

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

UNA PERSPECTIVA HISTÓRICA

PEDRO MONTOTO GARCÍA (USC)

2 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
1.2.	¿Qué es inteligencia?	7
	Bibliografía	9

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.

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 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.

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

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 ambos constructos son distinguibles por sus propiedades y materiales internos, aunque no los observables externamente, por lo que sí existe una posible distinción razonable.

1.2. ¿Qué es inteligencia?

Bibliografía

Aristotle (350 a.C.). *Politics, book One*. Trad. por Benjamin Jowett. MIT. URL: <http://classics.mit.edu/Aristotle/politics.1.one.html>.

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>.

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/>.