

El concepto de inteligencia artificial según allen newell

Serapio Cazana Canchis¹

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Recepción: 20 de marzo, 2017

Aceptación: 15 de abril, 2017

Palabras claves:

Filosofía de la tecnología

Inteligencia artificial

Psicología teórica.

RESUMEN

Gottfried Wilhelm Leibniz a inicios del s. XVIII diseñó una máquina de cálculos, en la que creía como posible se expresara en términos lógicos. En esta misma tarea se embarcó Jacques de Vaucanson quien construyó estatuas que manifestaban conductas básicas. Esto ahora no extraña, pero en aquella época debió parecer asombroso y hasta mágico o herético. Actualmente, desde los inicios de la inteligencia artificial tanto en su teoría como en su aplicación tecnológica, se viene gestando la posibilidad de generar un tipo de inteligencia no humana y que sea capaz de emular o de realizar las mismas conductas hasta llegar al propio pensamiento, parte nuclear de la inteligencia humana, donde cabe preguntarse ¿es esto posible? Este ensayo analiza esta posibilidad desde el concepto de inteligencia artificial de Allen Newell.

The concept of artificial intelligence according to Allen Newell

ABSTRACT

Keywords:

Philosophy of technology

artificial intelligence

Theoretical psychology.

Gottfried Wilhelm Leibniz at the beginning of the s. XVIII designed a calculating machine, in which he believed it possible to express himself in logical terms. Jacques de Vaucanson, who built statues that manifested basic behaviors, embarked on this same task. This is not strange now, but at that time, it must have seemed amazing and even magical or heretical. Currently, from the beginnings of artificial intelligence both in its theory and in its technological application, the possibility of generating a type of non-human intelligence and capable of emulating or performing the same behaviors until reaching one's own thinking has been developed, nuclear part of human intelligence, where it is possible to ask is this possible? This essay analyzes this possibility from the concept of artificial intelligence of Allen Newell.

¹ Licenciado en Filosofía por la Facultad de Teología Pontificia y Civil de Lima. Master en Filosofía por la Universidad de Salamanca. Docente de la Universidad San Ignacio de Loyola, Lima, Perú. Email: secazana@usil.edu.pe ORCID: <https://orcid.org/0000-0002-9755-6085>

Introducción

Como narra Paul M. Churchland en su libro *Materia y conciencia* (1992; p.184), el filósofo alemán Gottfried Wilhelm Leibniz (1646–1716) no sólo construyó una máquina que podía sumar y restar, sino también afirmaba que existía la posibilidad de lograr un lenguaje perfectamente lógico, en cuyo sistema todo el pensamiento humano podría expresarse mediante cálculos.

Posteriormente, Jacques de Vaucanson (1709–1782) construyó varias estatuas que manifestaban algún tipo de conducta; como por ejemplo, un pato que podía comer, beber, graznar y chapotear en el agua. Incluso había diseñado una estatua de forma humana que podía tocar la flauta. Quizá todo ello ahora se pueda conseguir en una tienda de juguetes para niños y no entraña ninguna rareza, pero a una mente humana del siglo XVIII ello le debió parecer asombroso y hasta herético. Hoy en día, después de varias décadas desde los inicios de la inteligencia artificial, tanto en su fase teórica como en su aspecto de realización tecnológica, a muchos pensadores también les parece que es algo prohibitivo pensar en la posibilidad de generar un tipo de inteligencia no humana y que sea capaz de emular o de realizar las mismas conductas.

Sin embargo, ¿qué posibilidades reales existen de que ello pueda conseguirse? Es importante notar que un componente fundamental para el desarrollo de la inteligencia artificial es la lógica y la matemática, por su capacidad para generar algoritmos. La imaginación literaria ya había plasmado hasta los perfiles psicológicos de máquinas que adquirirían características del pensar humano, pero una cosa distinta era contar con los elementos técnicos para sus concreciones iniciales. Pero podríamos decir que además de la imaginación que da libre vuelo a las imágenes, está

la imaginación teórica, aquella que toma el camino de la lógica, la visualización de la concreción técnica de un determinado proceso y automatismo.

En efecto, la ausencia de dispositivos tecnológicos constituyeron durante mucho tiempo la principal barrera para llevar a cabo la imaginación lógica de Leibniz o del matemático inglés Charles Babbage (1791–1871) –quien diseñó una especie de computadora capaz de realizar todas las operaciones lógicas y aritméticas elementales– pero que no pudo construir la máquina al cien por ciento precisamente porque la finura tecnológica no había logrado el necesario desarrollo. Con el pasar del tiempo la imaginación abstracta, lógica, pudo abrirse paso con la gran ola de la tecnología electrónica que irrumpió en las primeras décadas del siglo XX. Así la lógica de las proposiciones, la lógica de la geometría, la aritmética, el álgebra eran susceptibles de fijarlas en un algoritmo para que realicen operaciones complejas casi por sí mismas. Los sistemas formales y la tecnología poco a poco se dieron un memorable encuentro.

El concepto de inteligencia artificial clásica

Si bien ahora se ha avanzado mucho en un concepto unificado de lo que es inteligencia artificial², en la década del 50 y en las siguientes los teóricos y realizadores trabajaron mucho para definir el concepto.

La Conferencia de Dartmouth, que se llevó a cabo en 1956 (Hanover, New Hampshire) significó el inicio de la inteligencia artificial. En dicho evento se reunieron matemáticos y lógicos con el fin de describir algunos rasgos específicos de la inteligencia humana y para ver la manera de simularla mediante por un computadora. Sin embargo, aunque el concepto quede aparentemente obvio, es importante tener una

² Martínez-Freire lo define del modo siguiente: “la inteligencia artificial puede caracterizarse de modo amplio como aquella rama de la informática que estudia la teoría, el diseño y construcción de máquinas inteligentes, entendiendo por tales los mecanismos capaces de ejecutar tareas que realizadas por los humanos atribuimos a su inteligencia. Por ejemplo, los computadores dotados

de programas para jugar al ajedrez, o con programas para demostrar teoremas lógicos o matemáticos, o con programas para diagnosticar e indicar el tratamiento para curar una enfermedad, pueden considerarse máquinas inteligentes”. La importancia del conocimiento. Filosofía y ciencias cognitivas, 2ª ed; p. 20.

definición lo más precisa posible de lo que ello quiere decir. Una de las razones estriba en el hecho de que en un debate epistemológico se precisa de un concepto de significación común para luego posicionarse en una opinión u otra. Además, si bien es cierto en las décadas que han seguido a dicha conferencia tanto el concepto como su realización han tenido algún ajuste de la perspectiva³, es oportuno considerar una formulación exhaustiva que hizo del término Allen Newell⁴, uno de los participantes en dicha conferencia.

Al inicio de su consideración del tema Newell enfatiza la diferencia entre lo que él considera inteligencia artificial y lo que es propiamente la mente. De hecho, afirma tener mucho más clara la naturaleza de la inteligencia artificial que los problemas que deben plantearse sobre el concepto de la mente humana.

Para el autor que estamos considerando, es importante definir lo que es un mecanismo, ya sea general o especial, orientado al campo de la computación. Así, resumidamente, un mecanismo es, por una parte, cualquier proceso físico determinante para un resultado. Además, un proceso abstracto también podría recibir tal nominación si éste es susceptible de realizarlo mediante un proceso físico. En ese sentido, cualquier Software de un ordenador constituye un mecanismo. También lo es una regla por medio de la cual se podría construir un artefacto, como por ejemplo la Máquina de Turing o el Algoritmo de Markov. Incluso un sistema formal, que podría ser un sistema lógico, pero susceptible de realización mediante una máquina sería un mecanismo. Sin embargo, no se trata sólo de mecanismos que procesen información de modo lineal y directo, sino también incluye la probabilidad y la estadística. Esto es importante porque la inteligencia artificial posterior había considerado la probabilidad en los sistemas de respuesta de los ordenadores.

³ Por ejemplo, la inteligencia artificial clásica y la inteligencia artificial conexionista.

⁴ Newell fue hijo de un profesor de la Universidad de Stanford. Vivió en San Francisco y se graduó en la universidad donde enseñaba su padre. Luego pasó un año en Princeton, y al pensar que no era un

Recalcando lo decisivo del proceso lógico y la concepción matemática, conviene precisar un poco más el tipo de mecanismo abstracto o formal que interviene de modo decisivo en la consecución de la inteligencia no humana. A este respecto, Churchland menciona que un sistema formal consiste en un conjunto de fórmulas y un conjunto de reglas que permiten manejar y transformar esas fórmulas en otras. Así, hay muchos sistemas formales, que establecen la estructura de una fórmula y el patrón mediante el cual pueden combinarse sus elementos. Es decir, cómo pueden relacionarse entre sí las fórmulas e incluso jerarquizar sus funciones. Particularmente, la lógica y la matemática son extraordinariamente poderosos.

Ahora bien, el mecanismo puede caracterizarse aún de modo más preciso, incluso Newell lo divide en áreas o etapas.

Cuatro áreas “geográficas” de la inteligencia artificial

Primera área: programación

La programación es un tipo de mecanismo en donde se especifican, por una parte los datos de entrada y por otra la determinación de los datos a obtener así como la forma en que éstos deben presentarse. Este tipo de tarea ya sería una forma de inteligencia, en la medida que resuelve problemas, aunque claro, todavía no como la del programador humano.

Sin embargo, el probador de teoremas ya era un avance en esa dirección. En un sistema como ese, la novedad consiste en que una clase de tareas se reduce a otra; es decir, pasa de una programación determinada a la prueba de teoremas, los cuales han servido de base para dicha programación. La demostración de suficiencia, más bien que la necesidad de un mecanismo para una tarea es característica del

buen matemático decidió trabajar en la RAND Corporation. Posteriormente conoció a Simon, con el que compartió su afinidad acerca de la inteligencia artificial e hicieron varios proyectos conjuntos.

trabajo empírico en inteligencia artificial, según Newell. Vale decir, un mecanismo tiene que ser capaz de ser funcional en una tarea específica de procesamiento de datos, eso sería lo fundamental, antes que una necesidad lógica diferente a la requerida por en este cometido. Pero volviendo al ejemplo, la tarea no quedaba resuelto sólo con la demostración del teorema, de dar cuenta de la consiste del proceso; aún quedaría la extracción del programa a partir de la prueba demostrada. Aunque ello parecía complicado de realizar, Newell lo veía posible a partir de un algoritmo, con ello se conseguiría un mecanismo que siempre realice esa tarea y sin excesivo esfuerzo. Lo cual ya sería una forma de trabajo intelectual.

Segunda área: poder responder a una pregunta

“La segunda área consiste en proporcionar la respuesta a una pregunta sobre una creencia mantenida por el programa” (Newell, 1980 p. 10). En la explicación de esta segunda forma el autor pone el ejemplo siguiente. En términos operacionales, si a un computador se le preguntara: “¿Es un X un Y?”. La respuesta no debe ser afirmativa o negativa, sino dar las razones para la “creencia” que pueda expresar el programa. Nótese que se usa un término comúnmente atribuido a los humanos, como es el de creencia. Pero la razón para atribuirlo a un sistema es que el dato es algo real y seguro del programa. El autor menciona al programa de Colby (1969) el cual daba una respuesta en base a una red de información que caía bajo su dominio. Obviamente, tenía que disponer de un algoritmo de programación capaz de identificar y seleccionar el dato que constituía la respuesta.

El programa también establecía ciertas funciones específicas que eran la fuente de la respuesta. Por ejemplo, afirmaciones o proposiciones como “los niños agresivos son

independientes”, o reglas condicionales como “si los niños son agresivos, sus padres lo castigan”.

A estas estructuras simbólicas se asociaban nociones numéricas consistentes. De esta manera la máquina podía crear una interacción con el mundo externo, asimilando tanto datos como preguntas, y mediante un algoritmo generar inferencias de todo su dominio y ampliar, cada vez más, el número de respuestas así como las justificaciones de tales respuestas. Nuevamente, aquí se evidencia una tarea intelectual, ya que se manifiesta la construcción de un sistema de creencias y en la respuesta a preguntas sobre ese sistema.

Tercera área: solución de problemas

La tercera área es la solución general de problemas. Ello consiste en la capacidad de resolver cualquier problema que pueda expresarse en lenguaje formal especializado con algunas pretensiones de generalidad.

Este tipo de programas ya comenzaron a aparecer a finales de los años sesenta e inclusive antes. Uno de ellos era el General Problem Solver⁵ (GPS) en el que participó el propio Newell. Una primera forma del programa ya estaba preparada en 1959, pero se fueron perfeccionando en los años siguientes. Pero el programa que resalta Newell es el de Fikes (1969) llamado REF-ARF, cuyo lenguaje de programación ya tenía un procedimiento algebraico, que lo convierte en un sistema de formulación de enunciado-de-problema, ya que incluye funciones en forma de condiciones. Ello hacía que el dispositivo terminara de revisar toda la información de la red, hiciera la comprobación de las fórmulas y revisara todas las condiciones de satisfacibilidad de tipo lógico. Obviamente, para ello se requiere traducir un problema al lenguaje formal del REF-ARF, y cuando lo consiga el sistema podrá operar eficazmente.

⁵ El GPS fue creado por Herbert Simon, J.C. Shaw, y Allen Newell. El propósito era construir una máquina capaz de resolver problemas de carácter general. Cualquier problema simbólico formal puede ser resuelto, en principio, por el GPS. Por ejemplo: probar teoremas,

resolver problemas geométricos, trabajar con lógica proposicional y jugar al ajedrez. Algunas de las reglas básicas de solución son éstas: 1) transformar un objeto en otro, 2) reducir la diferencia entre dos objetos, y 3) aplicar un operador a un objeto.

Es decir, podrá responder después de hacer una revisión total de su base de datos. Hará una búsqueda y realizará una selección de acuerdo con el algoritmo empleado. En todo ello había un sistema de restricciones, de relaciones, pero además un proceso lógico en la codificación de la información que iba constituir el problema. De ahí la importancia que hizo el autor al decir que un mecanismo también es una formalidad abstracta capaz de realizarse en una máquina. Esa regularidad de relaciones o indicaciones precisas consistía en un proceso lógico. Ya que implica técnicas de satisfacibilidad lógica. Un caso podría ser, por ejemplo, un sistema que únicamente sume y reste y no permita la multiplicación, además de que no opere con enteros.

Cuarta área: coordinación

La cuarta área de la inteligencia artificial “comprende la coordinación de sistemas ojo-y-mano” Newell, 1980; p.15). Ello consiste en construir un sistema de organizar la información o los datos en bloques. En este caso se trata prácticamente de un robot. El salto en esta área es que los bloques son reales, y los objetos a seleccionar tridimensionales y sólidos. Para ello el computador tenía que poseer una especie de sistema sensorial, un ojo, una cámara de televisor y efector físico, como una pinza, para poder clasificar los objetos. Ello, obviamente, implica las tres áreas anteriores de la inteligencia artificial.

En ello hay una representación formal, simbólica y discreta. Pero dicha representación formal y su “sensibilidad” es construida por el sistema mismo precisamente cuando se presenta la dificultad frente a la tarea. La importancia de esto es que la coordinación genera un proceso que no estaba en el diseño original, en la programación. Más bien, este tipo de conducta lo genera, lo crea el sistema cuando se presenta la dificultad. Lo que sí está programado es el mecanismo que hace posible esta especie de “nueva conducta”. Evidentemente, este es uno de los avances de gran trascendencia en el campo de la inteligencia artificial.

Estas cuatro áreas proporcionan alguna descripción “geográfica” general de lo que era la inteligencia artificial en los setenta. Cabe indicar que muchas de estas propuestas todavía eran teorías lógicas no concretas en el tiempo que Newell escribió su obra, pero conceptualmente ya lo había teorizado de manera muy clara.

Los críticos de la inteligencia artificial dirían quizá, una cosa es el mecanismo de procesamiento de datos pero la inteligencia, propiamente, es diferente y sólo pertenece a los humanos. Newell está consciente de ello, de que existen tareas como componer una canción, pensar y escribir una obra de teatro, que poseen aspectos diferentes de los puramente intelectuales, al menos de las características que se ha señalado en las cuatro áreas indicadas. Sin embargo, la tesis de Newell es que cada una de esas actividades –como podrían ser escribir una novela– podría ser el núcleo desde el cual preguntar sobre cuáles son los *mecanismos* que lo realizan. Entendemos por *mecanismos* lo indicado antes, que lo podría constituir cualquier proceso físico que sea determinante para un resultado o también un sistema formal, abstracto, que sea susceptible de realización mediante un mecanismo físico.

Métodos débiles: buscando espacios concretos

Existe, en efecto, una correlación causal entre programa y tarea. Ésta última depende de la configuración del programa en sentido exacto, el cual no puede hacer tareas para lo cual no haya sido programado. En ese sentido, uno de los aspectos del programa sería su diferenciación con el resto de su género. Sin embargo, en esa multitud de programas, cada uno con su particular y privativa configuración, hay algunas descripciones que recogen las características generales de dicho número de programas. Por poner un ejemplo simple, hoy en día los programas de Word, Excel, Power Point, entre otros de la familia Microsoft tienen tareas distintas, cada uno cumple su finalidad, pero en los sistemas de acceso, archivo, compatibilidad, no sólo hay

semejanza sino equivalencia de funciones y manejo. Ello permite establecer una norma general y la consiguiente coherencia de un método que valga para todos.

A ello habría que añadir también la reflexión que hace Churchland con respecto al papel del soporte físico. Así como las redes neuronales soportan una infinita cantidad de procesos que podrían llamarse, por así decirlo, “programas”, el computador se comporta de modo similar: “Un programa específico convierte al ordenador en una máquina de “propósitos especiales”. Y dado que una cantidad potencialmente infinita de programas diferentes, podemos hacer que el ordenador se comporte como una cantidad potencialmente infinita de máquinas de “propósitos especiales” (1992; p.154).

Aunque Newell no identificó ninguna demostración concluyente de que esas ideas generales pudieran ser recogidas en forma de métodos, afirma que la inteligencia artificial podría ser concebida como el campo dedicado al descubrimiento y recolección de un conjunto de métodos. En otras palabras, diríamos que la inteligencia artificial consiste en encontrar lo común en lo diverso. Ello hace compartir un parentesco con los análisis numéricos o la programación matemática, la cual, como indicamos al inicio es la dimensión formal determinante para este nuevo concepto y su consiguiente plasmación técnica.

Para avanzar en ello el autor propone una concepción de método como una receta o procedimiento para realizar una clase de tareas. Un método, en este sentido, tendría tres componentes:

El enunciado del problema: un enunciado en algún lenguaje que describe la situación a la que se aplica el método, y el problema al que éste puede proporcionar solución.

El procedimiento: un programa (es decir, una secuencia condicional de acciones), basado en algún repertorio de operaciones primitivas.

La justificación: las razones por las que se piensa que si satisfacen las condiciones del enunciado del problema, entonces el procedimiento debe alcanzar la solución (Churchland, 1992; pp.17-18).

A partir de ello se desarrollan varios procedimientos que podrían resumirse en descubrimiento y desarrollo. El autor los llama métodos débiles ya que muchos de ellos no se pueden demostrar, ya que no podían en su momento recibir una concreción técnica. Lo que sí se podía hacer era describirlos de modo preciso en su concepción formal, así como indicar las condiciones que deberían tener poder realizarse. De modo que las características generales del método débil podrían resumirse en estos enunciados:

- Mostrar un conjunto de métodos –es decir, los métodos de inteligencia artificial desarrollados hasta la fecha.
- Demostrar que son métodos: proporcionar enunciados de problemas, procedimientos y justificaciones.
- Demostrar cómo se pueden escribir programas efectivos de inteligencia artificial en términos de estos métodos.
- Establecer que, en cierto sentido, la parte del programa total representada por el método proporciona una fracción sustancial (cuanto más, mejor) de la potencia total del programa.
- Demostrar que los mismos métodos son utilizados en más de un programa, reduciendo así la diversidad de programas independientes en el campo.
- Demostrar que los métodos abarcan una parte sustancial del territorio total de la inteligencia artificial.

Otro problema adicional de los programas de inteligencia artificial es las justificaciones que se tiene que hacer. En primer lugar, Newell indica que “a diferencia del ejemplo de la ecuación de segundo grado, los métodos de inteligencia artificial están expresados en un lenguaje de

procedimiento, y sabemos poco sobre cómo expresar su validez” (Churchland, 1992; p.26).

Es decir, cuando un programa escrito necesitaba hacer una tarea, demostrar que el proceso era correcto era menos importante que la comprobación experimental y la subsiguiente detección y supresión de errores. Eso es importante porque el procedimiento podría “exceder” la capacidad de la programación inicial. En segundo lugar, “la mayoría de los programas de inteligencia artificial no son algoritmos completos para las tareas que se proponen solucionar (esta es una de las razones por las cuales prefiero llamarlos por otro nombre: “métodos”)” (Churchland, 1992; p.27). Incluso cuando se logran establecer algunos algoritmos, como los de generar y comprobar, no siempre serían tan prácticos como lo exige el proceso. Ello constituye una especie de paradoja, en la medida que la validez formal permite diseñar el procedimiento pero éste, a su vez, cuando está en ejecución puede comportar características que no podrían ser formuladas por el lenguaje inicial. No sería un asunto de liberación del sistema formal y hasta de contrariedad, sino al parecer, se trata del énfasis en un campo u otro.

Psicología teórica e inteligencia artificial

Por otra parte, Newell sostiene que estos modelos, o teorías (él los usa indistintamente), están presentes tanto en la inteligencia artificial como en el hombre. En ambos hay sistemas simbólicos susceptibles de ser identificados, con nombres específicos en situaciones específicas, tanto por observación como por experimento. Así, los problemas de todos los modelos comparten principios semejantes. Todos tienen que ver con alcance, adecuación a los datos básicos seleccionados, capacidad de ampliación, potencia predictiva, comunicabilidad, simplicidad, entre otros rasgos determinantes.

En este sentido, las teorías, si deben ser útiles, no son descripciones aisladas de un proceso, sino que tienen relaciones recíprocas de deducción, y son inferidas para tratar conjuntos de datos. Así, se va creando familias de teorías,

ciertas estructuras de lenguajes comunes. Y ello se extiende no sólo en el conocimiento del individuo sino de colectivos.

En caso de los humanos los ejemplos del autor aluden al concepto tan característico de la filosofía alemana, *Der Zeitgeist* (el espíritu de los tiempos) o también el “paradigma” de Kuhn, los cuales son familias de modelos relacionados o incluidos entre sí. Al inicio pueden ser distintas y estar separadas, pero poco a poco van estableciendo sus lazos comunes. Ello implicaría un tratamiento descriptivo común entre las concepciones humanas y las de los procesadores informáticos.

En cuanto a la neurofisiología, cuyo paralelo o equivalencia se buscaba, también ahí se puede hablar en diferentes lenguajes pero al final se confluye en una estructura común. Se puede hablar en términos de redes neuronales, campos de variación continua, relaciones de fase, etcétera. De manera que Newell paulatinamente desarrolla su concepto de conciencia buscando los mecanismos comunes en la conducta humana y en la máquina. Y de modo tácito va demostrando por qué es propio hablar de un tipo de inteligencia no humana. El principio es el siguiente, si a un tipo de conducta lo llamamos inteligente, en razón de ciertas características, entonces también podemos llamar inteligente a cualquier proceso que reúna dichas características.

Cabe notar que la psicología teórica también compartía estas explicaciones sobre la naturaleza de la conducta humana. La psicología tenía al menos seis descripciones básicas que Newell encuentra concordante con los procesos que él llama inteligencia artificial:

- La teoría $S \rightarrow R$, que representa distintos conjuntos de estímulos y respuestas, que varían según la fuerza de conexión y las probabilidades de respuesta.
- Teorías estocásticas del aprendizaje, en donde el estado del organismo está representado por medio de las probabilidades de respuesta, con operadores sobre esas probabilidades en forma de leyes.

- Modelos de Markov, los cuales consideran al organismo en un estado discreto, con leyes que se establecen en forma de una matriz de transición de probabilidades entre estados.
- Conjunto de fuerzas, considera a la persona como un campo vital, en donde hay donde varios impulsos relacionados.
- Comunidad interactuante de subagentes, que incluye al Id, al Ego, y el Súper Ego de Freud, descritos en forma de personalidad.
- Procesador de información, que no sólo recibe datos sino que genera nueva información.

En efecto, todos estos lenguajes o teorías no son correctos o incorrectos en el sentido que se puedan probar mediante pruebas empíricas, sino son generadores de teorías, las cuales tienen “un aire de familia”, por lo que comparten aciertos y debilidades. Las teorías que este sistema genere serán útiles o no (se evita usar el término válido o correcto) en la medida en la medida que puedan enfrentarse exitosamente con los problemas que pretendan explicar. Por lo demás, su referencia a lo biológico obedece al hecho de que en inteligencia artificial se toma al hombre como procesador de información, el cual también usa un sistema de símbolos discretos.

Sin embargo, estas teorías, referidas al hombre, no abarcan la totalidad de los fenómenos, si bien describen procesos sustanciales no pueden describir la complejidad de los asuntos humanos. Pero el comprender al hombre como procesador de información no deviene únicamente de la inteligencia artificial clásica, sino también de la psicología en general, de los sub campos: solución de problemas, psicolingüística, formación de conceptos, aprendizaje verbal, memoria inmediata, percepción, entre otros. En ese sentido, la conducta humana podría entenderse como la consecuencia de un sistema compuesto de memorias que contienen símbolos discretos e impresiones simbólicas, conjuntamente con procesos que manipulan estos símbolos. Entendido el símbolo como una entidad con cierta propiedad funcional, que tiene información

codificada y, a su vez, también se pueden establecer familias de relaciones simbólicas.

De esta manera, el hombre procesa información y soluciona problemas, la inteligencia artificial hace lo mismo y ayuda a explicar el complejo proceso de la conducta humana. Sin embargo, en tareas de mayor complejidad, como la inducción, por ejemplo, no existen teorías de cómo se llega a ese tipo de conducta abstracta. Para ello se requeriría mayor evidencia conductual como ayuda para la identificación de teorías o modelos, pero como indicamos al inicio, la disciplina informática todavía tiene pocas décadas y no se sabe qué aspectos de lo humano podrá explicarse a partir del estudio de lo que es la inteligencia.

En consecuencia, la importancia de referir a la conducta humana es que “lo que se desea es crear un modelo de procesamiento de información que se conduzca como el humano que soluciona problemas, concebido a través de las inferencias realizadas a partir de los datos verbales y de movimientos oculares” (Churchland, 1992; p.54). Ello confirma que desde sus inicios se tuvo en cuenta lo peculiar, pero al mismo tiempo se orientaban a buscar la similitud subyacente entre el hombre y la máquina.

La naturaleza general de la mente

Con respecto a la mente, muchos de los teóricos de las neurociencias, como Feigl, sostenían que la relación principal entre mente-cuerpo tenía que ver con la naturaleza lógica de las leyes que conectaban lo sensorial y los procesos neurofisiológicos. Pero, ¿qué podría decir la inteligencia artificial con respecto a la mente, de aquello que Newell consideraba saber muy poco? Si bien su centro está en la relación que existe entre los procedimientos y las tareas, en el plano de la psicología teórica, términos asumidos por Newell, las mentes son sistemas de procesamiento de información que están realizadas sobre sistemas físicos, al menos de ello se podría estar totalmente seguros, y ello va a tener consecuencias muy grandes para la filosofía de la mente.

En la misma línea va el debate de Margaret Boden con el cartesianismo sobre la preeminencia de la *res cogitans* frente a la *res extensa*: “Una formulación alternativa de la misma cuestión esencial es ésta: ¿cómo es posible que los fenómenos mentales sean a la vez irreductiblemente psicológicos y, sin embargo, totalmente dependientes de una base causal mecánica (el cerebro y el sistema nervioso)?” (1984; p. 547). Y si el soporte mecánico es comparte una estructura común con otros sistemas de procesamiento de información, sería propio llamarlo “inteligente”.

Ahora bien, la pregunta que se hace Newell es la siguiente: “¿Qué es lo que, en función de estos problemas, diferencia a una teoría de procesamiento de información sobre la actividad mental de cualquier otra teoría psicológica?” (1984; p.76). En todo caso, para el autor esta correlación abre cuestiones de naturaleza filosófica. Siguiendo a Putnam, se toma el problema de manera análoga, en el sentido que las cuestiones de la filosofía de la mente pueden ser abordadas teniendo en cuentas las máquinas de Turing, las cuales tienen tanto efectores como receptores.

Pero la equivalencia o analogía de funcionamiento no agota el tema, ya que si bien la neurofisiología es el soporte para los procesos mentales, éstos podrían tener otra naturaleza, como sería el caso de que la electrónica sea irrelevante para la naturaleza de la inteligencia artificial. “La conducta de la máquina viene especificada por el manual de programación, y esto está totalmente realizado fuera del “hardware” subyacente” (1984; p.76).

Es decir, los mecanismos o sistemas equivalentes no son el argumento completo, sino que habría otras instancias que explicarían los aspectos sustanciales. No obstante, si la inteligencia artificial produce realizaciones similares a las de la mente y, además, si la inteligencia artificial depende de algunas características generales del sistema físico que lo soporta, cabe preguntarse si lo mismo hace una programación biológica para producir sus ideas,

sus guiones de cine, sus novelas o el plan para pintar una pintura del arte clásico.

El tema es eminentemente filosófico cuando ya no sólo en el ámbito de la ficción, sino en la realización tecnológica, se generan procesos inteligentes, aunque se trate de analogías. La máquina de Turing ya sugería la temprana preocupación filosófica sobre la posibilidad del “pensar” de la máquina. Debido a la complejidad tecnológica para estas concreciones o avances la filosofía no ha retomado las implicancias teleológicas de la inteligencia artificial, pero el planteamiento continúa.

Quienes defiendan la unicidad y singularidad de la inteligencia y la mente, claro está, referido a los humanos, podrían alegar las diferencias sustanciales actuales, incluso varias décadas después de que haya comenzado la inteligencia artificial clásica. Sin embargo, la formulación teórica ya está, sólo faltaría superar los obstáculos empíricos. De hecho, Newell afirma que si se acepta la inteligencia artificial como psicología teórica, la ejecución de un programa que se propone representar alguna conducta humana, es simplemente el cálculo de las consecuencias de una teoría psicológica.

Sin embargo, desde la filosofía ha habido algunos recelos por aceptar que seres no humanos sean capaces de pensamiento. Pero en los últimos años algunos filósofos tienen una lectura más abierta al tema. En su debate con la postura de John Searle, Pascual Martínez-Freire afirma que el hecho “de que sólo los cerebros producen mente constituye una clara petición de principio, ya que contiene en sí mismo el rechazo de una máquina que piense, conclusión que pretende derivar más tarde. Además tal axioma supone un ingenuo chauvinismo humano y la creencia en la radical y substancial superioridad de la química del carbono, propia de los seres vivos, sobre la química del silicio, propia de los computadores” (2007; p.124). Por su parte, Newell considera que en la inteligencia artificial los esfuerzos están dirigidos a la búsqueda, la inducción, el descubrimiento. En cambio, en filosofía, si tenemos en cuenta la segunda mitad del siglo XX, incluso antes, la reflexión se centró en las

operaciones mentales, pero, “estos tratamientos están saturados de un análisis del uso del lenguaje, como si fuera posible utilizar el sentido lingüístico plasmado en una comunidad lingüística como fuente de datos objetivos” (1984; p.83). No obstante, ya en los tiempos actuales, la filosofía cada vez más toma en cuenta los resultados de la ciencia experimental, entre ellos los que tienen que ver con la neurobiología.

Conclusiones

Volviendo a las preguntas iniciales, ¿cuál es el real alcance de la inteligencia artificial? Si bien muchas de las predicciones sobre las máquinas pensantes no se han cumplido y están lejos de realizarse no por ello se podría decir que el abismo es insalvable entre las inteligencias de las máquinas y las de los humanos.

Hay que tener presente que si en 1956 nace esta nueva ciencia en comparación con la física, o la química, por ejemplo, la inteligencia artificial es bastante joven. Y así como para las teorías de Leibniz o Vaucanson la barrera era tecnológica las lógicas y las matemáticas actuales están creando sistemas que podrían dar otro salto en cualquier momento. Recordemos que en principio, desde el punto de vista de la inteligencia artificial cualquier sistema formal puede automatizarse. Cualquier barrera tecnológica sólo sería cuestión de escala, tiempo y dispositivos mecánicos.

Como observa Churchland: “La pregunta que debe responder el programa de investigación de IA (inteligencia artificial) no es si los ordenadores correctamente programados pueden simular la conducta permanente producida por los procedimientos computacionales que se encuentran en los animales, incluyendo los que se encuentran en el ser humano. En principio, al menos, pueden. La pregunta importante es si las actividades que constituyen la inteligencia consciente son todos *procedimientos computacionales* de alguna clase. El supuesto rector de la IA es que lo son, y su meta es hacer programas verdaderos que lo simulen” (1992; p. 157).

Bibliografía

- Boden, M.A. (1984). *Inteligencia artificial y hombre natural*. Madrid: Tecnós.
- Churchland, P.M. (1992). *Materia y conciencia*, Barcelona: Gedisa.
- Copeland, J. (1996). *Inteligencia artificial*. Madrid: Alianza.
- McCorduck, P. (1991). *Máquinas que piensan*. Madrid: Tecnós.
- Martínez-Freire, P. (2007) *La importancia del conocimiento. Filosofía y ciencias cognitivas*, 2ª edición, Netbiblo, A Coruña.
- Newell, A. (1980). *Inteligencia artificial y el concepto de mente*. *Revista Teorema*. Vol.27. 111 pp.