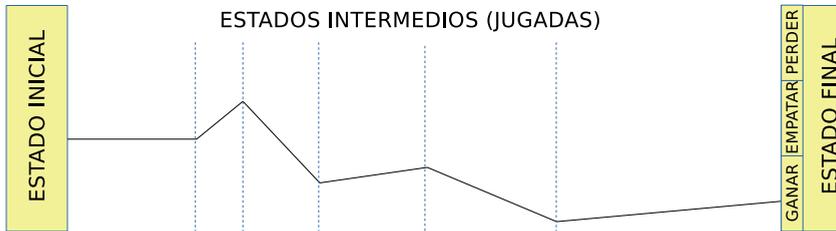


Figura 79. ¿Cuáles son los estados intermedios para llegar del estado inicial al final

Otro ejemplo interesante se ve en el problema 16.

Problema 16: Salvar el foso

Rodeando un castillo hay un foso de 5 metros de anchura, con la forma que se ve, y está lleno de cocodrilos hambrientos. Usted está afuera y quiere salvar el foso, para lo que dispone de 2 tablones angostos de 4,9 metros de largo. No dispone de nada más, ni de cuerdas, ni de pegamento, ni clavos, ni herramientas. Absolutamente nada más. ¿Cómo puede hacer para llegar al castillo?

Bien sea que hayas encontrado la solución o que la hayas buscado en los apéndices, te darás cuenta que hay un estado intermedio por el que es obligatorio pasar para llegar del estado inicial al final. Este estado interme-

dio es una especie de invariante de la solución, también llamada punto fijo. Cualquier solución (hay un pequeño conjunto de variantes, en función de la anchura de los tablonés) debe pasar por allí.

Otro ejemplo muy conocido es resolver el problema de las Torres de Hanoi. Hay tres torres, la central y la izquierda vacías, mientras que la derecha tiene N anillas de tamaños sucesivos, como se muestra en la figura 80-a. El objetivo es pasar todas las anillas a la torre de la izquierda (Figura 80-b) cumpliendo dos reglas: solo se puede manipular una anilla a la vez (Figura 80-c); y no puede haber una anilla grande encima de una anilla pequeña (Figura 80-d).

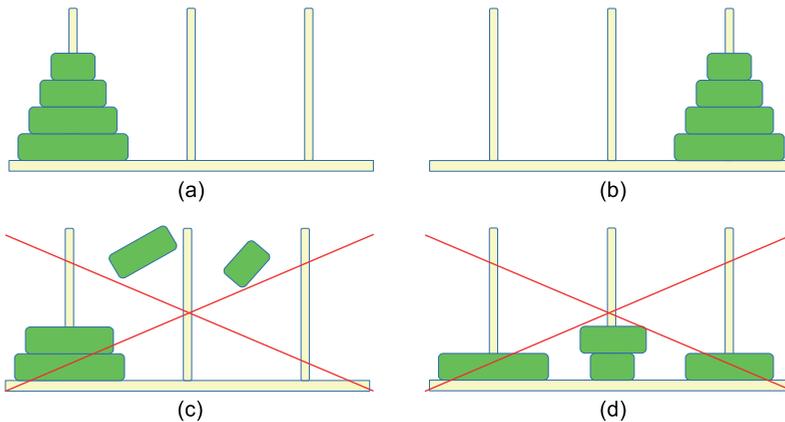


Figura 80. Torres de Hanoi: (a) posición inicial; (b) posición final; (c) prohibido mover más de una anilla a la vez; (d) prohibido poner una arandela encima de otra de menor tamaño

Sabemos la posición inicial, sabemos la posición final y necesitamos averiguar el algoritmo que me lleva de una a otra, cumpliendo ciertas restricciones. En este caso también hay un punto intermedio fijo: en algún momento del juego tenemos que pasar por la posición de la figura 81, de modo que la anilla más grande se pueda llevar de su torre inicial a su torre final.

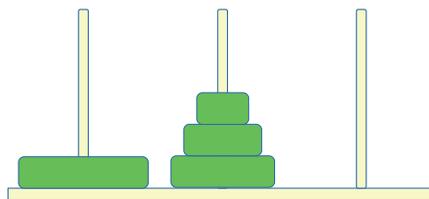


Figura 81. Punto fijo en Hanoi

Y cuando comprendes que esa jugada es ineludible, eso te permite diseñar el algoritmo, que es recursivo:

Recuadro 7: Solución a las Torres de Hanoi, en Ruby

```

def mover(numeroDeAnillas, torreInicial, torreFinal, torreAuxiliar)
  if(numeroDeAnillas == 1)
    # moverla, lo cual depende de la estructura de datos de las torres.
  end

  mover(numeroDeAnillas-1, torreInicial, torreAuxiliar, torreFinal)
  mover(1, torreInicial, torreFinal, torreAuxiliar)
  mover(numeroDeAnillas-1, torreAuxiliar, torreFinal, torreInicial)
end

```

Dicho en español, mover N anillas de la torre inicial a la final se puede lograr con tres pasos: mover $N-1$ anillas de la torre inicial a la auxiliar, mover 1 anilla de la torre inicial a la final (el punto fijo) y mover $N-1$ anillas de la torre auxiliar a la final (Figura 82).

En este algoritmo no se ha explicitado la condición de parada que se da cuando el número de anillas es 1, para no llenarlo de detalles accesorios dependientes de la implementación. Y esa es la única jugada que hacemos realmente: mover una anilla de una torre a otra, lo cual es trivial.

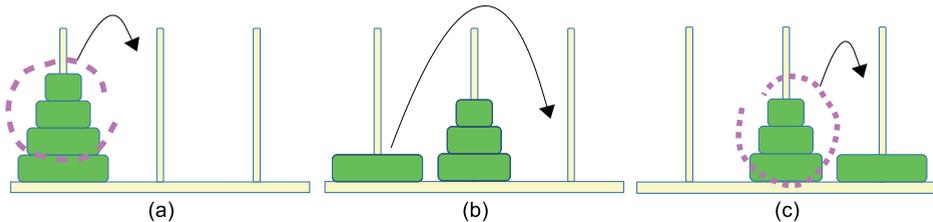
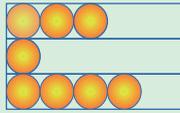


Figura 82. Solución: (a) mover $N-1$; (b) mover 1, que es el punto fijo; (c) mover $N-1$

Como vemos, encontrar las invariantes de un problema es importantísimo para guiarnos hacia la solución o, en general, hacia el conjunto de soluciones válido. Por ejemplo, Dijkstra popularizó una demostración que dice que todo lenguaje de programación debe tener obligatoriamente 3 construcciones para ser Turing-completo: secuencia, decisión y repetición (Estructurado_Wiki, 2017). Si vas a diseñar un nuevo lenguaje de programación, no te olvides de estas 3 estructuras de control. Si se te olvida alguna, tu lenguaje no tendrá capacidad expresiva para resolver los problemas de algoritmia que pueden resolver los demás lenguajes. Todos los lenguajes deben pasar por allí, si quieren ser útiles.

Hay muchos juegos que contienen una invariante que, si la conoces, te guía a ganar siempre (ver el problema 17).

Problema 17: NIM



Hay 3 filas de monedas y dos jugadores. Por turnos, cada jugador escoge una fila y de allí se lleva todas las monedas que quiera. Gana el que se lleve la última moneda. ¿Qué debes hacer para ganar?

Solución general a un problema

Cuando los datos de entrada son arbitrarios. Ejemplo: encontrar un algoritmo que resuelva cualquier problema de tipo *knapsack* (todavía no existe tal algoritmo general) o que resuelva a Pepito sus problemas con el almuerzo (esto es lo que hizo en su día al-Jwārizmī inventando el álgebra).

Solución general a todo tipo de problemas

Y con datos de entrada arbitrarios o incluso con ruido o ambiguos o mal especificados. Bueno, en eso estamos en el presente capítulo. Este es realmente el “único” problema de la IA.

Problemas clonables

También hay otra forma de clasificarlos dependiendo de si el problema es o no es clonable. La mayoría de los problemas de juegos, como el ajedrez y el Go, son clonables mientras que la mayoría de los problemas de la vida práctica no lo son.

La idea de la clonación es la siguiente: estoy jugando una partida de ajedrez contra el computador. Entonces puedo sacar un clon (una copia) del tablero y del estado de mi oponente (el computador) de tipo copia profunda⁷⁵, es decir, que no haya ningún tipo de conexión, intersección o interacción con la partida real. Sobre ese clon ensayo una jugada a ver cómo me va. Puedo sacar un número indefinidamente grande de clones, para intentar probar un gran número de jugadas, usando algún método mío o, mejor aún, algún algoritmo de IA. Cuando ya haya explorado lo suficiente y tenga claro cuál es mi mejor jugada posible, borro todos los clones y regreso a mi partida original, efectuando allí mi jugada.

En problemas de la vida real no suele ser factible clonar nada, habitualmente porque hay variables o estados desconocidos, inaccesibles o muy complejos de modelar. Por ejemplo, si compro acciones en la bolsa no hay forma de clonar el sistema para hacer ensayos no destructivos. Cualquier operación que haga sobre la bolsa tendrá consecuencias en el futuro. No hay forma de borrar ninguna acción realizada.

⁷⁵ Deep copy.

Lo mismo ocurre si intentamos enseñar algo a un niño. Cualquier error metodológico o de conocimiento que cometamos, quedará memorizado en el niño. No hay vuelta atrás. No es posible clonar para probar tácticas y seleccionar la mejor.

No obstante, hay muchos niños. Se pueden ensayar unas metodologías con unos y con otros, y comparar los resultados. Lo cual, en cierto modo, equivale al clonaje. Pero en la bolsa no está claro que se pueda hacer. Incluso aunque hay bolsas en cada país, se influyen mutuamente. No las podemos considerar clones separados. Y si dejamos pasar mucho tiempo entre un experimento y otro para que se disipen los efectos anteriores, la humanidad, la economía y las tecnologías habrán cambiado por lo que no es sensato considerar las situaciones muy separadas en el tiempo como clones.

La física experimental se basa en la suposición de que es posible clonar exactamente los experimentos (el péndulo funciona igual en todos los laboratorios de la Tierra), aunque eso solo es una suposición. Nadie sabe si el Universo guarda un estado global o si las coordenadas absolutas de un experimento afectan al resultado. A decir verdad, parece poco plausible. Sin embargo, en algunos experimentos (como cuando hay enmarañamiento cuántico⁷⁶) podría ser así.

Si recordamos el apartado “¿Qué es un simulador?”, cuando el problema es clonable, estoy creando un **tiempo interno** para realizar las simulaciones en los clones, que luego desaparece sin dejar rastro cuando selecciono la mejor acción y la ejecuto en el problema real. Allí es que el **tiempo externo** avanza un *tic* de reloj. En este sentido, los simuladores nos dan inteligencia gracias a que permiten hacer clones de los problemas y revisar así muchas hipótesis antes de tomar alguna decisión.

También podemos entender los problemas clonables como el requisito previo para que pueda trabajar (aunque sea solo en teoría) una Máquina de Turing No Determinista. Esta máquina debe fabricar clones del problema para ensayar en él todas las soluciones posibles de forma paralela y simultánea. La evolución también hace algo parecido: fabrica clones, con pequeñas mutaciones y ensaya cada uno a ver cómo se comporta solucionando un determinado problema. En este último caso la diferencia estriba en que lo que se clonan son las soluciones, en vez del problema.

Daniel Dennett (2000) nos plantea algo bastante relacionado en su libro *Tipos de Mentes*. Allí nos dice que hay cuatro tipos de mentes que, en orden progresivo de complejidad, son:

⁷⁶ *Quantum Entanglement*: cuando dos partículas interactúan, se mantienen en un estado cuántico compartido que perdura incluso aunque se separen millones de kilómetros.

- **Tipo I, mentes *darwinianas*.** Cada individuo implementa una idea, un algoritmo acerca de cómo comportarse en el mundo y que está escrito en *hardware*, fijo para siempre (habitualmente en los genes). Si el algoritmo es malo, el individuo morirá rápidamente sin dejar hijos. Cuanto mejor sea el algoritmo, el individuo tendrá más hijos que se parecerán a él (con algunas mutaciones que le darán variedad a los algoritmos). La evolución opera aquí seleccionando a los individuos con las mejores ideas. Cuando vemos animales resolver problemas, hay gente que dice que ellos no son inteligentes sino que usan su instinto. Esta es una interpretación equivocada. Lo que ocurre es que la inteligencia no está en el individuo sino en la población⁷⁷, y el aprendizaje ocurre por medio de la evolución.
- **Tipo II, mentes *skinnerianas*.** El algoritmo implantado en los genes es un poco más flexible. Si una acción recibe recompensa, el algoritmo aumenta la probabilidad de repetirla. Mientras que si recibe un castigo no mortal, el algoritmo disminuye su probabilidad. Esta realimentación conduce a un aprendizaje.
- **Tipo III, mentes *popperianas*.** Cada individuo ha construido en su cerebro un modelo del mundo, es decir, un simulador. El modelo se entrena con lo que el individuo percibe. Cuanto más inteligente sea más detalles del mundo captura en su modelo. Si es poco inteligente, el modelo podría congelarse y dejar de aprender, incluso ante eventos externos que no casen en absoluto con su modelo (sistema de creencias consolidado). Cuando el individuo tiene que tomar una decisión, primero lleva la acción que desea tomar al simulador para ver si los resultados son buenos o malos. De este modo mueren las ideas, pero no los individuos. El simulador permite hacer clones de la realidad y ello supone un gran salto en la inteligencia de los seres que lo poseen, típicamente los mamíferos.

Al respecto, un simulador es una arquitectura que permite probar modelos del mundo. El proceso de aprendizaje consiste en construir nuevos modelos y también en ajustar sus parámetros. Ser más o menos inteligente depende de tener modelos más o menos sofisticados, que van a predecir mejor o peor el futuro. Si quieres saber qué pasa cuando se sube el IVA del 15% al 18% puedes comparar los dos modelos en la figura 83. El primero es lineal, el segundo es complejo y tiene en cuenta muchos otros factores, como la tasa de desempleo, la tasa de comercio informal, la tasa de fraude y el cambio de la moneda.

⁷⁷ Como argumentaré más adelante, la inteligencia siempre es un fenómeno colectivo.

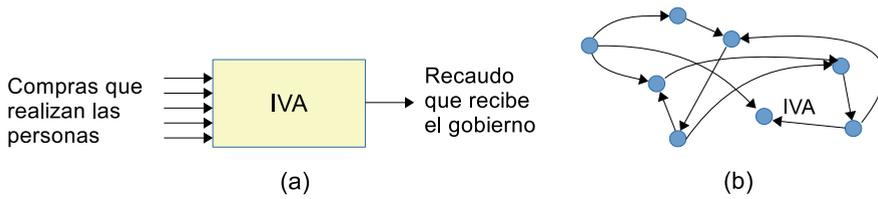


Figura 83. Dos modelos de la economía: a) lineal; b) complejo

No hace falta ir a campos tan abstrusos como la economía. En la vida cotidiana mucha gente tiene modelos erróneos. Por ejemplo, piensan que si se cae comida al suelo, disponen de 3 segundos para recogerla antes de que se contamine (Figura 84). Aquí se trata de un mero ajuste de parámetros: el primer modelo es equivocado pues supone que en el suelo hay solo unas pocas bacterias, y que ellas corren detrás de la comida. Mientras que el segundo modelo tiene un valor más realista de 10^9 bacterias por metro cuadrado, y cualquier cosa que caiga al suelo se unta inmediatamente.



Figura 84. Dos modelos de contaminación de la comida que cae al suelo

Fuente: Dibujos realizados y cedidos amablemente por Helga R. Calvo R. (2017).

Los niños pasan muchas horas entrenando sus simuladores. Es lo que llamamos jugar. Y, una vez adultos, los simuladores pueden servir para algo más, como indica Yuval Noah Harari (2016): explorar mundos imaginarios muy simples con alguna característica destacable. Son las ideologías económicas (con la idea del dinero para intercambio generalizado), religiosas (con un dios que premia y castiga), sociales de diversos tipos (con ideas como la libertad, el derecho al trabajo o la solidaridad) e incluso deportivas (clubes de fútbol). Lo interesante que propone Harari es que es muy fácil inducir a las personas a que ejecuten en su simulador estas ideologías (lo que llamamos imaginar) y, una vez que ello ocurre, los respectivos líderes consiguen controlar grandes poblaciones de humanos con muy poco esfuerzo. Los ani-

males no tienen esta capacidad por lo que leones y chimpancés están en lucha física permanente para mantener su liderazgo dentro de su manada que, a su vez, no puede ser muy grande.

De modo que el simulador que traemos de fábrica en nuestro cerebro es una característica sobresaliente que nos permite desenvolvernos mejor en el mundo, pero tiene un punto flaco y es que da acceso a otras personas para que nos manipulen. E incluso gracias a él se forman los superorganismos sociales. De alguna forma es también el policía interno que nos vigila, una vez que dejamos que se ejecute en él una cierta ideología, para que nos alineemos con lo que la sociedad espera de nosotros.

Por último, esta idea de Harari de que el *hackeo* del simulador cerebral permite crear ficciones que agrupan a los humanos en superorganismos, encaja muy bien con el concepto de memes ideado por Richard Dawkins (1994) y desarrollado por tantos otros científicos, entre los que destacan Daniel Dennett, Susan Blackmore y Robert Aunger (2002). Los memes son unidades básicas de información tales como teorías científicas, mitos religiosos, modas, temas musicales, obras de arte y prácticamente cualquier idea que podamos pensar. Los memes son los individuos de un proceso evolutivo, pues cumplen con las cuatro condiciones requeridas para que haya evolución. Hay una población de memes, que sacan copias de sí mismos saltando de un cerebro humano a otro. Siempre habrá variabilidad, pues las ideas se transmiten con errores o se combinan unas con otras. En este escenario, los cerebros humanos serían entonces las máquinas de reproducción de los memes. Lo que Harari aporta muy brillantemente es el mecanismo de presión selectiva, que nunca había sido explicado de forma convincente por los otros autores. Los memes que más se reproducen son los que fabrican las ficciones que tienen más éxito en el cerebro de los humanos, generando grupos sociales cohesivos que propagan esas mismas ideas.

- **Tipo IV, mentes gregorianas.** La mente anterior puede predecir el futuro, mientras que esta también puede modificarlo, pues es capaz de construir y usar herramientas (desde martillos hasta computadores) incluyendo un lenguaje para comunicarse con otros individuos. Aquí Dennett comete un error de categorización, pues este tipo de mente existe, efectivamente, pero no es la culminación de las tres anteriores sino que corre paralelo con ellas. La razón de ello es que construir herramientas no es tanto una habilidad cerebral, que por supuesto lo es, sino sobre todo la disponibilidad de tener una mano o algo similar, capaz de manipular el mundo. No es un asunto de potencia de cómputo sino de

versatilidad en el sistema de entrada-salida, usando la jerga de los computadores. Este nivel no lo pudieron alcanzar los delfines, a pesar de que son muy inteligentes y seguramente tienen una mente *popperiana*. Sí lo pudieron alcanzar los humanos, los simios en general y los elefantes con su trompa. Les queda muy difícil a los caballos y las vacas. Y de la misma manera que la Máquina de Turing Universal logra computación completa, la mano logra entrada-salida completa, ya que puede construir cualquier mecanismo incluyendo microscopios para ver lo muy pequeño, telescopios para ver lo muy grande, sensores para recibir información de entrada que antes no era accesible (por medio de infrarrojo, ultravioleta, rayos X, vibraciones o detección de productos químicos), e incluso otras máquinas para fabricar otros mecanismos imposibles de hacer a mano, donde también se incluyen máquinas de fabricación universales como las impresoras 3D, y máquinas de cómputo.

De hecho, hay muchos animales que construyen herramientas, aunque la mayoría lo hace porque están impresas en sus genes o, como dice Dawkins (1982), forman parte de su fenotipo extendido. Ejemplos son los termiteros, las telas de arañas, las madrigueras y los nidos. En la figura 85 podemos ver el nido de un tipo de araña, cuyos genes construyeron la tapa. Esta tapa se podría considerar como una parte de su cuerpo, a la que no hay que alimentar con oxígeno y azúcar. Pero, al fin y al cabo, también es una herramienta que le sirve para que no entre la lluvia o enemigos a su nido. Otros ejemplos de herramientas, esta vez no impresas en genes sino transmitidas culturalmente, son las piedras para abrir semillas que descubren algunos simios. El lenguaje de cada especie también puede considerarse una de estas herramientas, y parece que tiene las dos componentes: una genética y otra socialmente aprendida (Pinker, 1994).



Figura 85. Nidos de araña con tapa, en la Chapada Diamantina, Brasil

¿CÓMO RAZONAN LOS ROBOTS?

Si quieres profundizar en las últimas tendencias, deberías leer el libro de Pedro Domingos, *Machine Learning* (2015). Para él, la inteligencia también es un asunto de predicción, con la diferencia de que, en aprendizaje de máquina, los algoritmos ya no los escriben los humanos, sino que el computador los aprende por sí mismo. Pedro clasifica estos algoritmos en cinco técnicas: razonamiento simbólico, redes neuronales, computación evolutiva, razonamiento *bayesiano* y razonamiento por analogías. Veremos unas nociones someras de cada una de ellas.

Razonamiento simbólico. Es puramente lógica matemática. En sus inicios se basaba en los lenguajes *LISP* y *PROLOG*, que se pensaba iban a ser la base de la inteligencia artificial. Se implementa con árboles de decisión o con sistemas de reglas clasificatorias de tipo *IF-THEN-ELSE*. Se necesita un experto que programe las reglas adecuadamente, aunque en el libro anterior también vimos que se puede hacer evolutivamente. Al ser exacto y determinista se le aplica el teorema de Gödel, es decir, no puede ser creativo, no puede inventar nuevos espacios y conceptos. Pero si puede explorar sistemáticamente todo un espacio ya existente. Tampoco puede lidiar con incertidumbres, probabilidades, especificaciones contradictorias ni ambigüedades, de modo que está muy poco adaptado a los problemas reales que tenemos que solucionar los humanos. Esto quizás lo entienden poco los filósofos: el problema del marco⁷⁸, la objeción de Ada Lovelace⁷⁹ y otros obstáculos a la inteligencia artificial solo aparecen en este tipo de algoritmos.

Para operar requiere únicamente disponer de una Máquina de Turing Universal. Y ya hemos visto en el libro anterior que estas máquinas son de muy baja complejidad. Esta es la única técnica que ofrece resultados exactos, lo que, paradójicamente, es su principal desventaja ya que ello implica diversas limitaciones que la impiden alcanzar la verdadera inteligencia, aunque puede ser un apoyo para las otras técnicas.

Redes neuronales. Son un modelo simplificado de la arquitectura *hardware* del cerebro humano. Cada neurona se puede considerar un nodo donde se efectúa un cómputo a partir de los valores que llegan a través de sus arcos de entrada, que simulan las dendritas de una neurona real. El resulta-

78 Diferenciar lo relevante de lo irrelevante para resolver un problema de forma eficiente. Por ejemplo, si tengo setas de varios colores y quiero saber cuántas son, ¿el color es relevante? ¿Y si quiero comerlas sin envenenarme? Cualquier problema tiene miles de aspectos irrelevantes, y descartarlos sistemáticamente puede consumir demasiado tiempo.

79 Los computadores no son inteligentes. Son inteligentes los humanos que los programan.

do se ofrece por un arco de salida, que simula el axón. Los arcos de entrada están etiquetados con unos pesos fijos, y están conectados a las salidas de otras neuronas. Cada neurona calcula el sumatorio de sus entradas multiplicadas por los respectivos pesos, a lo que aplica una función de activación y si el resultado supera un umbral, la salida se activa.

Las redes neuronales se suelen diseñar con varias capas de neuronas, típicamente tres. La primera capa recibe las entradas del problema, la segunda capa hace un procesamiento, y la tercera capa genera las salidas, o sea, los resultados.

Hay una primera fase de entrenamiento supervisado, donde se le presentan a la red neuronal problemas con sus soluciones. La red entonces ajusta los pesos de sus conexiones entre neuronas, usando algoritmos como el *backpropagation*. También hay otras redes capaces de hacer aprendizaje no supervisado, donde se les presenta problemas de entrada, pero no se les dice cuál es la solución correcta.

La principal limitación de estas redes es que los algoritmos de aprendizaje donde calculan los pesos, suelen ser lentos y no convergen cuando hay más de tres capas. Sin embargo, se ha inventado un nuevo tipo de arquitectura que solventa este problema: el *autoencoder*. Consiste en una red de tres capas que intenta predecir sus entradas, es decir, la tercera capa está conectada a la primera. El entrenamiento es no supervisado y cuando ya están entrenadas, la segunda capa contiene aproximadamente la misma información que la entrada pero comprimida. Se puede forzar que haya compresión poniendo menos neuronas en la segunda capa que en la primera. Recordemos que comprimir es una forma de aprender o, si se quiere, de generalizar. Por ejemplo, si en la primera capa hemos puesto los *pixeles* de la foto de un gato, y hemos entrenado la red para que en la última capa aparezca más o menos esa misma foto, entonces en la capa intermedia habrá algo esencial de la foto a partir de lo cual se puede reconstruir el original.

Otra forma más sofisticada de forzar la compresión es poniendo muchas más neuronas en la segunda capa que en la primera pero obligando a que la mayoría de ellas no se activen. Con ello se consigue que el entrenamiento sea más fácil y, a la vez, que el resultado sea un conjunto de características separadas de la entrada. Ello da lugar a que la entrada quede codificada de una forma similar a una matriz dispersa, también usada por Hawkins en Numenta (2005). Es interesante destacar que el cerebro humano realiza algo similar, aunque de forma distribuida.

Nada impide apilar *autoencoders*. Es decir, usar un segundo *autoencoder* para comprimir aún más la información que entrega el primero. O, dicho de otra manera, para asociar de otra forma las características que entrega el

primero. Y luego utilizar un tercero para hacer lo mismo con el segundo, y así sucesivamente. De este modo, si a la entrada ponemos los *pixeles* de la foto de un gato, quizás el primer *autoencoder* extraiga características básicas de la imagen, como segmentos, zonas de color similar y así. El segundo *autoencoder* puede extraer siluetas, separar el primer plano del fondo y otras operaciones resultantes de analizar las características entregadas por el primero. El tercer *autoencoder* quizás sea capaz de reconocer la forma del gato analizando los datos entregados por el segundo.

Hay un éxito impresionante con las nuevas arquitecturas de redes neuronales, gracias a los *autoencoders* y al gran número de capas que se pueden poner u entrenar. Y eso es lo que actualmente se llama *Deep Learning* (Géron, 2017). Estas redes son las que ganaron el campeonato mundial de Go, las que reconocen gatos en videos, las que aprenden a jugar y ganar cualquier juego de computador por sí mismas.

La complejidad de estas redes neuronales es bastante baja pues no necesitan computación completa ni para ejecutar la red ni para entrenarla, y estructuralmente son solo un grafo dirigido.

Algoritmos evolutivos. Dado un problema, el programador crea un patrón general para la solución, llamado cromosoma. Ese patrón es un contenedor de datos y programas que se requieran para calcular la solución. Inicialmente los valores concretos de los datos y programas son desconocidos, y lo que se busca es averiguar cuáles son. Se parte de una población de cromosomas generados al azar que se hacen evolucionar, para lo cual se calcula la aptitud de cada uno —es decir, lo bueno o malo que es— seleccionándose los mejores con mayor probabilidad, para reproducirlos. La reproducción puede hacerse de muchas formas, pero lo que se busca es que los hijos se parezcan a los padres, pero que no sean completamente idénticos. Eso se logra con la mutación, que consiste en hacer un pequeño cambio a un padre, y con el cruce, que consiste en tomar una parte de un cromosoma y otra parte de otro, combinándolos entre sí. Los hijos se introducen en la población y opcionalmente hay otro proceso de selección para matar algunos cromosomas. Así se completa una generación. El algoritmo corre por muchas generaciones hasta que aparezca alguna solución aceptable.

Puede haber cromosomas que mantengan hipótesis contradictorias, lo cual es bueno, pues se deja que la evolución decida cuál es la mejor. Y la solución mejor encontrada seguramente no es la mejor posible, pues conforme se acercan al óptimo la presión selectiva disminuye.

Es un algoritmo lento, pero muy general, pues se puede aplicar a cualquier tipo de problema, incluso los que están mal definidos. Se puede considerar que cada cromosoma es una hipótesis respecto a la solución de un

problema, y que el proceso evolutivo genera nuevas hipótesis y selecciona las que están mejor adaptadas para solucionar el problema.

La principal limitación es que, una vez definido el cromosoma, el espacio de búsqueda queda allí limitado, es decir, la creatividad del algoritmo se desarrolla dentro de ese espacio, pero no puede crear espacios nuevos. Hemos visto muchas variantes de estos algoritmos en el libro anterior.

Estos algoritmos están inspirados en la evolución biológica que descubrió Darwin. Pero en el cerebro humano ocurren también procesos similares para la toma de decisiones: consideración de un conjunto de hipótesis, fabricación de otras por medio de pequeñas variantes, selección de las mejores, iterar unas cuantas veces y al final elegir la mejor.

La complejidad de un algoritmo evolutivo es prácticamente la misma que la reproducción. Si un objeto sabe reproducirse, su evolución es prácticamente inevitable.

Razonamiento bayesiano. Se tiene un conjunto de hipótesis mutuamente excluyentes y se tienen creencias sobre lo bien que funciona cada una de ellas en un problema específico. Las creencias se modelan con probabilidades, también llamadas probabilidades *a priori*. Habitualmente se escoge la hipótesis en la que se crea más (la que tenga la mayor probabilidad de éxito) para tratar de solucionar el problema. Una vez que sabemos la solución al problema, tanto si hemos acertado como si no, actualizamos nuestras probabilidades *a priori*, usando el famoso teorema de Bayes.

Razonamiento por analogías. Aquí Pedro agrupa varios algoritmos no completamente relacionados: la búsqueda del vecino más cercano, las máquinas de soporte vectorial y el razonamiento analógico completo. El objetivo de razonar por analogías es poder llegar a conclusiones aproximadas aunque casi no haya datos de entrenamiento. Basta con un ejemplo sobre un tema específico, para poder trabajar con esta técnica y comenzar a obtener resultados, así sean aproximados.

Hay también combinaciones de todo ello, como la lógica difusa, lógica multivariante y la lógica temporal, que están basadas en el razonamiento simbólico, pero cada proposición tiene un *tag* añadido que expresa algo nuevo que con la lógica formal no se puede, como por ejemplo, la probabilidad de que la proposición sea válida, el intervalo de tiempo de validez, el grado de verdad, y cosas similares. Y se han unificado varios marcos teóricos, aunque no todavía los cinco.

Pero eso no significa que aquí terminen todas las posibilidades. Quedan muchos temas por explorar, y para saber cuáles, nada mejor que entender en qué se basan los razonamientos humanos.

¿CÓMO RAZONAMOS LOS HUMANOS?

Para poder fabricar inteligencia artificial es una buena idea fijarnos cómo funciona la inteligencia humana. Es interesante conocer el detalle de las conexiones neuronales, pero no con el objetivo de imitarlas como intentan algunos proyectos de investigación. Eso no tiene mucho sentido. El objetivo será sacar modelos que capturen lo importante de la inteligencia, eliminando los detalles accesorios. Recordemos que, inspirados en las aves, hemos fabricado aviones, pero ninguno necesita de detalles como las plumas.

Vuelvo a insistir en que la inteligencia biológica es el resultado de un proceso evolutivo, y que el problema que trata de resolver es sobrevivir y reproducirse dentro de grupos sociales, y no calcular integrales triples. Eso puede llegar después, como un efecto secundario.

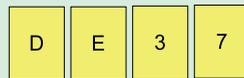
¿Cómo apareció por evolución la inteligencia humana? Hay varias hipótesis y me centraré en la más ampliamente aceptada. Como argumenta Dawkins (1994), la cooperación dentro de un grupo no puede tener un sustento genético. Y la cooperación es lo mínimo indispensable para que el grupo exista, porque de lo contrario le iría mejor a cada uno viviendo aislado de los demás. Sin embargo, es un hecho que existe la cooperación en los grupos de primates e incluso de otros mamíferos. Cuando un grupo coopera para conseguir un objetivo común, es inevitable que alguien trate de fingir cooperación sin realmente trabajar para el grupo, obteniendo de todos modos los beneficios de pertenecer a él. Como bien explica Pinker (1997), ello debió crear una carrera de armamentos entre cerebros que se hacían cada vez más grandes tratando de detectar los engaños de los otros miembros de la sociedad para evitar abusos, a la vez que intentando engañar a los otros para su propio beneficio. Parece ser que esa es la causa de que el cerebro del *Australopithecus Afarensis*, de 400 cm³ se agrandara en 4 millones de años hasta llegar a nuestros 1500 cm³ en promedio, a la vez que aumentaba el tamaño de la sociedad en la que vivíamos.

En esa carrera de armamentos, para lograr engañar a los demás, puede ser una buena táctica engañarse primero a sí mismo, como propone Trivers (2011), de modo que el otro no pueda leer en el rostro alguna incongruencia y así no dejar ningún resquicio a la duda. En general no es razonable esperar que un animal quiera engañarse a sí mismo, pues al alejarse de comportamientos óptimos podría poner en riesgo su vida. Sin embargo, en ambientes sociales es distinto. Juan Camilo Ramírez fue un estudiante asiduo a EVALAB, aunque su trabajo de grado finalmente no lo realizó aquí, y ha seguido trabajando en temas evolutivos durante sus estudios de doctorado. En Ramírez y Marshall (2015) nos muestra un modelo donde el autoengaño

evoluciona y prospera. Como efecto secundario, un cerebro tan grande diseñado para predecir y anticipar el comportamiento de sus congéneres puede servir para muchas otras cosas. Eso es lo que se llama *exaptación*. El cerebro humano no fue diseñado para las matemáticas y la ingeniería (los computadores son más precisos y más rápidos), pero puede servir para ello. Y es importante saber que, debido a su origen cooperativo, la forma de razonar de los humanos está siempre impregnada de un tinte social (ver problema 18).

Problema 18: Detectar engaños

a) Nos muestran las siguientes cuatro cartas que tienen una letra por un lado y un número por el otro y nos aseguran que “si una carta tiene la letra D en un lado, entonces tiene el número 3 en el otro”:



¿Qué cartas hay que girar para saber si esa afirmación es verdadera?

===

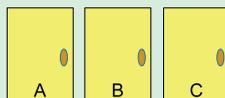
b) Supongamos que usted es un camarero en un bar y está haciendo cumplir la regla “si una persona está bebiendo cerveza, ella debe haber cumplido 18 años”. Usted puede preguntar a las personas que quiera qué están bebiendo o cuál es su edad. En la barra hay cuatro personas: un bebedor de cerveza, un bebedor de naranja, una persona de 25 años y una persona de 16 años.

¿Qué debe preguntar y a quién?

A la mayoría de la gente los problemas abstractos no se le dan bien, pero los sociales sí. Por eso las matemáticas suelen ser difíciles para casi todos los jóvenes, aun cuando sean muy inteligentes. Muchas ramas de las matemáticas tienen un fundamento básico que puede tener un origen social, como contar objetos para intercambiarlos o repartirlos. Pero otras son demasiado abstractas y hay que reconocer el gran avance que supuso su mera formulación. Un ejemplo es el teorema de Bayes, y en general, toda la estadística (ver el problema 19).

Problema 19: Las tres puertas

Estamos en un concurso de televisión, donde nos muestran 3 puertas. La mecánica del concurso es la siguiente: detrás de una de ellas está el premio, mientras que en las otras no hay nada. El presentador, quien sabe perfectamente en qué puerta está el premio, nos pide que elijamos una, pero todavía no la abre. Para aumentar la emoción, el presentador abre otra puerta, y nos muestra que está vacía. Por último, el presentador nos pregunta si queremos permanecer con la puerta elegida en un principio o deseamos cambiarnos a la tercera puerta que aún permanece cerrada.



¿Cuál es la mejor estrategia, permanecer o cambiar?

Teniendo en cuenta que inevitablemente nuestro razonamiento lleva incorporado un tinte social, que incluso puede verse como una especie de defecto de fábrica, podemos presentar una primera clasificación de cómo resolvemos problemas los humanos:

- **Fe.** Suponer que elegir cierta acción nos va a traer más beneficios, pero realmente no tenemos ninguna prueba, certeza, indicador ni experiencia de ello. Esto es equivalente a una variable estocástica. No es recomendable actuar por fe ni individualmente ni colectivamente.
- **Intuición.** Tenemos pruebas, experiencia e indicadores de nuestras ideas, pero no las hemos sabido formalizar. Hay bastante probabilidad de éxito y es recomendable seguir intuiciones a nivel individual, pero no podemos comunicar de manera efectiva este tipo de razonamientos a un colectivo de personas.
- **Razonamiento estructurado.** Hemos reflexionado sobre nuestras intuiciones y hemos logrado verbalizarlas y modelarlas en un proceso lógico inductivo, deductivo u otro. Esta es la forma de inteligencia más fuerte, en el sentido de que podemos someter a pruebas el modelo. Además, estos modelos son comunicables al colectivo de personas, por ejemplo, por medio de la universidad.

Pongamos ejemplos de los tres casos:

- A Juan le gusta el fútbol, pero a su equipo del alma le está yendo muy mal. Han perdido los 10 últimos juegos por su pésimo estilo, el presidente está en la cárcel por una cuestión de sobornos a árbitros, la hinchada no asiste ya al estadio, tienen muchas deudas, el entrenador acaba de despedirse porque no le están pagando, y hay varios jugadores lesionados. A pesar de ello, Juan tiene fe en que va a ganar el partido de esta noche. Eso es la fe. Irracional. Todos los indicadores en contra, pero Juan tiene fe. La fe a veces se basa en la tradición, es decir, seguir en lo que siempre se ha hecho o se ha visto hacer a los seres cercanos. A veces se confunde con el deseo. Pero definitivamente es un comportamiento humano. Es irracional pero vemos cómo a diario los humanos lo usamos para tomar decisiones. A veces incluso graves decisiones. De modo que hay que considerarlo como parte de la inteligencia humana. Una forma de modelarlo es como una variable estocástica, que toma un primer valor al azar, y ya no lo cambia jamás. Probablemente nadie quiera que la fe se incluya dentro de un *software* de inteligencia artificial, pero ese es otro asunto.
- Luisa acaba de despertar de un letargo de 50 años en que estuvo bajo criogenia. Y mira sorprendida un televisor bastante extraño, pero

televisor, donde va a comenzar un partido internacional entre Estados Unidos y Albania, de un juego llamado *calamunca*, completamente nuevo para ella, pero que causa furor por estos días. Ella, en los breves minutos que lleva despierta, intuye que va a ganar Estados Unidos. Y acierta. Más adelante, si le dan suficiente tiempo para reflexionar sobre ello, Luisa nos dirá que las reglas del juego le son desconocidas y que le costó un buen rato medio entender de qué se trataba, pues ni siquiera había una pelota rodando. Pudo identificar dos equipos por el color de los uniformes, y un montón de zonas en el suelo, que también tenían colores diversos. Y por los gritos de los aficionados sabía que estaba ocurriendo una jugada importante. Pero los movimientos de los jugadores eran completamente incomprensibles para ella. Sin embargo, sabía qué países eran Estados Unidos y Albania hace 50 años. El primero grande y rico, el segundo pequeño y pobre. Esas cosas no suelen cambiar rápidamente. Y en un país grande y rico hay más posibilidades de invertir en deporte, y hay más gente que haga deporte, y por ello hay más probabilidad de encontrar mejores deportistas. Además, en la mayoría de deportes que ella conoce, el primer país ganaba con holgura al segundo hace 50 años. Entonces basó su respuesta en esta información. Luisa está verbalizando ahora todo esto, pero en el momento de hacer la predicción, simplemente la hizo, sin razonarla detalladamente como ahora. Eso es lo que llamamos intuición. Tenemos un montón de conocimientos almacenados que usamos cuando las circunstancias no nos dan tiempo a hacer un razonamiento detallado. Y en muchos casos, acertamos. La intuición sí sirve, especialmente cuando hay que tomar una decisión bajo la presión de un tiempo escaso. Obviamente unos humanos tienen mejor intuición que otros. La intuición forma parte de la inteligencia humana. Y sería muy deseable que la inteligencia artificial también usara la intuición cuando opere con restricciones de tiempo. Pero no es comunicable, no se puede enseñar en las universidades, depende de un montón de experiencias personales.

- Luisa lleva ya cinco años en su nueva realidad, y también se aficionó al *calamunca*. En su país hay muchos equipos y también hay un sistema de apuestas con el que ella está ganando bastante dinero. Nos cuenta que poco a poco fue extrayendo un modelo de cada jugador, de todos los equipos. El modelo incluye la velocidad, los reflejos, la coordinación con otros y la visión lateral, que son los factores relevantes del juego. Una vez que sabe las alineaciones de cada equipo, corre muchas veces el modelo añadiendo ruido en todos los parámetros para verificar si emerge una solución robusta. La apuesta la hace sobre

los equipos ganadores de su modelo, y el valor de la apuesta depende de la dispersión de los resultados. Crear un modelo de una situación real, ajustar correctamente sus parámetros y tomar decisiones sobre el mundo real basadas en los resultados de la ejecución del modelo constituyen un ejercicio de inteligencia de un nivel más alto. No es la única forma de inteligencia, pues también se pueden tomar decisiones basadas en métodos deductivos, inductivos y otros.

Hemos visto tres niveles de inteligencia. ¿Puede haber más? En principio, nada lo impide, pero no tenemos más pistas al respecto. Probablemente surjan como combinatoria de los métodos ya enumerados. Quizás alguien descubra métodos nuevos. Y a partir de aquí solo podemos elaborar más el tercero, el razonamiento estructurado, que a su vez tiene varias formas:

- **Deducción.** Es el método más exacto, donde se parte de verdades generales para demostrar verdades particulares. El ejemplo más obvio es el *modus ponens*, que podemos ilustrar con un ejemplo: *si este jarrón se cae, se romperá. El jarrón se ha caído. Por tanto, el jarrón se ha roto.*
- **Inducción.** Aquí se hace al revés. Se parte de verdades particulares y se llega a una verdad general. Es un método exacto en matemáticas, y es convincente, pero no exacto, en las ciencias. Para hacernos una idea de por qué jamás la inducción será exacta⁸⁰ en ciencias, nada mejor que un chiste: *un matemático, un físico y un ingeniero viajan juntos por primera vez por Australia. Van en un autobús y divisan un cisne de por allí. El ingeniero dice: ¡mirad, en Australia todos los cisnes son negros! El físico le replica: no. Lo que sabemos es que en Australia hay un cisne negro. Y el matemático sentencia: no, no, no. Lo único que sabemos es que en Australia hay un cisne que es negro por un lado.* Es decir, nunca hay seguridad de que hayas logrado capturar la esencia de un fenómeno, y que el sol haya salido mil veces seguidas por el este no te da garantías de que mañana vuelva a hacerlo. También se le llama a ello el razonamiento (erróneo) del pavo que, el día antes de navidad piensa: *estos humanos son muy amables pues han estado alimentándome y cuidándome durante muchos meses. Nada malo me ocurrirá con ellos.*
- **Reducción al absurdo.** Se supone la hipótesis falsa y se razona con ella hasta llegar a una contradicción, lo que significa que la hipótesis no puede ser falsa. Un ejemplo en matemáticas sería la demostración de Euclides que dice que los primos son infinitos: primero suponemos que los primos son finitos $\{p_1, p_2, \dots, p_N\}$. Los multiplicamos todos y le

80 El filósofo David Hume alertaba de ello en el siglo XVIII.

sumamos 1 al resultado, obteniéndose un número $q=p_1 \cdot p_2 \cdot \dots \cdot p_N + 1$. Este nuevo número q es mayor que todos los demás y no puede ser primo debido a la suposición inicial. Por tanto es no-primo, y eso significa que existe un número primo p_x que es divisor de q . Desde luego, ese primo p_x es divisor de $p_1 \cdot p_2 \cdot \dots \cdot p_N$ y dado que $q=p_1 \cdot p_2 \cdot \dots \cdot p_N + 1$ entonces también se requiere que p_x sea divisor de 1 , lo cual es imposible. Por tanto, la suposición inicial es falsa, y los primos son infinitos.

- **Diagonalización.** Sirve para hacer demostraciones con conjuntos de infinitos elementos, típicamente para demostrar que la cardinalidad de dos conjuntos es distinta, aunque en ambos sea infinita. Ello creó la matemática de los transfinitos, que sirvió para demostrar que hay más números reales que naturales. Y también es el núcleo de la demostración de Gödel y del problema de la parada de Turing.
- **Mutación.** Es decir, prueba y error, que es lo más básico. Con este método no se pueden hacer razonamientos, pero sirve para descubrir verdades probables (que después habrá que demostrar con otro método).
- **Amplificación del absurdo.** Es, básicamente, el método *socrático*: hablar con una persona de un tema sobre el que mantiene una opinión absurda, animarla a que lo desarrolle, hacerle preguntas, para finalmente que ella misma se dé cuenta que llega a una contradicción. Es especialmente interesante y fácil de practicar con niños. Un computador no necesitaría de este método, pues al llegar a una contradicción la detectaría rápidamente. Pero los humanos somos propensos a vivir mucho tiempo con disonancias cognitivas sin darnos cuenta. El propósito de este método es entonces amplificarlas hasta que sea insostenible mantenerlas y fuercen a un cambio de ideas.
- **Aumentar el número de dimensiones.** Al aparecer una contradicción por alguno de los métodos anteriores, se debe reformular todo. Para ello hay que ser creativos (usando mutaciones y analogías) buscando enmarcar el problema en un sistema de hipótesis más amplio, con más dimensiones (grados de libertad), que evite la contradicción. Este es el método típico de la física fundamental para crear nuevas teorías. Viene sugerido por la metáfora de que hay nudos hechos con cuerdas vulgares que no pueden desatarse en nuestro mundo de tres dimensiones, pero si pudiéramos sumergirlos en un mundo 4D entonces sí se podrían desatar. Un ejemplo real es que como la física de Newton entra en contradicción cuando se encuentra que la velocidad de la luz es constante aun cuando el observador esté viajando a cualquier velocidad, entonces se relaja la suposición de que el espacio y tiempo son sistemas de referencia fijos y se les permite estirarse y encogerse,

lenguaje (típicamente una máquina de estados finitos) con la que realizar ya la búsqueda de la solución.

- **Analogía, metáfora o isomorfismo.** Es el más potente hasta ahora, y se usa casi sin darnos cuenta. Consiste en crear un paralelismo entre dos áreas distintas de conocimiento, de modo que los razonamientos que se hagan en una, se pueden trasladar a conclusiones en la otra. Uno puede pensar que la metáfora es una figura retórica confinada al ámbito literario. Nada más lejos de la realidad. La ciencia está plagada de metáforas e incluso el immaculado edificio de las matemáticas las usa sin reparo. Por ejemplo, Kolmogorov (1973) emplea un lenguaje muy bonito, muy matemático, conciso, limpio, a la vez que, con frecuencia usa expresiones aparentemente poco rigurosas: “surge de modo natural este nuevo concepto”, “es muy natural considerar este nuevo concepto así”, “análogamente”, y expresiones similares. Hay que afrontarlo. La metáfora es una forma de razonamiento muy imbricada en el pensamiento humano y la única que nos permite dar el salto entre el mundo de las ideas y el mundo real. Cuando pasamos de sumar dos manzanas con tres manzanas, a sumar $2+3$, estamos haciendo uno de estos saltos. Hofstadter (1995) es el principal investigador que hace énfasis en que realmente este es el método que más usamos los humanos. En la figura 87 tenemos un ejemplo de muchos *puzzles* que realmente son el mismo, si uno sabe buscar el isomorfismo adecuado.



Figura 87. Puzzles isomorfos

- **Búsqueda de los límites.** Cuando no se tiene la menor idea de cómo solucionar un problema, es una buena estrategia formular los aspectos teóricos que impiden encontrar la solución. Por eso es que hay un capítulo en el libro anterior que explica cuáles son los límites fundamentales de la computación, que nos impiden llegar a la verdadera inteligencia artificial. Una vez que se conocen los límites se puede intentar ver en qué circunstancias no son aplicables. O se puede intentar formular el problema en otros términos, donde desaparezcan las premisas que soportan los límites. En el capítulo siguiente sobre la consciencia se verá un ejemplo de ello: una vez que aceptamos que ningún proceso físico, químico o biológico tiene la capacidad causal para generar consciencia, ello significa que hay que buscar otro tipo de procesos, concretamente, computacionales.

El razonamiento estructurado humano está parcialmente cubierto por las técnicas de *Machine Learning* que explicamos en el apartado anterior. Pero no completamente, y aquí queda todavía mucho trabajo por hacer, explorando las técnicas recién enumeradas. Y quizás haya más.

Además, esta es una clasificación algo engañosa. Los humanos solo razonamos así cuando nos enfrentamos a un examen o algún problema formal o en ámbitos académicos como la escritura de un artículo o dar clases. En la vida diaria nadie usa el método deductivo para saber si puede cruzar sin peligro una calle, ni usa la reducción al absurdo para elegir pareja. Está bien documentado que los humanos tomamos decisiones basados en las emociones (Ariely, 2008) y luego racionalizamos el proceso. Y que cuando en teoría de juegos se habla de jugadores racionales, no entran allí los humanos que son capaces de hacerse daño a sí mismos con tal de llevar a cabo una venganza, porque eso no es racional. Entonces, todos estos métodos tienen que ver con un ideal de racionalidad que perseguimos para la inteligencia artificial.

El filósofo David Hume (Quinton, 1999) proponía que la razón es, y debe ser, esclava de las pasiones, como la prudencia, el amor o la envidia. La razón no puede tomar decisiones por sí misma, sino que cada persona tiene pasiones y ello le lleva a desear hacer cosas. De este modo, la razón le ayuda a encontrar la mejor forma de hacerlas. Esto ha dado lugar a una nueva área de investigación, los ordenadores emocionales, uno de cuyos primeros autores fue Picard (1998).

INTELIGENCIA COLECTIVA

Toda inteligencia es colectiva,
ÁNGEL E. GARCÍA BAÑOS

Últimamente han aparecido varios libros muy interesantes sobre inteligencia colectiva, con ejemplos de cómo un conjunto de agentes (personas, hormigas, bacterias, hongo del fango en laberintos...) pueden resolver problemas que ninguno de ellos por separado podría.

En cuanto reflexionamos sobre ello resulta que toda inteligencia es colectiva, pues cuando decimos que el ser humano es inteligente, habitualmente lo que queremos resaltar son los logros de toda la humanidad, que se basan principalmente en un sistema educativo de replicación de conocimientos. Si yo puedo desarrollar un *software* interesante es porque millones de personas antes que yo han trabajado, ideado y proporcionado un conjunto de teorías y de herramientas físicas (incluyendo los computadores) que me permitieron dar mi pequeño pasito hacia adelante.

Pero si nos centramos en un solo ser humano realizando una tarea para la que se requiere inteligencia, debemos percatarnos que quien realmente está pensando es un colectivo de neuronas.

Y probablemente por ello es que un *software* monolítico diseñado con una intencionalidad concreta no nos creemos que pueda llegar a ser inteligente pues no supera la objeción de Ada Lovelace: ese *software* hace ciegamente lo que el programador le pidió que hiciera, por lo que la inteligencia reside en el programador, no en el *software*.

La cosa cambia si diseñamos un *software* formado por un colectivo de programas, de cuya interacción surgen el pensamiento o las acciones que vamos a ver que expresa hacia el mundo. Ese *software* sí nos puede sorprender y quizás podamos atribuirle algún tipo de inteligencia. La sorpresa (o el no-determinismo) viene de la interacción no coordinada entre muchos algoritmos.

Al respecto los algoritmos evolutivos son una buena opción porque están constituidos por una población de entes (estructuras de datos o programas) que compiten entre sí y se reproducen. Solo uno de los entes se expresará hacia el mundo, solo uno de ellos propondrá su solución al problema que le planteemos. Pero no es posible predecir cual, especialmente porque cada programa habrá sufrido modificaciones aleatorias por interacción con otros programas (mutación y cruce), de modo que ni siquiera el programador podrá saber de antemano cuál va a ser la solución que proponga el algoritmo evolutivo.

Y esto tiene un claro paralelismo con la forma como funcionamos los humanos. Ante un problema, es común que debatamos internamente entre varias soluciones posibles. Es posible que ideemos otras nuevas (mutaciones) o combinemos varias (cruce). Finalmente una de esas ideas internas la llevaremos a cabo, y de lo malo o bueno de su resultado nos quedará una experiencia que modificará la forma como realicemos este proceso en un futuro. Incluso es posible que hagamos algo un día y cambiemos de idea al día siguiente, lo que quiere decir que la población de ideas no desaparece incluso aunque haya una vencedora, sino que continúan en una lucha interna por salir a expresarse al exterior.

En el libro *The wisdom of crowds* (2014), James Surowiecki nos cuenta cuándo comenzó a pensarse en inteligencias colectivas. Parece ser que el estadístico Galton estaba en un mercado de ganado y presencié un concurso: había un buey y la gente debía tratar de adivinar cuanto pesaba, rellenando un papelito con sus datos. Al finalizar el concurso pesaron el buey, miraron todos los papelitos y dieron un premio al que más se acercó en su pronóstico. Galton pidió los papelitos y sacó el promedio de todos ellos. Este resultado se acercó mucho más al peso real del buey. Con este ejemplo se muestra que el colectivo sabe más que el mayor de los expertos.

¿O no?

Porque la verdad es que eso parece prácticamente imposible. La explicación que ofrece Surowiecki es que para lograr este resultado se necesitan cuatro cosas:

- Que no haya un control central tomando decisiones.
- Que el colectivo de personas tenga mucha diversidad.
- Que las personas no se comuniquen entre ellas.
- Que haya una forma de agregar el conocimiento de todas las personas.

La falta de control centralizado es un tema recurrente en vida artificial ya que, principalmente, ofrece robustez ante fallos (si el presidente se enferma no pasa nada, porque no hay presidente).

La bondad de la diversidad también es un tema recurrente en vida artificial. Si el colectivo es homogéneo entonces los resultados son malos porque todas las personas pueden estar sesgadas en la misma dirección. Pero si el colectivo incluye personas con y sin conocimientos, con y sin experiencia y de diversas procedencias, entonces es de esperar que las opiniones particulares conformen una campana de Gauss, donde los sesgos hacia arriba se compensen con los sesgos hacia abajo, y que el resultado promedio sea muy acertado. Para ello es importante que no haya comunicación dentro del colectivo, pues de otro modo aparecerían correlaciones que impedirían la

formación de la campana de Gauss (típicamente aparecerán leyes de potencia), por ejemplo, si mucha gente decide imitar al que aparentemente más sabe, o al más charlatán.

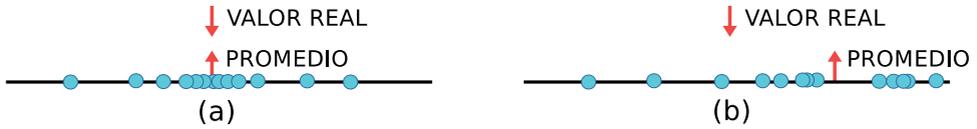


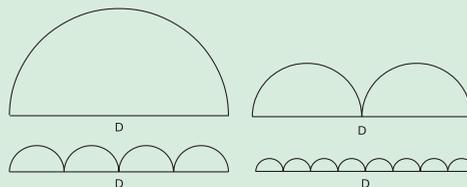
Figura 88. Dos posibilidades en el experimento de Galton

Suena bien pero, en mi opinión, Galton simplemente tuvo suerte. No hay ninguna garantía de que la campana de Gauss esté centrada en el valor correcto (como se ve en la figura 88-a). Al respecto ver también el problema 20. No hay ningún indicador que sirva para evaluar la calidad del colectivo, *a priori*. No se pueden esperar milagros de la inteligencia colectiva, y menos si se ejecutan experimentos de esta manera, suponiendo que el promedio lo va a resolver todo. Y, como más adelante reconoce el autor del mencionado libro, el experimento fracasa si se usan solo niños (Figura 88-b). O sea que la diversidad no es garantía de éxito. Hacen falta expertos. Los expertos tendrán información imperfecta y sesgos que, si no hay comunicación, tenderán a anularse. Pero todos ellos tienen conocimiento y quieren llegar a averiguar el valor real, por lo que no ocurrirá lo que vemos en la figura 88-b.

Problema 20: Aplanando la circunferencia



Dibujemos un segmento de longitud D y usémoslo como diámetro de una semicircunferencia. En su interior dibujemos dos semicircunferencias de la mitad de diámetro cada una. Repitamos infinitas veces el proceso. Al final, las semicircunferencias se confundirán con el segmento D dibujado inicialmente.



La longitud de la primera semicircunferencia es $\pi D/2$. La suma de las dos siguientes semicircunferencias es $2\pi D/4/2$. La de las 4 siguientes es $4\pi D/4/2$, y así sucesivamente. Cuando tengamos N semicircunferencias, su longitud total será $N\pi D/N/2 = \pi D/2$. Esta longitud es igual a la del segmento D , dibujado inicialmente. Por tanto $\pi D/2 = D$ de donde se deduce el verdadero valor del número π : $\pi = 2$.

¿Está usted de acuerdo?

De alguna forma, en los algoritmos evolutivos también ocurre algo así. Hay una presión selectiva empujando a que los individuos se aproximen a la mejor solución. Si se escogieran individuos al azar y se hicieran promedios no habría garantías de acercarse al óptimo. Entonces, como en toda nueva ciencia, hay confusiones iniciales. Los expertos hacen falta, pero se requiere que haya una variedad de ellos. Los ignorantes solo introducen ruido. Y por ello es que en los casos de éxito los colectivos usan sus conocimientos de una manera estructurada para lograr resultados que ninguno de sus miembros a nivel individual podría obtener. “Estructurada” significa que se usan expertos en múltiples áreas, que hay comunicación entre ellos, y que los expertos son más relevantes en la toma de decisiones que los legos. La comunicación no suele ser jerárquica como en las empresas tradicionales, sino que se permite un flujo libre de ideas y de acciones. Y en casos más complejos se requieren equipos de expertos específicamente seleccionados para atender situaciones de emergencia, como cuando en una fábrica de aviones tienen problemas completamente nuevos (Miller, 2010) o cuando quisieron rescatar a un deportista de *kitesurf* que se perdió en el Mar Rojo (Nielek, 2013) sin saber nadie su posición, sino únicamente con expertos en diversas áreas como *kitesurf*, corrientes marinas y geografía local. Aquí también muestran cómo el ruido introducido por los no expertos puede amplificarse por medio de realimentaciones positivas. Concretamente, la gente miraba las noticias en la televisión y publicaba su opinión sobre ellas en los foros de Internet. A su vez los periodistas leían los foros y daban su versión por televisión. De este modo, cualquier rumor se amplifica exponencialmente, dificultando la entrada de informaciones reales.

Pero ¿qué tipo de comunicación se requiere entre expertos que evite el sesgo de autoridad⁸¹? Esta es una pregunta fundamental, pues los agentes que conforman el colectivo tienen una inteligencia muy limitada, de modo que ella emerge gracias a la red. Pensemos de nuevo en el cerebro humano, ya que también es un colectivo: la capacidad de cómputo de una neurona es bajísima, a la vez que la arquitectura de conexiones de toda la red es tan compleja que no puede estar codificada en el ADN. De modo que se puede sospechar que hay un algoritmo parcialmente aleatorio que construye las conexiones. Su parte determinista es muy sencilla de especificar con poca información (que estará en algún lugar del ADN), y su parte aleatoria hace que cada cerebro sea distinto de otro pero aun así, todos funcionan razonablemente bien.

81 Cuando todo el mundo imita al que tiene o parece tener más conocimientos en el tema.

Averiguar cuál es ese algoritmo de generación de la red significará entender por completo el funcionamiento del cerebro. Seguramente falta poco para lograrlo, ya que después de la “década del cerebro” (1990-2000) sigue habiendo muchos otros proyectos centrados en entender el cerebro humano e incluso simularlo computacionalmente. Se puede mencionar la iniciativa BRAIN de USA, el HBP de Europa, el BlueBrain de Suiza, el Brain/MINDS de Japón, el BRAIN de China. Pero también hay iniciativas de empresas particulares y actualmente es posible acceder a APIs de simulaciones de cerebro disponibles en Internet, como el TensorFlow de Google con el que se ha derrotado al campeón mundial humano de Go; el Brain Simulator de GoodAI, corriendo sobre tarjetas gráficas con interfaz CUDA; NuPic de Numenta que simula las columnas de neuronas en el córtex; Microsoft, IBM, Yahoo, junto a muchos otros, tienen sus propios proyectos de inteligencia artificial, aunque más orientados al *Big Data* que a la simulación del cerebro, pero son caminos también muy válidos.

De todos modos lograr entender y fabricar inteligencia colectiva no es sencillo. Eric Bonabeau, un experto en el tema predijo en el 2009 que Wikipedia iba a fracasar, y hoy todavía no entiende el éxito que sigue teniendo. Y de la misma forma que Galton pensaba que el colectivo es más inteligente que cada individuo, Miller (2010) cree que “ninguno de nosotros es tan tonto como todos nosotros juntos”. Es decir, a veces la inteligencia colectiva fracasa rotundamente y pone ejemplos de ello:

- Los hormigueros la mayoría de las veces hacen un trabajo extraordinario que llamaríamos inteligente: construyen sus nidos, buscan comida y se defienden de sus enemigos, a pesar de que la inteligencia de cada hormiga es minúscula y muchas de ellas parecen empujar la comida en el sentido equivocado e incluso olvidar para donde van. Emerge un comportamiento inteligente a pesar de la estupidez de cada una de ellas. Sin embargo, a veces también emerge estupidez, como cuando hay una larga columna de hormigas caminando y dejando un rastro de feromonas. En su deambular llegan a cruzarse por algún sitio donde ya habían pasado, detectando sus propias feromonas y siguiendo el rastro. El resultado es la llamada “marcha de la muerte”, pues darán vueltas en círculos sin descanso, hasta fenecer.
- La langosta caníbal, de la que hablábamos en el capítulo “Complejidad”, es otro ejemplo de comportamiento colectivo emergente que lleva a un desastre.
- Los comportamientos colectivos humanos no están exentos de ese problema. Las burbujas económicas ocurren cuando todos los individuos quieren ganar dinero, con lo que producen un crecimiento ex-

ponencial del valor de algún bien, típicamente los bienes raíces. Y lo aceptan a sabiendas de que conducirá a una excesiva sobrevaluación del bien, hasta que llegue un momento en que nadie quiera comprarlo, cayendo en picado su precio. Y entonces se dice que la burbuja ha explotado. Todos quieren ganar, pero el comportamiento colectivo hace que todos pierdan.

Hay casos donde no es razonable dejar en manos de un reducido número de expertos la solución a un problema difícil que afecta a un colectivo mayor, y entonces se recurre a hacer votaciones democráticas, donde a cada persona le corresponde un voto. Así son los sistemas de elección de dirigentes en los países democráticos y, como muchas veces vemos, eso no significa que siempre den con una solución correcta. Además, se ha detectado que los resultados de las votaciones democráticas son predecibles con bastante exactitud cuando un grupo de expertos usa la primera técnica que vamos a ver enseguida. De modo que quizás haya que buscar una transición a un sistema democrático más sofisticado⁸². Porque hay varias técnicas para diseñar inteligencia colectiva. Las que se conocen hasta ahora son:

- **Mercados de predicción.** Cada experto hace una predicción que se transforma en una apuesta en dinero y que, al final, se convierte en un premio para el que se acerque más. Con ello se logran evitar sesgos emocionales (cuando hay dinero de por medio, nadie va a predecir que va a ganar su equipo de fútbol favorito si las probabilidades son bajas). Y quien no tiene información fiable, no apuesta nada. De modo que la información privada fiable que alguien pueda tener sale a la luz, porque se le da un incentivo económico para ello.
- **Método Delphi.** Se hace una encuesta a expertos sobre lo que piensan que va a ocurrir en un determinado tema. Y se les pide argumentos que sustenten sus opiniones. Después se publican los resultados así como los argumentos, pero de forma anónima, es decir, sin que se sepa lo que piensa cada uno de los expertos. Los expertos pueden cambiar de opinión cuando ven los resultados de la encuesta anterior así como los argumentos que apoyan cada predicción, de modo que se repite la encuesta y su publicación varias veces, hasta que se llegue a un consenso. Con ello se permite que los expertos se comuniquen,

82 Isaac Asimov anticipaba un ejemplo caricaturesco de ello en su cuento “Sufragio universal” (Asimov, 1991), que relata en un futuro distante cómo un gran computador selecciona a una única persona en todo USA que representa las características del ciudadano promedio, para que sea el único votante en las elecciones.

pero sin que medie entre ellos ninguna emoción ni ninguna presión por el respeto a la autoridad⁸³. Esto produce una realimentación positiva de las buenas ideas.

Este método es bastante intuitivo y se ha inventado varias veces. Por ejemplo, Ray Dalio (2017) lo ideó para tomar decisiones en su empresa operadora del mercado de capitales, que es la mayor de su clase en el mundo, aunque con una gran diferencia: en vez de opiniones anónimas, inculca en sus empleados la sinceridad y la transparencia, que son valores emocionalmente más difíciles de conseguir. La verdad es que es más sencillo lograr lo mismo por medio del anonimato. Lo que sí ha conseguido es ponderar las opiniones de cada funcionario en función de su credibilidad, es decir, cómo le haya ido en situaciones pasadas.

- **Votaciones.** Muchas veces se recurre a las votaciones para establecer decisiones sobre temas difíciles. Puede que entre el grupo haya expertos, pero sus votos se van a diluir entre los de la gente común, y eso va a llevar a malas decisiones. Pero hay una forma sorprendentemente fácil de separar los expertos de quienes no lo son (Vedantam, 2017), haciendo a cada persona una doble pregunta de tipo ¿qué es X? y ¿qué crees que la gente piensa que es X? Los expertos pueden diferenciar las respuestas, mientras que el común de la gente contesta lo mismo a ambas. Se mostrará con un par de ejemplos:
 - ¿Cuál es la capital de Brasil? y ¿cuál crees que piensa la mayoría de la gente que es la capital de Brasil? Si en la primera pregunta la gente contesta Brasilia y Río de Janeiro, mientras que en la segunda contesta solo Río de Janeiro, entonces la respuesta correcta es Brasilia. Y quien sabe diferenciarlos es el experto.
 - En un estanque crecen nenúfares, duplicando el área ocupada cada día. Si ocupan todo el estanque en 48 días. ¿En cuántos días ocuparon la mitad del estanque? y ¿cuántos días crees que piensa el resto de la gente que los nenúfares ocuparon la mitad del estanque? Si en la primera pregunta la gente contesta 24 y 47, mientras que en la segunda contesta solo 24, entonces la respuesta correcta es 47. Y el experto es quien sabe diferenciar las dos respuestas.
- Obviamente hay que hacer algún tipo de estadística pues es determinante que no haya nadie (o casi nadie) que responda Río de Janeiro a

83 Todos nos dejamos influir con facilidad. Los experimentos de Asch (1974) muestran personas que cambian su declaración de lo que han observado en una escena debido a que todos los demás testigos decían lo contrario (eran actores con los que se había convenido que dieran falsas informaciones).

la primera pregunta y Brasilia a la segunda. De la misma forma, cuanto más gente responde 47 a la primera pregunta que gente que responde 47 a la segunda pregunta, aumenta la probabilidad de que la respuesta correcta sea 47. Y hay más formas de distinguir sistemáticamente a los expertos, como puede verse en Pal, Harper y Konstan (2012).

Se ha diseñado *software web* de apoyo a estas metodologías, para no tener que reunir a todo el grupo de personas en el mismo sitio a la misma hora.

Y se puede considerar una forma muy general, que estoy seguro de que te va a sorprender si estás al día en las últimas tendencias en programación. Se necesitan dos fases: primero, distribuir la información entre todos los agentes que conforman la inteligencia colectiva. Y después, agrupar los resultados que cada agente ha generado. ¡Efectivamente!, es la función *collect()* de Ruby, que primero se hizo popular como el algoritmo *map-reduce* (MapReduce_Wiki, 2017) para el cálculo del *PageRank* de Google. La distribución de datos (el *map*) puede hacerse de diversas maneras, pero el objetivo es la eficiencia, por lo que la más efectiva es el árbol. Y la forma de reducir depende completamente del resultado que se desea obtener. Con este tipo de algoritmos se pueden calcular resultados de votaciones, promedios o valores de mercado que son función de la oferta y la demanda, por mencionar unos pocos. Y, por tanto, el *map-reduce* también sirve para conseguir inteligencia colectiva.

Otra pregunta válida es si podremos darnos cuenta de que existe inteligencia en el nivel del superorganismo, por encima del nuestro individual. Para entender lo difícil que es percatarse de ello, hagamos un paralelismo con una célula de nuestro cuerpo a ver si se da cuenta de que pertenece a un organismo más complejo. Terriblemente difícil, pero nosotros contamos con algo más de inteligencia que la pobre célula. Y así propongo dos indicadores de que exista una inteligencia en el nivel superior:

- Si es más fácil conseguir objetivos en grupo que solo, ello demuestra que hay sinergia en el grupo, y es un indicador débil de inteligencia en el superorganismo.
- Vamos a ir un poco más allá: si los problemas del individuo se resuelven como por arte de magia, sin necesidad de realizar ninguna acción, ello es un indicador fuerte de que el individuo se encuentra sumergido en un superorganismo inteligente, que puede predecir el futuro y anticipar las soluciones que requieran sus individuos constituyentes. Y ello ya empieza a ocurrir en nuestro mundo occidental gracias a la interacción de muchos humanos con las tecnologías de la información y las comunicaciones.

Pensemos cómo puede surgir inteligencia colectiva en una universidad. Allí hay un conjunto de profesores $\{p_1, p_2, p_3, \dots\}$, un conjunto de escuelas formadas por profesores $\{e_1, e_2, e_3, \dots\}$ y luego un enorme conjunto de comités de escuela, de programa, de planta física, de laboratorios y de investigaciones $\{c_1, c_2, c_3, \dots\}$. También se construyen comités transversales con representantes de otros comités en una jerarquía de niveles que da vértigo $\{t_1, t_2, t_3, \dots\}$. Si recordamos el primer capítulo acerca de las formas de generar complejidad, aquí vemos la más básica: combinatoria. Algunos comités generan normas para regular la acción de los profesores y de los otros comités; otros piden o emiten informes; otros solicitan u ofrecen recursos; y todo ello debe interpretarse como bucles de realimentación, con su consiguiente generación de complejidad. Como cada bucle funciona en paralelo con los otros, el resultado global es emergente, en el peor sentido de la palabra: una sorpresa continua. Porque aparecen realimentaciones inesperadas: dado que un profesor puede estar en varios comités, la información puede fluir de una manera misteriosa entre ambos o puede haber acciones de un comité que favorezcan a otro. Esto en definitiva es el comienzo de un comportamiento colectivo inteligente pero también puede dar lugar a injusticias y presiones indebidas.

La inteligencia colectiva no es fácil de lograr, y lo poco que se consigue es con mucho esfuerzo. Habitualmente los dirigentes de un grupo no saben qué hacer para cumplir objetivos, debido a la complejidad de los problemas. Suelen tener deseos, típicamente mejorar la capacidad de supervivencia del superorganismo, logrando que tenga más beneficios monetarios o expandiéndose a nuevas sucursales, pero no saben cuál es el mejor plan para llevarlos a cabo. Los dirigentes dan órdenes hacia abajo expresando sus deseos, esperando que los de abajo sepan cómo cumplirlos. Los de abajo repiten lo mismo hacia el siguiente nivel inferior. Habitualmente se acaba con operarios cumpliendo órdenes descoordinadas y posiblemente contradictorias. Pero algo harán. Las contradicciones se resuelven iterando, o sea, repitiendo una y otra vez el proceso hasta que salga algo positivo, y con ayuda de comunicación local, es decir, sin seguir el conducto jerárquico regular, sino en conversaciones de pasillo con amigos que están en otras unidades con las que se interactúa de alguna manera; por cierto, es aquí donde hay algo de inteligencia a nivel individual. Se pierden muchos recursos y tiempo en ello, pero el dirigente no se da cuenta, no le importa o incluso cree que todo ha ido bien.

A veces alguien entrevé un problema bajo una perspectiva global, e incluso propone una solución. Por ejemplo, un formato mejorado para la contratación de nuevos profesores. El formato es aceptado porque es intuitivo, fácil y rápido de usar. Pero si la secretaria que diseña los formatos no es la

misma que la que los utiliza, es decir, si quien tiene el control local no obtiene beneficios inmediatos, lo más probable es que ese formato desaparezca en la siguiente convocatoria, volviéndose al antiguo. Por ello, una estrategia para hacer nacer la inteligencia colectiva es cerrar bucles, es decir, que la persona que genere un problema sea también la que sienta los resultados de su acción. Si no hay bucles cerrados no habrá acciones correctoras.

En un hormiguero es lo mismo: las hormigas encuentran una miga de pan y algunas empujan en la dirección correcta y otras no. Ni siquiera son conscientes de esta discrepancia ni intentan resolverla. Habitualmente consiguen llevar la comida al nido, pero la ineficiencia es alta. Uno podría preguntarse por qué finalmente ganan las hormigas que empujan en la dirección correcta. Seguramente la respuesta está en la evolución, que hizo que dejaran de existir los hormigueros con los genes muy equivocados. La evolución, en ausencia de una presión selectiva fuerte, se conforma con resolver problemas aunque sea de forma ineficiente.

Pero la evolución no nos ayuda en los sistemas sociales. Sería increíblemente costoso y socialmente inaceptable, teniendo como objetivo el diseño de un formato óptimo de contratación, crear un conjunto de empresas o universidades con distintas estructuras, distintas reglas, distintos formatos, y poner una presión selectiva muy alta, por ejemplo aumentando el presupuesto de las que gradúen más estudiantes y disminuyéndolo a las que gradúen menos. Si se hace eso, muchas universidades desaparecerán, y las que queden serán las que tengan mejores formatos. Pero esta forma de generar formatos óptimos es increíblemente costosa. Además, la eficiencia de cualquier organización depende de un montón de otros procesos de modo que no hay garantías de que se logre el formato óptimo a no ser que se genere un número enorme de universidades.

¿Cómo lograr entonces inteligencia colectiva? A veces un individuo tiene una buena idea. Y además de probarla, debe convencer a los otros de que lo es y de los beneficios que traerá. No solo tiene que ser creativo sino también comunicativo. Ese tipo de cosas van mejorando el sistema, pero son muy lentas.

Las piezas que forman parte de un superorganismo deberían ser muy flexibles. O sea, lo que importa no es que la pieza pueda resolver un problema, sino que sea lo suficiente flexible como para resolver un conjunto amplio de problemas. Las neuronas son así, lo mismo que las hormigas y las personas. Un ejemplo no biológico sería las piezas de un *puzzle* versus las de un *Lego*. El *puzzle* resuelve un único problema, típicamente, una bonita foto de un paisaje. El *Lego* resuelve una multitud de problemas actuales y futuros. En este sentido, las personas en una organización son muy flexibles: pueden

realizar distintas tareas, se les puede cambiar de oficina y de funciones (las hormigas no son tan flexibles). Mientras que los sistemas de información son muy inflexibles, pues necesitan los datos en un cierto formato y orden, la *interface* con el usuario es rígida, los cálculos son siempre los mismos y el formato de salida de resultados suele ser solo uno (la mayoría de las veces es el peor formato posible, HTML, que no es interoperable programáticamente con otros sistemas de información). La ventaja del *Lego* radica en su capacidad combinatoria: las piezas se pueden conectar con otras piezas de muchas maneras. Análogamente, la tabla periódica de los elementos permite, por combinatoria, generar muchas moléculas. Si todos los átomos fueran como los gases nobles, sin capacidad de combinarse, la vida sería imposible. Por eso es que si se emplean sistemas de información inflexibles es imposible que nazca la inteligencia colectiva. Y sin inteligencia, el superorganismo en ciernes morirá, pues no podrá adaptarse a los nuevos problemas que aparezcan en su entorno.

El superorganismo necesita identificar problemas globales, pero los individuos que lo conforman solo pueden hacerlo de forma local. Es imposible para ellos identificar problemas globales. Lo único que pueden hacer desde su localidad es tratar de ver lo más lejos posible. Si en una universidad cada departamento optimiza localmente sus procesos, el resultado suele ser un montón de contradicciones e ineficiencias a nivel global. Cuando todos los individuos se dan cuenta de ello, se sentarán a negociar para ponerse de acuerdo en la manera de interoperar. Y habremos resuelto problemas locales un poco más amplios. Sin embargo, los problemas globales son más difíciles de percibir. Por ejemplo, suponiendo que la universidad prepare egresados con muy buena formación, ¿debe aumentar al máximo posible la salida de egresados? Localmente a nivel de la universidad pareciera que sí, pero ¿y a nivel social? Si estos egresados no son los que requiere la sociedad, lo único que estaremos haciendo es aumentar el problema del desempleo. Otra pregunta en la misma tónica: muchos egresados consiguen trabajo en el extranjero. ¿Es razonable que el país haya dedicado tantos recursos para formar personal que finalmente va a trabajar para otro país? A un nivel local de país, la respuesta es no. Pero a un nivel más amplio, planetario, la respuesta es sí. ¿Dónde está el nivel global? ¿Hay que situarse en el plano del universo completo? ¿Del multiverso?

No tengo respuesta, pero la pregunta se parece mucho a la que hace Dennett acerca del *yo*, como punto gordo. Dependiendo de lo amplio que seas de miras, el *yo* que quiero que sobreviva, evolutivamente hablando, puede ser mi cuerpo, mi familia, mi barrio, mi ciudad, mi país, mi planeta...

NUEVA PERSPECTIVA

Con todos los conceptos vistos hasta ahora, vamos a proponer una nueva forma de clasificar la inteligencia. Cuando hablábamos de complejidad, decíamos que todo es cuestión de grado y en el caso de la inteligencia también es así. Hay dos grados principales: inteligencia individual e inteligencia colectiva (Figura 89).

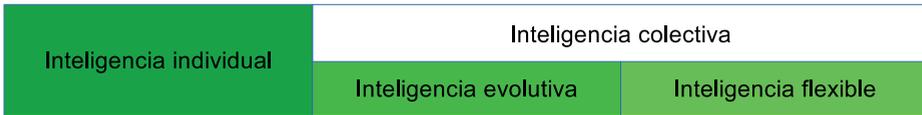


Figura 89. Inteligencia individual versus colectiva

Mucha gente no llamaría inteligente a la individual, ya que consiste en objetos únicos que no se reproducen y que tienen objetivos muy específicos. Son algo así como inteligencia solidificada. Estoy pensando en las herramientas como el destornillador, que solo sirve para apretar y soltar tornillos. El martillo para golpear clavos. Otras herramientas más abstractas que se encuadran dentro de esta categoría son el álgebra inventada por al-Jwārizmī, el algoritmo A^* inventado por Peter E. Hart, Nils J. Nilsson y Bertram Raphael en 1968 para hacer búsquedas en grafos, y el *software* de Deep Blue, especializado en jugar al ajedrez. No le pidan a Deep Blue que conduzca un automóvil por una ciudad.

Una vez que aparece la reproducción, también lo hace la inteligencia colectiva que, a su vez se puede clasificar en dos tipos: inteligencia evolutiva e inteligencia flexible (Figura 89). En la primera los individuos siguen siendo algoritmos fijos, que mueren para dar paso a otros mejores de manera que la población va aprendiendo. En este tipo, la población es más inteligente que el individuo. Es lo que ocurre con todos los algoritmos evolutivos actuales, sean biológicos o computacionales. En la segunda, los individuos se agrupan para formar superindividuos que no mueren, sino que son flexibles, se adaptan y aprenden, gracias a que llevan una población dentro de sí. Los individuos de esta población cautiva quizás si mueran en el proceso de aprendizaje, pero el superindividuo no. Es lo que ocurre con las personas y sus neuronas en el cerebro. Y la agrupación de individuos en superorganismos crea sucesivos niveles de inteligencia que no tienen por qué tener ningún techo (Figura 90).



Figura 90. Emergencia de niveles de inteligencia, por agrupación de individuos en superorganismos

De hecho, el cerebro no es un lugar apacible donde simplemente transcurren los pensamientos humanos, como uno podría crear. Todo lo contrario. Las neuronas también están envueltas en una lucha terrible para sobrevivir. Seguramente por eso es que automodifican continuamente la forma de expresar su ADN (Feng, Fouse, Fan, 2007). Las que no logran nunca ninguna correlación entre sus entradas y sus salidas mueren (o modifican sus conexiones), y eso ocurre continuamente, aunque la tasa más alta es desde los 3 hasta los 16 años de vida de un niño, lo cual se puede interpretar por medio del teorema de *No-Free-Lunch* que veíamos en el libro anterior. El cerebro neuronal hace una exploración del espacio de búsqueda que le tocó, o sea, su ambiente al nacer, y descarta las neuronas y las conexiones que no le sirven para resolver problemas. Acepta que no puede resolver todos los problemas posibles y se centra únicamente en resolver de manera eficiente los problemas que encuentra en su medio ambiente. Por eso conviene darles a los niños muchos estímulos sensoriales, afectivos y cognitivos. De lo contrario, un niño aislado en un entorno pobre desarrollará un cerebro con pocas capacidades.

Estamos viendo que en el momento en que un objeto de inteligencia individual es capaz de reproducirse, pasa a formar parte de la inteligencia colectiva evolutiva. Y en el momento en que una población de inteligencia evolutiva logra aislarse del entorno y crear una representación interna del mundo, es decir, un simulador, entonces conforma una inteligencia flexible. Pero también hay un flujo en dirección contraria: las inteligencias colectivas en su quehacer pueden crear herramientas como los destornilladores y los martillos, es decir, inteligencias individuales. De hecho, un individuo de una población evolutiva es solo eso: una herramienta, una inteligencia solidificada, capaz de resolver solamente un problema. A su vez también ocurre que las inteligencias flexibles, los cerebros humanos, se agoten a ciertas edades, y pierdan la flexibilidad. Hemos escuchado personas que dicen: “lo hago así porque siempre lo hice así”. O que no pueden entender problemas nuevos. O que se empeñan en resolver nuevos problemas con herramientas viejas. O, simplemente, que les asusta el cambio. Se convierten en algoritmos fijos, como el A*. Se vuelven inteligencias individuales, inteligencias cristalizadas, solidificadas, fósiles. Herramientas.

Además, no existe un cerebro humano sino una población, la humanidad. De modo que tenemos ya otro nivel de inteligencia colectiva evolutiva. Nada impide que sigan emergiendo nuevos niveles en una especie de espiral, que genera lateralmente a su vez muchas herramientas. Y estas herramientas le sirven como apoyo inteligente para seguir creciendo. Ahora mismo estamos creando una herramienta, la inteligencia artificial que, en cuanto adquiera capacidades de autocopia, emprenderá su propio rumbo. Otras posibles consecuencias de la capacidad de autocopia en máquinas pueden verse en (CdeCiencia, 2015).

Conforme van apareciendo más niveles por agrupación de individuos del nivel anterior, el aprendizaje en cada uno se va desligando de la muerte. Los individuos de nivel superior no tienen por qué morir para aprender, aun cuando los de niveles más elementales sigan haciéndolo. En nuestras sociedades humanas parece ser que la muerte sí tiene que ver con el aprendizaje de la sociedad. Recuerdo que cuando comencé mis estudios universitarios se jubilaron muchos profesores que ofrecían asignaturas obsoletas, y dejaron el campo abierto a una modificación curricular que modernizó los estudios. En este ejemplo estoy hablando de muerte académica, que también ocurre cuando se renueva un comité, un congreso de diputados o un presidente en algún país. El superorganismo cambia de comportamiento y aprende porque sus partes constituyentes se renuevan. La democracia, tan frágil y tan asaltada en estos tiempos, tiene esa virtud. Permite la renovación, permite el aprendizaje.

RESUMEN

Siendo un concepto difuso, se empieza a aceptar que la inteligencia es la capacidad de predecir el futuro, anticiparse a él e incluso modificarlo con el objetivo de sobrevivir, lo que indica que es un producto de la evolución. Y en ese afán de anticipación del futuro, a largo plazo la inteligencia solo produce aumento de entropía, es decir, nos acerca más rápido al final del universo.

La inteligencia, al ser una cuestión de predicción, también lo es de velocidad. Cuanto más rápidamente predigas los acontecimientos futuros y cuanto más hacia adelante en el futuro te refieras, más tiempo tendrás para reaccionar por adelantado. En este sentido, las plantas son menos inteligentes que los animales. Las plantas son bastante complejas y se han descubierto fenómenos asombrosos de adaptación, comunicación e inteligencia (Gage, 2017), pero su lentitud no las permite reaccionar a problemas inmediatos. También en este sentido, los computadores son mucho más rápidos que las personas, por lo que dejo al lector que saque sus conclusiones.

La inteligencia es un producto de la evolución. Si no hay evolución, no puede haber inteligencia. Y es un resultado colectivo. Si vemos un individuo inteligente, con seguridad está formado por múltiples unidades similares (neuronas u otros sistemas). La inteligencia se desarrolla más cuanto más complejo sea el ambiente. Si el ambiente es simple, la inteligencia que allí se desarrolle también lo será, precisamente porque es un resultado de la evolución. Es decir, al tratar de sobrevivir en un cierto ambiente, hay que predecirlo y con ello surge la inteligencia. En resumen, la complejidad que se requiere para que surja la inteligencia es la misma que se necesita para que aparezca la evolución, y su nivel de sofisticación estará limitado por la complejidad del ambiente en el que se desarrolle.

El triunfo de un *software* como campeón mundial de Go, gracias a las nuevas técnicas de *Deep Learning*, ha marcado un punto de inflexión donde la inteligencia artificial supera a los humanos en tareas muy complejas y sin que sus programadores mantengan el control de qué hace o cómo lo hace.

Hemos visto las formas de razonar de los humanos, y que nuestro cerebro está orientado a resolver problemas sociales, por razones evolutivas, aunque podemos forzarlo a resolver problemas matemáticos y de lógica, pero allí muchas veces nos equivocamos. En general los humanos actuamos orientados por nuestras emociones, a las que luego buscamos una explicación lógica más o menos hipócrita. Además, para medir la inteligencia de un ente hace falta un árbitro más inteligente, de modo que ello puede dar lugar a problemas de *aliasing*.

Hemos realizado una clasificación de problemas de inteligencia artificial según su grado de dificultad, y una de las clases más interesantes son los problemas clonables. A su vez hemos visto una clasificación de tipos de mentes, y la *popperiana* es capaz de resolver los problemas clonables, gracias a que tiene dentro un simulador. Ya habíamos visto cómo funciona un simulador en un capítulo anterior. También está el tipo de mente *gregoriana*, que posee un sistema de entrada-salida universal, como por ejemplo, la mano.

Por último reflexionamos que toda inteligencia es colectiva, es decir, una agrupación de agentes independientes y diversos, que se comunican entre sí y tratan de tomar decisiones. A lo que emerge de allí, que puede ser bueno o malo, le llamamos inteligencia. Hemos mostrado ejemplos de inteligencia colectiva así como varias formas de diseñarla y un par de indicadores de su presencia.

PARA SABER MÁS

- Steven Pinker (1997). *Cómo funciona la mente*. Bogotá: Ediciones Destino.

En la primera parte explica cómo funcionan ciertos módulos cerebrales (principalmente la visión). En la segunda parte (a partir de la página 318 aproximadamente, capítulo “Buenas ideas”), explica los comportamientos humanos desde una perspectiva evolutiva, incluyendo sus bases genéticas (preprogramadas) y sus limitaciones. Un libro fascinante, para entender quiénes somos los humanos y por qué somos así. Un libro que hay que leer, si se pretende trabajar en inteligencia artificial. Libro divulgativo y fácil de leer, a pesar de su tamaño.

- Andy Clark (1999). *Estar ahí*. Barcelona: Ediciones Paidós Ibérica S. A.

Muestra el completo cambio de perspectiva que están tomando los investigadores en inteligencia artificial. Explica funciones cerebrales desde un punto de vista evolutivo. Descarta el enfoque reduccionista *top-down* habitual en ingeniería. Resuelve el grave problema de la “representación”, que nadie sabe abordarlo en inteligencia artificial clásica, convirtiéndolo en “representación no-simbólica, orientada a la acción”. Da una clara introducción al caos y a la emergencia. Habla de los nuevos robots del MIT “rápidos, baratos y fuera de control”. La cognición adulta parece lógica y proposicional pero realmente está basada en recursos corporales de tiempo real (metáforas de fuerza, acción y movimiento). Excelente libro. Abre la mente a los investigadores de la inteligencia artificial, para que exploremos nuevos caminos.

- Daniel C. Dennett (1999). *La peligrosa idea de Darwin*. Madrid: Círculo de Lectores.

Excelente libro para entender la evolución en un sentido amplio. La evolución está presente por todas partes, no solo en biología. Muchos argumentos, algunos sorprendentes. Explica incluso cómo hacerse rico, usando la evolución de una manera inusual. Libro divulgativo y fácil de leer, a pesar de su tamaño.

- Daniel C. Dennett (2000). *Tipos de mentes: hacia una comprensión de la conciencia*. Barcelona: Debate.

Empleando un enfoque evolutivo, explica los distintos tipos de mente que pueden encontrarse en el reino animal, incluyendo los humanos, y su razón de ser.

- Eric Bonabeau, Marco Dorigo y Guy Theraulaz (1999). *Swarm intelligence*. New York: Oxford University Press.

Explica en detalle los algoritmos basados en inteligencias de enjambre, incluyendo las colonias de hormigas.

- Dan Ariely (2010). *The upside of irrationality. The unexpected benefits of defying logic at work and at home*. USA: Harper Collins.

Ariely tiene varios libros donde ilustra con ejemplos la irracionalidad humana. En este comienza contándonos cosas como que cuando se pagan incentivos para alcanzar metas los resultados suelen ser mejores cuanto mayor es el incentivo. Y ello es lo racional. Pero llega un momento donde el pago es tan alto que pone nervioso a las personas, y su desempeño disminuye notoriamente. Les ocurre a los animales, a los estudiantes, a los deportistas, a los CEO y a los banqueros. La conclusión debería ser limitar los sueldos de los banqueros para lograr que sean más responsables y eviten las crisis que estamos viendo.

Después continúa hablándonos sobre el significado del trabajo. Se consiguen más objetivos cuando el trabajo se entiende y tiene algún sentido, mientras que se abandona antes cuando no tiene sentido, a pesar de que paguen igual en ambos casos. O sea que si eres un CEO que quiere desmotivar a sus empleados, la recomendación es destruir su trabajo enfrente de ellos. Muchos animales también buscan una recompensa en la comida sobre todo si ha sido difícil de conseguir, aunque los gatos son una excepción. Los *blogs* son tan populares porque esperamos que alguien nos lea. Si supiéramos que nadie nos lee, no los escribiríamos. Dividir el trabajo en parcelas especializadas hace que sea más eficiente, pero a la vez, los trabajadores pierden el sentido de lo que hacen. Lo mismo ocurre con el *software* muy particionado y manejado por muchas personas para generar un producto.

Y así sigue con una larga colección de temas donde los humanos nos mostramos en toda nuestra irracionalidad.

REFERENCIAS

Libros, artículos y enlaces web

- Ariely, D. (2008). *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. USA: Harper Collins.
- Asch, S. (1974). Fuerzas de grupo en la modificación y distorsión de juicios. *Estudios básicos de la psicología social*, pp. 351-364. Barcelona: Hora.
- Asimov, I. (1950). *Yo, robot*. Barcelona: Editorial Edhasa.
- _____. (1991). *Con la Tierra nos basta*. Madrid: Martínez Roca.
- Aunger, R. (2002). *The electric meme*. New York: The Free Press.
- Bonabeau, E. (2009). Decisions 2.0: The Power of Collective Intelligence. *MITSloan Management Review*, 50(2), pp. 45-52.
- Carroll, L. (1995). *Matemática demente*. Barcelona: Tusquets Editores.
- Clarke, A. C. (1973). *Profiles of the Future: An Inquiry into the Limits of the Possible*. New York: Harper & Row.
- Csikszentmihályi, M. (2008). *El yo evolutivo*. Barcelona: Kairós.
- Dawkins, R. (1982). *The extended phenotype*. Oxford: Oxford University Press.
- _____. (1994). *El gen egoísta*. Barcelona: Salvat.
- Dennett, D. (1995). *La conciencia explicada: una teoría interdisciplinar*. Barcelona: Ediciones Paidós.
- Domingos, P. (2015). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic Books.
- Dubois, D. M. (ed). (1999). *Computing Anticipatory Systems (CASYS'99)*. USA: American Institute of Physics.
- Estructurado_Wiki (2017). Teorema del programa estructurado. Wikipedia. Recuperado el 21 de agosto de 2017. Disponible en: <https://goo.gl/yDptTw>
- Feng, J., Fouse, S. y Fan, G. (2007). Epigenetic regulation of neural gene expression and neuronal function. *Pediatric Research*, 61(5 Pt 2), pp. 58R-63R. DOI: <https://doi.org/10.1203/pdr.0b013e3180457635>
- Gage, G. (2017). Electrical experiments with plants that count and communicate. TED. Recuperado el 27 de noviembre de 2017. Disponible en: <https://goo.gl/AWUiz4>
- Géron, A. (2017). *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media.
- GoodAI (2015). *Brain Simulator*. Recuperado el 12 de junio de 2017. Disponible en: <https://goo.gl/Zht8GL>

- Google (2015). *Tensorflow*. Recuperado el 11 de junio de 2017. Disponible en: <https://goo.gl/M4zPhN>
- Greene, B. (2001). *El universo elegante*. Barcelona: Crítica.
- Harari, Y. N. (2016). *Homo Deus. Breve historia del mañana*. Madrid: Debate.
- Hawkins, J. (2005). *Numenta*. Recuperado el 12 de junio de 2017. Disponible en: <https://goo.gl/2S9NiK>
- _____. y Blakeslee, S. (2004). *On Intelligence. How a New Understanding of the Brain Will Lead to the Creation of Truly Intelligent Machines*. New York: Henry Holt and Company.
- Hernández-Orallo, J. (1999). Beyond the Turing Test. *Language and Information*, 9.
- Hofstadter, D. R. (1979). *Gödel, Escher, Bach: un eterno y grácil bucle*. Barcelona: Tusquets Editores.
- _____. (1995). *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. New York: Basic Books.
- Iyyer, M., Manjunatha, V., Guha, A. et ál. (2016). *The Amazing Mysteries of the Gutter: Drawing Inferences Between Panels in Comic Book Narratives*. Recuperado el 7 de abril de 2017. Disponible en: <https://goo.gl/Qz9UQu>
- Karpathy, A., Salimans, T., Ho, J. et ál. (2017). *Evolution Strategies as a Scalable Alternative to Reinforcement Learning*. Recuperado el 7 de abril de 2017. Disponible en: <https://goo.gl/kBMJFJ>
- Kidd, C. y Palmeri, H. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, 126(1), pp. 109-114. DOI: <https://doi.org/10.1016/j.cognition.2012.08.004>
- Korb, K. B. y Nicholson, A. E. (2004). *Bayesian Artificial Intelligence*. London: CRC Press.
- Malone, T. W., Laubacher, R. y Dellarocas, C. (2010). The Collective Intelligence Genome. *MIT Sloan Management Review*, 51(3), pp. 21-31.
- Miller, P. (2010). *The Smart Swarm: How Understanding Flocks, Schools, and Colonies can Make us Better at Communicating, Decision Making, and Getting Things Done*. New Jersey: Avery Publishing Group.
- Mischel, W. y Ebbesen, E. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology*, 16(2), pp. 329-337. DOI: <https://doi.org/10.1037/h0029815>
- Nielek, R., Wawer, A. y Wierzbicki, A. (2013). Collaborative Problem Solving in Emergency Situations: Lessons Learned from a Rescue Mission. *Advanced Methods for Computational Collective Intelligence*. London: Springer. DOI: <https://doi.org/10.1007/978-3-642-34300-1>

- Neapolitan, R. E. (2004). *Learning Bayesian Networks*. New Jersey: Pearson - Prentice Hall.
- Nilsson, N. J. (1998). *Artificial Intelligence: A New Synthesis*. San Francisco: Morgan Kaufmann.
- Pal, A., Harper, F. M. y Konstan, J. A. (2012). Exploring Question Selection Bias to Identify Experts and Potential Experts in Community Question Answering. *ACM Trans. Inf. Syst*, 30(2). DOI: <http://dx.doi.org/10.1145/2180868.2180872>
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco: Morgan Kaufmann Publishers.
- Picard, R. (1998). *Los ordenadores emocionales*. Barcelona: Ariel.
- Pinker, S. (1994). *El instinto del lenguaje*. Madrid: Alianza Editorial.
- _____. (2003). *La tabla rasa*. Barcelona: Paidós.
- Punset, E. (2010). *El viaje a la felicidad*. Barcelona: Planeta.
- Quinton, A. (1999). *Hume*. Bogotá: Editorial Norma.
- Ramírez, J. C. y Marshall, J. A. R. (2015). Self-deception can evolve under appropriate costs. *Current Zoology*, 61(2), pp. 382-396.
- Rosen, R. y Klir, G. (1985). *Anticipatory Systems*. New York: Pergamon.
- Russell, S. y Norvig, P. (2014). *Artificial Intelligence: A Modern Approach*. Essex: Pearson.
- Stamos, D. N. (2008). *Evolución. Los grandes temas: sexo, raza, feminismo, religión y otras cuestiones*. Madrid: Biblioteca Buridán.
- Stephens, G. J., Silbert, L. J. y Hasson, U. (2010). Speaker-listener Neural Coupling Underlies Successful Communication. *PNAS*, 107(32), pp. 14425-14430. Recuperado el 7 de abril de 2017. Disponible en: www.pnas.org/cgi/doi/10.1073/pnas.1008662107
- Suárez, É., Bucheli, V. A., Zamara, R. y García, A. (2015). Collective Intelligence: Analysis and Modeling. *Kybernetes*, 44(6/7), pp. 1122-1133, DOI: <https://doi.org/10.1108/K-11-2014-0245>
- Surowiecki, J. (2004). *The Wisdom of Crowds*. USA: Doubleday.
- Trivers, R. (2011). *The Folly of Fools: The Logic of Deceit and Self-deception in Human Life*. New York: Basic Books.
- Whitby B. (1996). *Reflections on Artificial Intelligence*. Exeter, UK: Intellect Books.
- MapReduce_Wiki (2017). *MapReduce*. Recuperado el 16 de julio de 2017. Disponible en: <https://goo.gl/MUJzMD>

Películas y videos

-
- Cameron, J. (1984). *The Terminator*. USA: Hemdale Film. Cinema 84, Euro Film Funding, Pacific Western.
- CdeCiencia (2015). *¿Qué es la vida?* Recuperado el 8 de octubre de 2017. Disponible en: <https://goo.gl/MFs7yp>
- Dalio, R. (2017). *How to build a company where the best ideas win*. TED. Recuperado el 7 de septiembre de 2017. Disponible en: <https://goo.gl/WoP2rM>
- Garland, A. (2015). *Ex Machina*. Reino Unido: DNA Films, Film4.
- Vedantam, S. (2017). Ep.64: I'm right, you're wrong. *Hidden Brain*. Podcast. Recuperado el 3 de octubre de 2017. Disponible en: <https://goo.gl/tHMPn6>

Tesis y trabajos de grado en EVALAB

- Narváez, S. (2016). *Las máquinas y el lenguaje: desarrollo e implementación de un sistema para la asociación entre el lenguaje natural y eventos en un mundo virtual utilizando NUPIC*. Cali: Universidad del Valle.

