

La noción de inteligencia después de *Computing Machinery and Intelligence*

UNA PERSPECTIVA HISTÓRICA

PEDRO MONTOTO GARCÍA (USC)

3 DE AGOSTO DE 2015

Resumen

Este trabajo pretende estudiar la evolución del concepto de inteligencia en los grupos de Inteligencia Artificial a partir de la publicación por parte de Alan Turing de *Computing Machinery and Intelligence* en Turing 1950, las impresiones y técnicas generadas a partir de éste artículo y otros por los investigadores en esta disciplina en la década de 1960 que han tenido un impacto en la vertiente filosófica de este problema y en última instancia proponer una clasificación de los tipos de problemas que se intentan resolver en la disciplina hoy en día contrastándolos con las intenciones de los fundadores de la disciplina en la década de los 60. Se hará una recensión de los problemas que la pregunta **¿Puede pensar una máquina?** genera, de los tipos de soluciones técnicas que se dan con los problemas que éstas enfrentan, técnicos y matemáticos, y de los nuevos problemas y conclusiones filosóficas a los que nos lleva ésta. Como guía organizativa se ha usado el trabajo de Nils Nilsson¹ (Nilsson 2009) sobre historia de la IA titulado *The Quest for Artificial Intelligence: A history of ideas and achievements*.

Keywords: Turing, artificial intelligence, history, algorithms, cognitive psychology, cybernetics

¹Investigador en el *Department of Computer Science* de la Universidad de Stanford, inventor de varios algoritmos de gran importancia en redes neuronales y co-creador del algoritmo A^* , que permite descubrir el mejor camino más corto entre dos nodos de un grafo mejorando la búsqueda exhaustiva ponderada (conocida como *algoritmo de Dijkstra*) mediante el uso de métricas heurísticas.

Índice general

1	Introducción Histórica a lo Inteligente y lo Artificial	3
1.1.	¿Qué es un constructo tecnológico?	5
2	La definición de inteligencia: el nacimiento de las ciencias cognitivas	9
2.1.	Las ciencias cognitivas y el estudio de la inteligencia	11
	Bibliografía	15

Introducción Histórica a lo Inteligente y lo Artificial

Todo lo que llamamos IA comienza con una pregunta de apariencia simple, *¿Es posible construir algo que pueda pensar?*, que en realidad presenta muchas cuestiones asociadas. Esta idea no es en absoluto reciente ni mucho menos, ya que existen leyendas griegas clásicas alrededor del siglo V antes de Cristo bien conocidas de estatuas de Pigmalión traídas a la vida por Afrodita, diosa de la vida y del amor, y de Hefesto, dios de la forja, construyendo ayudantes dorados para los dioses. De todo esto nos llega noticia a través de la *Política* de Aristóteles, que crea probablemente por accidente uno de los primeros ejemplos de ciencia ficción político-social, planteando la cuestión de qué ocurriría si tuviésemos *máquinas/autómatas/seres artificiales* inteligentes:

For if every instrument could accomplish its own work, obeying or anticipating the will of others, like the statues of Daedalus, or the tripods of Hephaestus, which, says the poet, of their own accord entered the assembly of the Gods; if, in like manner, the shuttle would weave and the plectrum touch the lyre without a hand to guide them, **chief workmen would not want servants, nor masters slaves**. Here, however, another distinction must be drawn; the instruments commonly so called are instruments of production, whilst a possession is an instrument of action. (Aristotle 350 a.C.)

Más ejemplos se dan posteriormente. El Talmud, compilado entre los siglos I y VI, habla de *golems* creados de tierra que hombres santos y doctos pueden infundir de vida. En el siglo XII el catalán Ramon Llull inventa lo que él llama *Ars Magna*, un conjunto

de discos de papel que convenientemente combinados y rotados permitirían dirimir cualquier discusión teológica mediante el uso de la lógica, con la intención de convertir a las personas de fe musulmana al cristianismo mediante la misma. En algún momento entre finales del siglo XV ó principios del XVI Leonardo da Vinci crea unos esquemas para un robot-caballero que sería supuestamente capaz de sentarse, levantarse y mover los brazos manipulado, eso sí, por un humano¹.

En el siglo XVII se empieza a percibir, en el pensamiento general de la época, la renovación de la corriente filosófica humanista-racional, que se convierte en un cierto consenso social en el cual todo puede llegar a mecanizarse. Se puede decir incluso que hasta ésta época, desde la Grecia clásica, vida e inteligencia sólo podía ser algo otorgado por dioses u otros seres omnipotentes más allá del universo físico, al darse que cualquier creación humana no puede superar una mala imitación de la vida otorgada por la divinidad. Hobbes, en este mismo siglo en su *Leviathan*, contempla la posibilidad de crear un ingenio mecánico que se comporte como un animal, pues todos los órganos para él tienen paralelismos con la mecánica: “el corazón no es más que un muelle, los nervios son cuerdas” y conceptos similares aparecen a lo largo de su obra.

En el siglo XVIII Jacques de Vaucanson presenta un autómatas que es capaz de simular un pato vivo en ciertos de sus aspectos, construido mediante un armazón metálico adornado con plumas de pato y componentes de relojería y mecánica, que era capaz de comer grano, beber y “digerir”. En realidad el producto de la digestión ya estaba en el interior del pato para la simulación. Jacques de Vaucanson también es el inventor de las primeras tarjetas perforadas entendidas como contenedoras de programas, secuencias de acciones, para entes mecánicos.

También en el siglo XVIII se creó el “Turco mecánico”, un autómatas que podía jugar al ajedrez como un maestro y que como exhibición del mismo fue enviado de gira por las cortes de Europa de la época retando y ganando a soberanos y estrategas. Más

¹Éste robot se presentó supuestamente funcionando en una fiesta de la época en la corte de Venecia en 1495 organizada por Ludovico Sforza, y más recientemente un empresario llamado Mark Elling Rosheim reconstruyó los diseños de Leonardo, probando que los mismos eran sensatos y funcionales.

tarde se supo que éste “autómata” debía su genialidad a un maestro de ajedrez humano que se ponía en su interior en cada partida. Nota bene: La compañía Amazon crea en 2013 un servicio que distribuye y automatiza tareas simples realizables por humanos a voluntarios que cobran por tarea realizada. El nombre del mismo, en homenaje, es Amazon Mechanical Turk y nos permite resaltar que el uso de la inteligencia humana en tareas simples aún no automatizables es, por tanto, una solución bien conocida para las lagunas que la Inteligencia Artificial aún tiene.

A partir del siglo XIX comenzamos a ver por doquier obras de teatro, narraciones y películas que hablan del *qué ocurriría* si las máquinas pudiesen pensar o actuar como humanos². Es evidente, por tanto, que muchos de los problemas de este *posible* que plantea el nacimiento de la IA no son nuevos y sin embargo, el mayor avance que se ha dado en este campo ha sido sin duda alguna los mayores avances se han dado en el siglo XX de la mano de los avances en maquinaria computacional, psicología, neurología y la vertiente artística en la ciencia-ficción.

Las principales preguntas que se han de responder para responder *¿Es posible construir algo que pueda pensar?* son dos: *¿Qué es pensar?* y *¿Qué es construir?*. Trataremos de ofrecer una visión histórica y panorámica de ambas cuestiones.

1.1. ¿Qué es un constructo tecnológico?

Para la construcción de inteligencia precisamos, por tanto, definir qué entendemos por construcción. Esto presenta un número de problemas, como veremos, que no entran por completo en el alcance de este trabajo, pero creemos que conviene reflejar el cambio de concepto que se ha dado en el último siglo, aunque los grupos de tecnólogos y científicos (ingenieros, matemáticos, biólogos, psicólogos y científicos sociales) que trabajan en Inteligencia Artificial no suelen tenerlo en cuenta.

²Como veremos, del pensar al actuar como humanos hay diferencias que además, se han ido acen-
tuando con el progreso en IA.

Normalmente definimos constructo o artefacto³ como una entidad, que dejaremos como definición genérica pues el *problema del ser* está completamente fuera del alcance del trabajo, en la que algunas o todas sus propiedades preexisten en la intencionalidad de un autor. Es un artefacto si éstas propiedades existen en la descripción de la intencionalidad del autor y/o son aceptadas como válidas en la descripción de lo construido por el autor, i.e. el autor determina unas propiedades y valida que el constructo para el que éste ha guiado o accionado el proceso son las que se espera del mismo. Ya desde los tiempos de Aristóteles, en su *Física* se intuye esta definición al decir que los “productos naturales” se generan por sus propios impulsos internos mientras que los “productos artificiales” precisan de una intencionalidad humana. Avicena criticaba en la edad media que la alquimia jamás podría conseguir “sustancias genuínas” como las presentes en la naturaleza precisamente por ser un constructo con intencionalidad humana. En este caso, para mantener el debate cerrado en torno a lo que nos proponemos definir, i.e. lo que es un constructo, asumiremos que no existe autor en los entes naturales y que son autogenerados en cierta medida por sí mismos o sus predecesores en el tiempo.

Podemos criticar que la “vida natural” no se diferenciaría en absoluto de una supuesta “vida artificial” pues la intencionalidad del autor al crear dicha vida natural, esto es, las propiedades de dicha forma de vida, debería poseer características en común con la vida “artificial”, habiendo sólo cambios en el proceso y los materiales fuente si es una imitación perfecta. Avances en biología, química e Inteligencia Artificial nos llevan en última instancia que lo “artificial” podría diferir sólo de lo “natural” en una cuestión de proceso y al no ser posible distinguir lo artificial de lo natural ontológicamente sin analizar los pasos previos que se han dado para obtenerlo podría convertir esta distinción en vacua: i.e. si no podemos distinguir sin conocer previamente el proceso, si llegamos a conocimiento del proceso posteriormente no podemos distinguir el proceso “natural” del proceso “artificial”. Las cualidades del objeto que lo hacen “natural” o “artificial” no son inherentes al objeto una vez la técnica de imitación

³Para una referencia básica en la filosofía de la tecnología es recomendable consultar Franssen, Lokhorst y Poel 2013 y Hilpinen 2011.

ha avanzado lo suficiente, como creemos que es la dirección que se está tomando actualmente. Este hecho será de especial relevancia cuando comentemos el juego de la imitación de Turing.

En el contexto de Inteligencia Artificial esto nos lleva, más adelante, al si existe distinción ontológica entre la inteligencia creada artificialmente y la inteligencia natural. Es decir, si logramos crear inteligencia, sea ésta lo que sea, ¿existe aún alguna diferencia entre lo que esta inteligencia es y la inteligencia que nosotros pretendemos? Por ello entramos en una distinción similar a la que propone John Searle entre “IA débil”, significando “la IA sólo imita la acción de una mente real”, e “IA fuerte”, significando “la IA *es* una mente real”. En el argumento que comentamos más arriba las consecuencias son similares, aunque la distinción se haga irrelevante. No debe entenderse este argumento como crítica a Searle pues en el argumento de Searle, que habla sobre el proceso inteligente como constructo versus el proceso inteligente como ente natural, ambos constructos son distinguibles por sus propiedades y materiales internos, aunque no los observables externamente, por lo que sí existe una posible distinción razonable.

En última instancia decir que todas las referencias que se han podido compilar para este trabajo no adoptan una postura respecto a las bases filosóficas de la distinción natural-artificial o de la teleología de los componentes de la inteligencia. La mayoría de la literatura se limita a presentar diferentes modelos matemáticos que suelen tomar como base la máquina de Turing, las redes neuronales artificiales, aunque otros llegan a argüir que es necesaria una imitación matemática de los procesos biológicos involucrados en las emociones, normalmente se trata de una lista de tipos de procesos matemáticos, es decir, herramientas para la imitación de la inteligencia según la definición de ésta que se dé.

La definición de inteligencia: el nacimiento de las ciencias cognitivas

La definición de lo que es inteligencia es naturalmente un apartado importante, aunque de difícil solución en el campo de la IA¹. En primer lugar, trataremos de diferenciar la propiedad “inteligencia” que nos parece de un platonismo peligroso al declarar que ésta existe como entidad, pero no es designable lingüísticamente ni en el universo físico, y reemplazarla por “comportamientos inteligentes” o “tareas para las que es necesaria inteligencia”. Se predica normalmente que ser inteligente es un atributo humano, aunque si tomamos ciertas definiciones de inteligencia como válidas podemos atribuir inteligencia a todos los seres vivos o incluso a seres carentes de vida como piedras, que, intuitivamente no parecen tener ninguno de los atributos necesarios para la inteligencia. Es por tanto un tema que merece ser tratado con mucho cuidado.

Tal y como se dice en Ashby 1956: “La cibernética² es a la máquina real (electrónica, mecánica, neural o económica) lo que la geometría es a los objetos materiales de nuestro espacio terrestre” y más adelante “La cibernética es entonces indiferente al reproche de que algunas de las máquinas que estudia no están incluidas en las que encontramos entre nosotros”. Dejando a un lado que Ashby asume que el comportamiento inteligente es claramente mecánico, sacamos de aquí la idea de que la Inteligencia Artificial, en primer lugar, no escapa a la creación de comportamientos inteligentes fuera del ámbito humano pues tal vez no interesa la reproducción de todos los comportamientos a los que atribuiríamos inteligencia y, en segundo lugar, que comportamiento inteligente no

¹Para ver una pequeña lista de definiciones históricas de inteligencia en diferentes ramas de las ciencias cognitivas ver Legg y Hutter 2007.

²*Cibernética* era, en general, el nombre de lo que hoy referimos como Inteligencia Artificial en los años 50 y 60

es sólo predicable de seres vivos, al ser la economía la ciencia que estudia las redes de intercambio. Se entiende todo sistema complejo, desde la maquinaria formada por múltiples máquinas simples hasta las redes sociales pasando por los sistemas neuronales como una máquina compleja cuyas leyes son:

1. finitas y cogentes
2. imitables por otras máquinas
3. sólo dependientes de lo observable externamente

Es decir, en el caso de Ashby aún no se había dado el giro cognitivista a respecto de los fenómenos mentales, y Ashby trata la mente como un constructo mecánico cuyas piezas son irrelevantes y de las cuales el funcionamiento observable es lo único relevante. Un funcionalismo que será, 50 años más tarde, destituido en las ciencias cognitivas por dos elementos: en la inteligencia artificial por las técnicas aplicadas a problemas reales, ya que la investigación en la explicación de la mente y el conocimiento será otorgada a los psicólogos, y en la psicología por el cognitivismo, una corriente que identifica como clave en el comportamiento inteligente a los procesos mentales no observables directamente. Este funcionalismo se da también en la definición de Turing 1950, ya que identifica comportamiento inteligente como “comportamiento percibido como humano por un ser humano”. Creemos que esta definición es limitante, pues sólo identifica comportamientos inteligentes o pensamientos con actuaciones humanas, siendo ambas identificaciones comportamiento-pensamiento y humano-inteligente problemáticas: la identificación comportamiento-pensamiento impide a la IA dar una descripción de lo que es pensamiento más allá de “es comportamiento” y la identificación humano-inteligente impide la resolución de problemas que están más allá del proceso de pensamiento humano en formas diferentes que un simple aumento en la velocidad o precisión de los mismos.

Veremos a continuación un pequeño resumen de la segunda mitad del siglo XX que explica la convergencia de diversos grupos científicos de lingüistas, psicólogos,

neurocientíficos, economistas, filósofos y computólogos en lo que hoy se conoce como *ciencias cognitivas*.

2.1. Las ciencias cognitivas y el estudio de la inteligencia

Definimos *ciencia cognitiva* como toda ciencia que se ocupa de uno o más aspectos de los fenómenos de la cognición. Como tales, las ciencias cognitivas forman un conjunto de ciencias y al mismo tiempo un campo de investigación interdisciplinar cuyo tema central es el estudio de la cognición humana, animal y mecánica (Martínez-Freire 2005, p.20), que es a la vez la fuente, se cree, de los comportamientos inteligentes. Se distinguen dos conceptos de cognición: cognición A, aquella que se refiere a la acción de tomar en cuenta una realidad o, dicho de un modo más apropiado a la terminología científica, como recepción de información y cognición B, como el uso y manejo de dicha información. Tanto en el primer y segundo caso se hacen asunciones sobre la capacidad del ser humano de recibir o hacer uso de la información, así como la existencia en cierto grado de dicha información. Existen múltiples concepciones de la información que sirven de base a conocimientos tan diversos como la medición cuantitativa del contenido de un texto fuente (p.e. la concepción informativa de Shannon como desorden) o la fundación de la capacidad deductiva de la lógica clásica (p.e. la concepción objetivista informativa de Corcoran).

Tal y como lo expresan Allen Newell y Herbert Simon en *Human Problem Solving* (1972) el ser humano y su capacidad intelectual, y al mismo tiempo las máquinas formales de cómputo (i.e. las máquinas Turing-equivalentes) pertenecen al género, por analogía biológica, ‘sistema de procesamiento de información’. La constitución de las ciencias cognitivas como grupo de interés en las funciones cognitivas de animales, humanos y máquinas, comienza en la llamada “conferencia Dartmouth”, una conferencia de matemáticos y lógicos que pone de manifiesto el auge de una disciplina informática

llamada inteligencia artificial, cuyo objetivo es la obtención de comportamiento que denominaríamos inteligente en sistemas artificiales: a saber, máquinas sentientes, capaces de actuación o razonamiento autónomo. Nota bene: Más allá de la cognición en animales y máquinas, se comienza a sospechar que los reinos de las plantas, los hongos e incluso los virus tienen algún tipo de cognición. Por ejemplo, se ha demostrado que las plantas segregan neurotransmisores cuando se identifica una situación de estrés (i.e. aumento en la acidez del suelo o carencia de agua) y mediante esos neurotransmisores se regula su propio crecimiento ³.

Otro de los momentos importantes para este grupo de interés es el nacimiento de la psicología cognitiva, manifestado en tres hechos: la fundación en 1960 del Harvard Center for Cognitive Studies, la publicación en ese mismo año de *Plans and the structure of Behaviour* y la publicación en 1967 del primer libro de texto de esta escuela de pensamiento psicológico. Con la psicología cognitiva se recupera el mentalismo, esto es, la existencia de una vida mental interna que es al menos parcialmente ajena a la materia y se deja de lado el conductismo y el funcionalismo, como mencionamos antes. Además, el postulado de los procesos mentales que es inherente a la psicología cognitiva, asume un cierto compromiso con el computacionalismo, la existencia de procesos que podrían ser simulados o reproducidos mediante máquinas, y acerca esta escuela a los propósitos de la inteligencia artificial.

Aunque inteligencia artificial y psicología cognitiva son nucleares a las ciencias cognitivas no son, evidentemente, las únicas disciplinas que se centran en el estudio de los mecanismos de consciencia e inteligencia que se dan en animales y máquinas. Ciertas disciplinas, como la sociología cognitiva, la pedagogía y la filosofía de la mente, entre otras, trabajan también con este rumbo aportando diferentes perspectivas y técnicas de trabajo. Se pone el énfasis en que las ciencias cognitivas asumen una naturalización materialista de sus postulados, i.e. no existen entidades fuera del mundo físico, por tanto la mente, el espíritu y cualquier entidad a la que pueda ser asignada la función

³<http://www.sciencedaily.com/releases/2015/07/150729085922.htm>

creadora de inteligencia debe tener una manifestación física y ningún tipo de atributo no-físico. Por lo tanto, todas las disciplinas que toman parte en las construcciones de las ciencias cognitivas si no asumen esta doctrina metafísica, por así llamarla, deben compatibilizarla con ella⁴.

⁴Es éste el caso de la psicología, cuyo estudio de los procesos mentales obvia en muchos casos los procesos fisiológicos que los desencadenan.

Bibliografía

- Aristotle (350 a.C.). *Politics, book One*. Trad. por Benjamin Jowett. MIT. URL: <http://classics.mit.edu/Aristotle/politics.1.one.html>.
- Ashby, William Ross. (1956). *An introduction to cybernetics*. Trad. por Jorge Santos. Trad. *Introducción a la cibernética, 1976, Ediciones Nueva Visión SAIC, Argentina*. Chapman & Hall Ltd., London.
- Franssen, Maarten, Gert-Jan Lokhorst e Ibo van de Poel (2013). «Philosophy of Technology». En: *The Stanford Encyclopedia of Philosophy*. Ed. por Edward N. Zalta. Winter 2013. URL: <http://plato.stanford.edu/archives/win2013/entries/technology/>.
- Hilpinen, Risto (2011). «Artifact». En: *The Stanford Encyclopedia of Philosophy*. Ed. por Edward N. Zalta. Winter 2011. URL: <http://plato.stanford.edu/archives/win2011/entries/artifact/>.
- Legg, Shane y Marcus Hutter (2007). «A Collection of Definitions of Intelligence». En: *CoRR abs/0706.3639*. URL: <http://arxiv.org/abs/0706.3639>.
- Martínez-Freire, Pascual (2005). *La importancia del conocimiento: filosofía y ciencias cognitivas*. Segunda Edición. Thema, Universidad de Málaga.
- Nilsson, Nils J. (2009). *The Quest for Artificial Intelligence: A history of ideas and achievements*. 1st. New York, NY, USA: Cambridge University Press. ISBN: 0521122937, 9780521122931. URL: <http://ai.stanford.edu/~nilsson/QAI/qai.pdf>.
- Turing, A. M. (1950). «Computing Machinery and Intelligence». En: *Mind* 59, págs. 433-460. URL: <http://cogprints.org/499/>.