

Estado de la cuestión de la inteligencia artificial y los sistemas de aprendizaje autónomo *

State of the Art in Artificial Intelligence and Learning Machines

CARLOS F. ALASTRUEY

Dpto. de Ingeniería
ETSIT
Universidad Pública de Navarra
Campus Arrosadia, 31006 Pamplona-Iruña, Navarra
Correo electrónico: karlos@unavarra.es
ORCID: 0000-0002-4522-4046

Este artículo está sujeto a una: Licencia "Creative Commons Reconocimiento-No Comercial" (CC-BY-NC)

DOI: https://doi.org/10.24197/st.Extra_2.2021.182-195

RECIBIDO: 02/06/2021
ACEPTADO: 14/08/2021

Resumen: En este artículo se muestra una panorámica actualizada de los avances, las posibilidades y el estado de la cuestión de la inteligencia artificial y algunas tecnologías relacionadas. El artículo tendrá la siguiente estructura:

- I. Introducción y panorámica general de la Inteligencia Artificial
- II. Ciencia de datos
- III. Comprensión del lenguaje
- IV. Vehículos autónomos
- V. Edge computing
- VI. Computación Cuántica
- VII. Desafíos en el horizonte

La Inteligencia Artificial se podría definir como un conjunto de algoritmos que establecen una serie de pasos a seguir por un agente artificial para realizar un amplísimo rango de tareas.

Abstract: In this paper an update of the advances, possibilities and state of the art of IA and some related technologies, is shown. The paper will have the following structure:

- I. Introduction and general view of IA
- II. Data Science
- III. Language understanding
- IV. Autonomous Vehicles
- V. Edge Computing
- VI. Quantum Computing
- VII. Challenges in the horizon

IA may be defined as a set of algorithms that establishes a series of steps to be taken by an artificial agent in order to perform a very wide range of tasks. These "series of steps" are not taken in lineal order, but conditional. Deep Learning is defined as the capacity of an artificial system to learn autonomously. Neural networks imitating the

* Este trabajo se ha realizado en el marco del Proyecto de Investigación "LOGGER, el Logos de la Guerra", liderado por Roger Campione y financiado dentro del marco del Plan Nacional.

Estas “series de pasos” no se siguen de forma lineal sino condicional. El aprendizaje profundo se define como la capacidad de un ente artificial de aprender autónomamente. Se definen redes neuronales que imitan la estructura de nuestro cerebro a fin de albergar un gran conjunto de algoritmos para la toma de decisiones complejas.

Palabras clave: Inteligencia Artificial; Learning Machines; Ciencia de datos; Edge computing; Computación cuántica.

structure of the human brain are defined in order to host a big number of algorithms for complex decision-making.

Keywords: IA; Learning Machines; Data Science; Edge Computing; Quantum Computing.

I. Introducción y panorámica general de la Inteligencia Artificial

La Inteligencia Artificial (en adelante IA) tiene por objetivo general trasladar el modo de pensamiento y razonamiento humanos a la computación. Particularmente se propone desarrollar en la máquina la capacidad de percepción, el aprendizaje a partir de la experiencia, relacionar conocimiento – incluyendo conocimiento abstracto o conceptos –, y resolver problemas.

Un aspecto vanguardista de la IA lo constituye la ciencia cognitiva, que se propone emular los procesos del cerebro humano, y en concreto las capacidades de razonar, oír, hablar y emocionarse. De hecho, una posible definición de la IA es aquella rama del conocimiento que se encarga de llevar a cabo procesos computacionales capaces de realizar tareas sobre la base de dos características humanas fundamentales: el razonamiento y la conducta. Existen otras definiciones alternativas de la IA que también contemplan al ser humano como referencia. Cuatro de ellas son citadas por Torra (2020).

La primera sería lograr que las máquinas dotadas de IA actúen como las personas (definición de McCarthy). El método para evaluar el rendimiento de un programa respecto a la IA sería por tanto usar de modelo al propio ser humano y comparar la forma de actuar. El test de Turing, que examinaremos un poco más adelante, también utiliza este enfoque. Como ejemplo, el bot conversacional Eliza entra en esta categoría.

La segunda tendría como objeto de análisis el razonamiento humano. Se trataría de conseguir máquinas que razonen como las personas. Es importante observar aquí que estamos hablando de que el objeto de una IA así entendida es el propio razonamiento humano, y no el resultado de dicho razonamiento. De esta desafiante área se encarga la ciencia cognitiva.

Una tercera definición, también basada en el razonamiento, tiene una filosofía sin embargo bien distinta: razonar racionalmente. Es decir, se supone que es posible identificar un modo racional de efectuar el razonamiento. Aquí

interviene la lógica para conceptualizar los modos de razonamiento y ser capaces de sistematizarlos en una máquina de IA.

En cuarto lugar, podríamos regresar a la primera definición, pero modificándola de acuerdo a la filosofía de la tercera: buscaríamos crear una IA que actuase racionalmente. Un ejemplo claro de este planteamiento lo constituyen los sistemas de IA que juegan al ajedrez. No nos interesa en última instancia cómo se calcula el resultado, sino que nos interesamos por el propio resultado: ganar.

Además, existen dos planteamientos de base en torno a la IA. Lo que se conoce como inteligencia artificial débil, y por contraste lo que llamamos inteligencia artificial fuerte. En un paradigma de IA débil, se parte de que las máquinas únicamente podrán simular que razonan, aunque en realidad no lo hacen. Quienes adoptan este punto de vista consideran que es esencialmente imposible crear una máquina consciente, por lo que a lo que puede aspirarse es a construir una simulación de un proceso cognitivo, pero en última instancia es imposible producir artificialmente un proceso cognitivo genuino.

Los que promueven la búsqueda de una IA fuerte plantean como meta crear una máquina dotada de una mente, entendida esta como un conjunto de estados mentales posibles. Por tanto, en el horizonte temporal se considera posible crear una máquina dotada de todas las capacidades mentales de los seres humanos, incluyendo el razonamiento y la imaginación. Es este el punto de vista adoptado mayoritariamente y el escenario sobre el que se trabaja en la actualidad.

La IA entendida como ciencia se ocupa de los procesos cognitivos, y por ello tiene una fuerte conexión con la neurología y la cognición. En este aspecto uno de los frentes más vanguardistas se define en relación con la ciencia cognitiva, es decir, emular los procesos tales como razonar, oír, ver, hablar, emocionarse. Por otro lado, la IA entendida como ingeniería se centra en la ingeniería del conocimiento, es decir, en el desarrollo de métodos para la resolución de problemas.

Una de las áreas en las que se usa profusamente la IA y tiene más influencia en nuestra vida diaria es en las redes sociales (Wang, 2019). Dichas redes diseñaron sistemas de IA que se perfeccionan continuamente con gran dedicación de recursos, a fin de emular el comportamiento de los usuarios de forma personalizada, tratando de maximizar el tiempo de uso, que el usuario interactúe con otros usuarios e invite a usuarios potenciales a unirse a la red (interactividad), así como efectuar compras dentro de la red.

Para aquellas áreas de la IA, como su aplicación en redes sociales, que buscan crear sistemas que actúan como humanos, el modelo es el propio ser humano. Como camino para diferenciar si un sistema artificial constituye IA de tipo humanoide o no, existe cierto consenso en la utilidad del test de Turing.

El test de Turing fue en principio denominado el juego de imitación por su creador Alan Turing (1950). Consiste en un test para comprobar la habilidad de una máquina para mostrar un comportamiento inteligente indistinguible del de un ser humano. Turing propuso que el test sería aplicado por un evaluador humano para examinar conversaciones entre un ser humano y una máquina que hubiera sido construida para producir respuestas en lenguaje natural. La conversación tendría lugar de forma escrita, mediante intercambio de texto, para que el resultado no dependa de la capacidad de imitar el tono de voz humana o condicionantes similares (de lo que se trata es de evaluar la inteligencia según un modelo humano de referencia). Si el evaluador no puede determinar a ciencia cierta quién es humano y quién es artificial en la conversación, se considera que la máquina ha aprobado el test de Turing. Hay que observar que el resultado del test de Turing no depende de que la máquina sea capaz de dar respuestas correctas a las preguntas que se le puedan plantear, sino de cuánto se parecen las respuestas de la máquina a las que un humano hipotético pudiera dar en semejantes circunstancias.

Se entiende por tanto que si un sistema artificial aprueba el test de Turing, puede pasar por humano. El problema podría ser lo que algunos autores llaman “artificial stupidity”, es decir, que para ser reconocido como humano, un sistema de IA tiene que poseer ciertas características de irracionalidad y tendría que cometer ciertos errores y mostrar ciertas lagunas e incluso cierta imprevisibilidad en su comportamiento. La prueba de Turing da positivo ante la estupidez y da negativo ante inteligencia sobrehumana y ante una coherencia infalible. Hay que tener en cuenta, por tanto, que el test de Turing no evalúa en realidad la inteligencia de la IA, sino su “humanidad”.

En su trabajo de 1950, Alan Turing propone considerar una cuestión filosófica: ¿puede una máquina pensar? Dado que “pensar” es difícil de definir, Turing decide sustituir dicha pregunta por otra directamente relacionada con la anterior pero expresada en términos menos ambiguos: ¿Existirán computadores que aprobarán el juego de la imitación (el posteriormente denominado Test de Turing)? Turing declara su convicción de que esta segunda pregunta se puede responder y dedica el resto de su artículo a refutar las objeciones principales a la primera pregunta, “¿puede una máquina pensar?”. Este trabajo de Alan Turing recibió y recibe amplio eco en la comunidad científica y se ha convertido en un concepto importante en la filosofía de la inteligencia artificial (Russell & Norving, 2003; Warwick, 2010). Algunos de sus críticos son a su vez objeto de controversia.

En otras áreas de aplicación fuera de las redes sociales se busca más específicamente la creación de sistemas de IA que siguen el modelo humano de pensamiento, razonamiento y aprendizaje, pero buscando que lleguen muchísimo más lejos en sus logros teniendo en cuenta la posibilidad de acceder

con enorme velocidad a una descomunal cantidad de datos de muy diversas categorías.

Para los seres humanos con inteligencia promedio es muy difícil percibir realidades o relaciones que se salgan de la intuición o de lo que la persona ha experimentado de una u otra forma. Por ejemplo, en el ámbito matemático, los espacios no euclídeos no resultan intuitivos para un ser humano no entrenado; así, las relaciones matemáticas – incluso las más sencillas – que son ciertas en un espacio euclídeo y falsas en un espacio no euclídeo – pueden suponer un gran desafío para la intuición de una persona normal. Como ejemplo, la suma de los ángulos de un triángulo plano, que da 180 grados, no se cumple si el triángulo está trazado sobre una superficie esférica, pongamos por caso. Cuando los objetos de estudio son mucho más sofisticados resulta prácticamente imposible para el ser humano darse cuenta de las limitaciones que presenta su intuición o su manera de resolver los problemas, a no ser que se dote de potentes herramientas de análisis. Sin embargo, para un sistema de IA es irrelevante que el sistema a analizar sea – en nuestro ejemplo – euclídeo o no, si cuenta con una definición correcta del problema.

Hay algunas grandes áreas en la investigación actual en IA. Citamos en primer lugar los sistemas basados en el conocimiento. Estos sistemas, cuando se aplican en campos específicos del quehacer humano, precisan incorporar conocimiento relevante a dicho campo. Por ejemplo, la aplicación de la IA a la medicina requiere establecer mecanismos para representar el conocimiento médico acumulado a fin de que la máquina pueda utilizarlo convenientemente en una tarea médica. Esto es así en las cuatro grandes áreas de aplicación de la IA en medicina hoy en día: diagnóstico de enfermedades, desarrollo de fármacos, tratamiento personalizado y edición genética (Amisha 2019).

En cuanto al diagnóstico de enfermedades, la situación actual – que empeora cada año – es que la demanda de expertos excede con mucho el número de expertos disponibles. Desarrollar la capacidad de realizar diagnósticos eficaces lleva a una persona muchos años de formación médica rigurosa. En los últimos años, los llamados algoritmos de aprendizaje profundo (Deep Learning machines) han realizado enormes avances en el diagnóstico automático de enfermedades, lo que hace que la diagnosis sea más barata y accesible para el gran público. De todas formas la aplicación del aprendizaje automático a la diagnosis médica está aún en sus comienzos, y ahora se están desarrollando sistemas más ambiciosos de diagnóstico que usan la combinación de fuentes diversas de datos (genómica, proteómica, datos del paciente, escáners, resonancias e incluso documentos a mano).

Respecto al desarrollo de fármacos, hay que tener en cuenta que es un proceso enormemente costoso y laborioso. Los procedimientos de análisis y extracción de conocimiento necesarios para diseñar un fármaco ya se ejecutan de forma mucho más eficiente por medio de sistemas que incorporan IA. Si se

generaliza el uso de la IA para este tipo de tareas, se ahorrarán años de trabajo (y el consiguiente retraso en disponer de remedios en caso de situaciones de pandemia, por ejemplo) y centenares de millones de euros de inversión. Una estrategia que se sigue en la actualidad es usar la IA para encontrar biomarcadores a fin de diagnosticar la enfermedad de modo seguro y barato. Ya hay indicios prometedores que señalan el éxito de la IA en el diseño de fármacos (Lowe 2019).

Otra área de investigación muy activa es la del aprendizaje automático o autónomo. Se trata de conseguir que la máquina aprenda de su experiencia y en particular que pueda obtener aprendizajes observando sus propios errores y aciertos. Ya hay diseñados algunos métodos que avanzan en este objetivo, y se espera que en los próximos años esta área se desarrolle y refine ya que afecta directamente al corazón de la propia IA.

Directamente relacionada con la anterior se encuentra la capacidad de una máquina para resolver problemas y situaciones complejas y diversas. A fin de desarrollar esta capacidad, un paso previo indispensable consiste en poder formalizar el problema o situación. Y una vez que el problema está formalizado, otra capacidad a desarrollar consiste en establecer y probar métodos de resolución, y también en comprobar que la resolución es satisfactoria al menos en parte.

La IA descentralizada o distribuida es asimismo un campo de rápido desarrollo de la IA en los últimos años (Shoham and Leyton-Brown 2009). Se trata de un enfoque que está demostrando ser adecuado para dar cabida al aprendizaje complejo, la toma de decisiones y la planificación. Al descentralizar las operaciones de cómputo y de gestión de datos en un conjunto de máquinas en una red, con la capacidad de actuar en paralelo, se habilita la posibilidad de resolver problemas que requieran usar datos a gran escala. La IA descentralizada consiste en nodos de procesamiento con aprendizaje autónomo, que se hallan distribuidos en escalas muchas veces enormes. A estos nodos se los conoce con el término de “agentes”. Los nodos de la IA distribuida realizan sus tareas de forma autónoma. Las resoluciones parciales que realiza cada nodo se van integrando paulatinamente mediante determinados protocolos de intercambio de datos entre nodos. Teniendo en cuenta esta arquitectura, las operaciones se pueden realizar de forma asíncrona, lo que facilita la optimización de recursos. No se requiere que cada nodo efectúe las tareas que se le asignan de forma sincronizada con otros nodos, lo que añadiría datos de sincronización a la red local y, lo que es más importante, provocaría en ocasiones tiempos de espera de parte de los nodos hasta que los otros nodos con los que están sincronizados terminan su labor asignada. Si un nodo termina su tarea, puede asumir otra a continuación sin tener que esperar; con ello se optimiza el uso de los recursos en la red de agentes. Dado el gran tamaño que puede alcanzar un sistema de IA descentralizada, este tipo de sistemas son robustos (es decir, insensibles a

indefiniciones, imprevistos o inexactitudes), y se construyen con la capacidad de adaptarse a los cambios en la definición de los problemas. Los sistemas de IA descentralizada no precisan que todos los datos necesarios estén alojados en un mismo lugar. Este hecho contrasta fuertemente con lo que se denominan sistemas de IA monolítica o centralizada, que se encuentran en fuerte conexión y proximidad con sus nodos de procesamiento. Además los sistemas de IA distribuida tienen la ventaja de poder seguir trabajando sin interrupción mientras se actualizan los datos sobre los que está realizando las tareas asignadas.

Algunas áreas de aplicación de la IA descentralizada o distribuida las constituye el comercio electrónico (e-commerce), las redes de telecomunicaciones (WLAN), gestión de tráfico de vehículos, gestión de flujo de stock en sistemas de venta por internet, el estudio de la vida artificial o vida simulada, y la supervisión y mantenimiento de las redes eléctricas (Catterson et al, 2012).

II. Ciencia de datos

Según Richard Feynman (2000), la Ciencia de Datos en realidad debería llamarse Ingeniería de Datos. Esta disciplina se ocupa de los métodos y procesos para extraer conocimiento a partir de gran cantidad de datos. La ciencia de datos comprende las áreas de estadística, minería de datos, aprendizaje automático y analítica predictiva. Las grandes aplicaciones de la ciencia de datos se encuentran en dos áreas. Por una parte en el marketing, donde se realiza estudio y maximización de la probabilidad de que alguien compre un determinado producto. Tanto en las redes sociales como en plataformas tales como Netflix y Spotify se aplican continuamente sistemas de IA con estos fines. Por otra parte la ciencia de datos se aplica en gobernanza, en áreas como seguridad, propiedad de datos, privacidad, movilidad urbana sostenible y ciudades inteligentes. Además hay organismos tanto públicos como privados que impulsan las aplicaciones de IA para promover la mejora de la rendición de cuentas de los gobiernos ante los ciudadanos.

Por big data (macrodatos) se quiere significar grandes volúmenes de datos no procesables con las herramientas tradicionales (L'Heureux et al, 2017). Los volúmenes de datos que se manejan actualmente por grandes y no tan grandes corporaciones, empresas y organizaciones no son almacenables en bases de datos tradicionales ni son procesables por procesadores estándar.

Los objetivos de la tecnología referente a big data son: extraer valor del big data, comprender y comunicar conclusiones, crear soluciones para el aumento de beneficios. Además se diseñan experimentos adicionales cuando se necesitan a fin de aprender sobre propiedades de los datos. A este respecto, la película documental "The Social Dilemma" muestra el diseño de tales

experimentos en el contexto del manejo de big data para el desarrollo de las redes sociales.

III. Comprensión del lenguaje

A fin de avanzar en la comprensión y uso del lenguaje humano por parte de la IA, se desarrolla una disciplina que denominamos PLN, o “procesamiento del lenguaje natural”. En el contexto de la IA, por “lenguaje natural” entendemos lenguaje humano. El PLN se relaciona fuertemente con las ciencias de la computación, la IA y la lingüística. Su objetivo principal es la formulación e investigación de sistemas computacionales eficaces a fin de promover la comunicación entre la persona y la máquina mediante el lenguaje natural.

Las principales áreas de aplicación del PLN son: traducción automática de textos, producir respuestas a preguntas humanas, síntesis de voz, recuperación de información, comprensión del lenguaje – incluyendo cómo usamos símbolos para expresar ideas y sentimientos; la máquina debe interpretar el mensaje y entenderlo en significado, contexto e intención -, reconocimiento del habla – en el contexto de la comunicación hombre/máquina -, y extracción de información – lo que se refiere a extraer información estructurada a partir de datos legibles.

Las tendencias actuales en la investigación para aplicaciones de la comprensión del lenguaje son: automatizar trabajos – tareas repetitivas en una empresa u organización -, análisis de sentimientos – comunicación con el público, evaluación de una campaña en marcha -, traducción automática del lenguaje hablado, sistemas conversacionales – para generar respuestas adecuadas (aquí cobra relevancia el test de Turing) -, y por último extracción de contenido realmente importante y relevante.

IV. Vehículos autónomos

Para el desarrollo del campo de los vehículos autónomos es esencial desarrollar la capacidad de predecir las intenciones y reacciones humanas. En este sentido una de las experiencias punteras hoy en el mundo es el grupo de vehículos autónomos del MIT, aplicado específicamente al automóvil, en donde se está configurando el vehículo autónomo de acuerdo al pensamiento humano. En la IA para vehículos autónomos se aplica la psicología social. Se busca que cada automóvil adapte su comportamiento a la personalidad de cada conductor, mientras que en los modelos previos se suponía que todas las personas desarrollaban el mismo tipo de pensamiento, lo cual obviamente no es cierto.

En el MIT se ha desarrollado el algoritmo SVO (Social Value Orientation) en el que se clasifican a los conductores en tres categorías: egoísta, cooperativo y altruista. Gracias al algoritmo SVO se espera que el coche

autónomo podrá predecir las intenciones humanas para lograr una mejora de la seguridad vial. Dentro de este apartado cobra vital importancia el aspecto de la moral del vehículo (lo que en la jerga especializada se conoce como the trolley problem o el problema del tranvía. La moral del coche autónomo depende de la decisión que tome la persona que lo programe. Así, la acción del vehículo en cuestión se hallará condicionada por lo aprendido durante el proceso de entrenamiento.

V. Edge computing

La Edge Computing o computación en el límite, se refiere a los procesos de computación diseñados y enfocados para que el procesamiento y el almacenamiento de datos se acerquen todo lo razonablemente posible al lugar donde se necesita, es decir, al lugar donde se va a acceder a dichos datos y se va a operar con ellos. Esto se hace a fin de mejorar el tiempo de respuesta evitando retardos indeseados producidos al atravesar los diferentes nodos de la red y sus sistemas de permisos y prelación. También se pretende así optimizar el uso del ancho de banda, ya que muchos datos no viajan por tramos prolongados de la red, sino que se mueven localmente.

El origen de la edge computing está en los servidores de plataformas de vídeo, a finales de los años 1990, que buscaban aliviar sus redes del altísimo tráfico de datos relacionado con el visionado de vídeo en los hogares. La edge computing busca por ello realizar la computación fuera de la nube, y en consecuencia llevarla al extremo de la red. Y esto es particularmente crítico en aquellas aplicaciones que requieren realizar el procesamiento de datos en tiempo real. Por ello la edge computing se lleva a cabo en servidores cerca del final de la red. Por “cerca del final de la red” se entiende a una distancia como máximo de dos nodos respecto al lugar donde se encuentra el usuario y donde, por tanto, va a tener lugar el uso y procesamiento de los datos.

El uso de las técnicas asociadas a la edge computing aumenta la velocidad de computación (la que resulta crítica por ejemplo en el reconocimiento facial, ya que los seres humanos tardamos aproximadamente medio segundo en reconocer mediante la vista a otra persona, y por tanto el uso de la edge computing puede imitar mejor dicha percepción).

También se produce una mejora en la eficacia, como por ejemplo en el procesamiento de vídeo. Si solo se usa la red local para la transferencia de grandes vídeos, la red general no llega a saturarse.

Las aplicaciones principales en la edge computing son: reducción de tráfico en la red y disminución por tanto de los costes de transmisión, la mejora del tiempo de respuesta de las aplicaciones, la disposición y uso de juegos en red, el funcionamiento de coches conectados entre sí y comunicación automática

entre vehículos circulando en proximidad o en una misma región o área, y el funcionamiento y optimización de ciudades inteligentes.

VI. Computación cuántica

Richard Feynman, en sus célebres conferencias sobre computación (Feynman 2000), analizó las limitaciones físicas de las computadoras teniendo en cuenta las leyes de la Física. De acuerdo con Bennett, la disipación de energía asociada al proceso de computación se puede hacer arbitrariamente pequeña, siendo sus límites las leyes de la mecánica cuántica y el principio de incertidumbre. Richard Feynman se concentra en concretar algún hamiltoniano para un sistema que pueda hacer las veces de computadora. Recordamos que el hamiltoniano en el formalismo de la mecánica cuántica, es un operador que actuando sobre un vector en el espacio de Hilbert que represente el estado físico del sistema, obtiene la observable energía. Por tanto, el operador hamiltoniano que busca Feynman servirá para determinar si la energía del sistema se conserva en un proceso cuántico de computación. Dado que las leyes de la física cuántica son reversibles en el tiempo, Feynman se enfoca en su estudio en máquinas cuánticas de computación que obedezcan tales leyes reversibles. Si buscamos una máquina con tal propiedad, la arquitectura de la computadora debe estar compuesta de una combinación de puertas lógicas exclusivamente reversibles, para que sea realizable a un tamaño extraordinariamente pequeño (de hecho, lo que se busca es llegar a la frontera de trabajar con átomos como unidades de la máquina de computación). En el estudio se busca cuál es la cantidad mínima de energía libre que debe consumirse para hacer funcionar una computadora ideal compuesta exclusivamente de puertas lógicas reversibles. Durante mucho tiempo se pensó en la comunidad científica que el aumento mínimo (límite mínimo absoluto) de entropía en una operación de computación era igual a \log_2 unidades, lo que representa una generación de calor igual a $kT \log_2$ a temperatura T, siendo k el número de operaciones simples para la realización de la operación en curso. Las computadoras basadas en transistores clásicos disipaban 10.000 millones de veces esa cantidad. La naturaleza, en el proceso de copiado del ADN, disipa 100 kT por bit copiado. Aunque cuando Feynman da sus famosas conferencias, la tecnología estaba increíblemente alejada del mínimo teórico de disipación de energía, se plantea una supuesta máquina que fuera capaz de escribir bits en átomos individuales, en lugar de los 100.000 millones de átomos que necesitaba la escritura de un bit en aquella época. Lo fascinante del asunto es que las conclusiones a las que llega el estudio de Feynman son aplicables a la computación cuántica de hoy en día, precisamente porque el único límite físico que se antepuso en el estudio fue el del propio átomo. Para ello Feynman supone que podemos contar con un “sistema” que es un átomo y que admite dos estados, de tal forma que dependiendo del valor del

estado binario del átomo, dicho estado representará un 0 o un 1. La aportación de Feynman consiste en haber abierto el camino teórico para el diseño de las computadoras cuánticas 30 años después.

El paradigma de la computación cuántica es esencialmente diferente al de la computación clásica. No se trata simplemente de construir máquinas más rápidas o potentes. La misma arquitectura de la máquina es radicalmente distinta y nueva respecto a la computadora clásica. En una máquina clásica, la unidad básica es el bit, que admite dos estados 1 o 0. En cambio, una máquina cuántica se compone de qubits, que son unidades que admiten también dos estados, pero que pueden de hecho encontrarse en ambos estados simultáneamente, de acuerdo con las leyes de la mecánica cuántica.

La capacidad de cómputo de un computador clásico crece linealmente con su tamaño. Si por ejemplo un computador con una potencia de cálculo digamos igual a 2, multiplica por 10 su tamaño, pasaría a tener una potencia de cálculo de 20. Pero en los computadores cuánticos la potencia de cálculo crece exponencialmente con el aumento de tamaño. Por lo que en el ejemplo anterior nuestro ordenador cuántico con una potencia inicial de 2, si viera su tamaño multiplicado por 10, pasaría a tener una potencia de cálculo igual a 1024, más de 50 veces superior al caso clásico. Este aumento es posible debido al fenómeno del paralelismo cuántico que se deriva del principio de superposición de estados de la mecánica cuántica. La capacidad – tanto de almacenamiento como de procesamiento - de un computador cuántico crece de forma exponencial con el número de qubits (unidades básicas del computador cuántico). No obstante, medir estados cuánticos entraña una dificultad muy relevante, ya que los dos estados que se superponen en un qubit, al ser medidos pueden dar medidas diferentes incluso para dos estados idénticos. Durante bastante tiempo esto se consideró una limitación fundamental para la realización de computadores cuánticos. Sin embargo Shor [1994] publicó algoritmos polinomiales que permitían en un computador cuántico factorizar números enteros y calcular logaritmos discretos. A partir de dicho avance, se produjo un gran aumento de la investigación en computación cuántica, particularmente en el desarrollo de algoritmos para una máquina cuántica como en el diseño de una máquina físicamente realizable. De acuerdo con García et al (2016), los resultados de Shor rompieron el sistema criptográfico RSA, el más utilizado en el mundo. Ese hecho impulsó el desarrollo de sistemas criptográficos cuánticos. Para garantizar la confidencialidad en los canales cuánticos se aprovecha el hecho de que los estados cuánticos no se pueden copiar o clonar. Algunos de los desarrollos que están teniendo lugar ahora en el terreno de la computación cuántica tienen relación precisamente con el área de las comunicaciones: la confidencialidad, la transmisión de información clásica a través de canales cuánticos y la transmisión de estados cuánticos propiamente dichos.

Las grandes perspectivas que supone alcanzar la supremacía cuántica (lograr construir computadores cuánticos con rendimiento superior a las máquinas clásicas) han impulsado enormemente la investigación en este campo. Muy recientemente (2019) se presentaron los primeros computadores cuánticos de uso comercial: el IBM Q System One, que combina computación cuántica y convencional en una máquina de 20 qubits. Por otra parte la revista Nature publicó en octubre de 2019 que oficialmente Google había alcanzado la supremacía cuántica.

Las áreas en las que se está investigando actualmente para desarrollar la computación cuántica y el transporte cuántico de datos son la electrónica molecular, el entrelazamiento cuántico, la teleportación cuántica, la computación basada en ADN y la criptografía cuántica entre otros (Wichert 2014).

El impacto que está teniendo la computación cuántica afectará a futuros desarrollos de la Inteligencia Artificial. Por ejemplo, el desarrollo práctico de redes neuronales requiere del tratamiento de enormes números de datos, imágenes y otras estructuras a fin de poder entrenar la red neuronal en cuestión. Los computadores clásicos, por sus limitaciones físicas, no pueden ofrecer suficiente potencia de cómputo para que una red neuronal funcione en la práctica según los requerimientos y exigencias al que está sometido un cerebro comparable al de un ser humano. La computación cuántica, sin embargo, abre una ventana completamente nueva para el desarrollo de IA basada en redes neuronales.

VII. Desafíos en el horizonte

Las previsiones son que el gasto mundial en IA se duplicará anualmente. En 2021 se prevé una inversión global de 57.600 millones de dólares, de acuerdo con la Fundación Santander (Santander [2020]). Podríamos para terminar señalar algunos desafíos tanto en positivo como en una vertiente negativa, en lo que al desarrollo de la IA se refiere en los decenios por delante.

En el “haber” de la IA podemos citar la liberación de tareas rutinarias y aumento de la productividad, la toma más rápida de decisiones mediante tecnologías cognitivas, la prevención de errores, o la asunción de riesgos por la IA en lugar de por seres humanos. Por otro lado, los desafíos para los que aún no se vislumbra una solución clara, son los altos costos, la dificultad para elaborar pensamiento artificial fuera de la programación, la falta de desarrollo en la IA actual de la compasión y la empatía, y la creación de relaciones de alta dependencia de los seres humanos ante las máquinas.

REFERENCIAS BIBLIOGRÁFICAS.

- Amisha, Paras Malik, Monika Pathania & Vyas Kumar Rathaur, (2019). Overview of artificial intelligence in medicine. *Journal of Family Medicine and Primary Care*, 8(7), 2328-2331.
- Catterson, V.M., Davidson, E.M., & McArthur, S.D.J., (2012). Practical applications of multi-agent systems in electric power systems. *European Transactions on Electrical Power*, 22(2), 235-252.
- Feynman, R., (2000) [1996]. Lectures on computation. *Perseus Publishing*.
- Fundación Banco de Santander, (2020). Los principales retos de la Inteligencia Artificial. <https://www.santander.com/es/sala-de-comunicacion/dp/los-principales-retos-de-la-inteligencia-artificial>
- García López, A; García López de Lacalle, J. & García Mazarío, F., (2016). Computación cuántica, *Máster en Ciencias y Tecnologías de la Computación, Universidad Politécnica de Madrid*.
- L'Heureux, A.; Grolinger, K.; Elyamany, H. F.; Capretz, M. A. M. (2017). "Machine Learning With Big Data: Challenges and Approaches". *IEEE Access*. 5: 7776–7797.
- Lowe, D., (2019). Has AI Discovered Drug Now? *Science*. September 5th.
- Russell, S. & Norvig, P., (2003) [1995]. Artificial Intelligence: A Modern Approach. *Prentice Hall*.
- Shoham, Yoav; Leyton-Brown, Kevin (2009). *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. New York: Cambridge University Press. ISBN 978-0-521-89943-7.
- Shor, P. W., (1994). Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, *SIAM J. Comput.* 26, 1484–1509 (arXiv:quant-ph/9508027).
- Torra, V., (2020). La inteligencia artificial, http://www.fgcsic.es/lychnos/es_es/articulos/inteligencia_artificial,

Instituto de Investigación en Inteligencia Artificial, Consejo Superior de Investigaciones Científicas.

Turing, A., (1950). "Computing Machinery and Intelligence", *Mind*, LIX(236), 433-460.

Wang, Peyton (2019), "The Impact Of Artificial Intelligence On Social Media", medium.com.

Warwick, Kevin (2010), "Hidden Interlocutor Misidentification in Practical Turing Tests", *Minds and Machines*, **20** (3): 441-454.

Wichert, A., (2014). Principles of Quantum Artificial Intelligence, *World Scientific Publishing*.