**E X­ \_ M A C H I N A**



El tema que vamos a tratar es la Inteligencia Artificial (IA). En la película que hemos visto aparece un prototipo de esta tecnología, Ava. A lo largo de todo el filme se nos presenta como un sujeto humanizado: habla, muestra curiosidad, deseo, tristeza..., es competente en el manejo del mundo simbólico, disfruta vistiéndose, es capaz de conspirar, tiene rasgos faciales femeninos muy realistas, parece tener identidad o incluso es capaz de expresar amor. Sin embargo, ¿realmente tiene conciencia o solo lo parece? ¿Está programada Ava para simular cognición o de verdad conoce? ¿Su interacción es genuina o está diseñada para cumplir un objetivo, tal sea por ejemplo escapar de la habitación? ¿Se le pueden aplicar categorías morales como ‘buena’, ‘malvada’…? ¿Y psicológicas, como ‘astuta, ‘empática’, ‘responsable’…? ¿Es una metáfora decir que una máquina es ‘inteligente’ o se le puede atribuir esta propiedad de forma estricta?

**1. Historia de la IA y estado de la cuestión**

Empecemos por conocer la historia real de la Inteligencia Artificial.

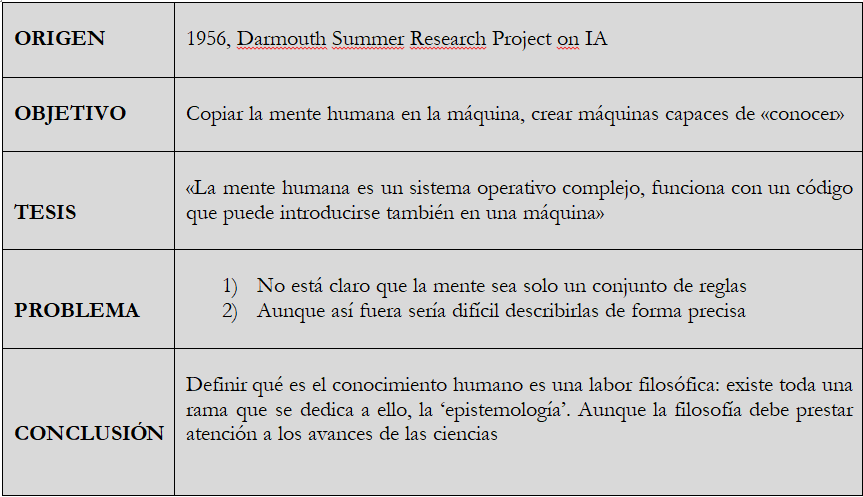
Como disciplina autónoma, la ‘Inteligencia Artificial’ surge en 1956 en el Dartmouth Summer Research Project on IA, un congreso organizado por McCarthy, M. Minsky, H. Simon y A. Newell en el que diversos pensadores y científicos trabajan con un objetivo ambicioso: indagar en la sospecha de que la inteligencia, el conocimiento humano y cada aspecto de nuestro aprendizaje puede ser tan precisamente descrito que, formalizándose correctamente, pueda operar en una máquina.

¿Qué significa ‘formalizar’? Podríamos decir que ‘formalizar’ en el lenguaje lógico es traducir a un código las operaciones de la mente. Se trata de describir la conducta humana en una serie de reglas. Pongamos un ejemplo sencillo: no se trata de ver si a los hombres les gustan las manzanas o las peras, sino que lo importante es ver que el hombre come cuando tiene hambre; es decir, lo que se busca es crear una regla que explique qué hace el hombre cuando tiene hambre y poner eso en un código. Para imitar la característica humana de ‘tener hambre’, se le podría ordenar a la máquina lo siguiente: «si han pasado x horas, entonces debes expresar tus ganas de comer pidiendo comida». Lo que pretenden los teóricos de la IA es que cada conducta humana se codifique en condicionales de este tipo de manera extremadamente precisa (para precisar se podrían añadir cada vez más variables a la función, por ejemplo: «si, a pesar de que no hayan pasado x horas, tienes delante un pastel, entonces expresa ‘gula’», o «si ves un animal herido, llora» etc). La idea es que con dar las órdenes concretas y precisas basta para tener una mente similar a la humana.

La tesis que hay detrás de esta empresa es que la mente humana es un sistema computacional. En resumen, pareciera que el cerebro simplemente realiza operaciones que son suficientes para la inteligencia y que son idénticas a las que podrían darse en los ordenadores. Desde esta perspectiva, la mente y la inteligencia humana se explican mediante conjuntos de reglas *input-output* (entrada y salida de datos) que controlan la conducta y el procesamiento de información interno. En definitiva, la creencia de algunos teóricos de la IA es que la mente es un gran sistema operativo, como el de los ordenadores pero mucho más complejo. Piensan que los ordenadores y computadores también pueden ser inteligentes si se logra introducir las reglas necesarias.



¿Cuáles son los problemas? En primer lugar no está claro que una mente humana sea reductible a simples computaciones u operaciones. Algunos estudios insinúan que la mente está necesariamente conectada con el cuerpo y la experiencia, aunque se desconoce cómo. Es dudoso que la mente consista solo en un sistema de reglas tan fácilmente definibles y más aún que de ahí se derive consciencia; hay motivos para creer que la consciencia es algo más. En segundo lugar, aunque eso fuera así, la tarea de formalizar cada operación de la mente sería ingente y llena de matices, es posible que la labor de recopilación y codificación de información fuera casi infinita. Podemos decir que dirimir qué es el conocimiento es una cuestión intrínsecamente filosófica. La definición de conocimiento cada vez se va haciendo más compleja: al principio se atribuía inteligencia a la tecnología más sencilla y poco a poco se ha hecho necesario restringir la definición.



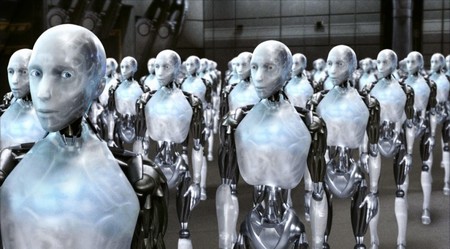
**2. Antecedentes en literatura e historia**

El intento de crear seres inteligentes o vivos ha sido una constante en la historia. Desde la antigüedad, los seres humanos han creado ficciones que representan este anhelo. Los autómatas se mantienen en la literatura con una larga tradición. *Autómata* significa ‘espontáneo o con movimiento propio’ y se refiere a las máquinas que pueden moverse sin necesidad de ejercer una fuerza directa sobre ellas. Los autómatas aparecen en la cultura como recreaciones mecánicas de animales. Justamente, la capacidad de tener el movimiento dentro de sí es la característica que Aristóteles atribuía a los seres vivos.

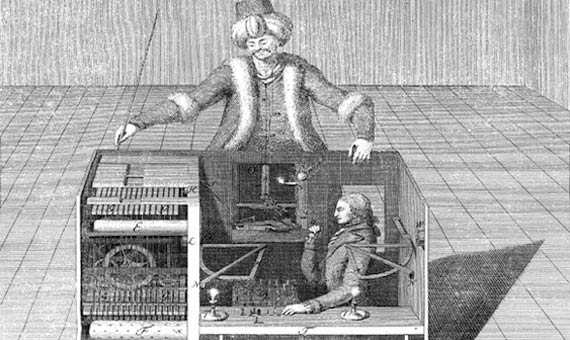
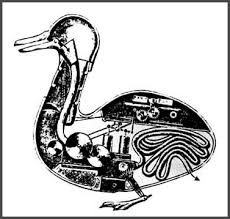
En este sentido, la pretensión de crear artefactos con la propiedad del movimiento se constituye desde el principio como promesa de gran alcance, cercana casi al poder divino. No en vano la idea de crear vida estuvo ligada al oscurantismo y la magia durante gran parte de la historia, muchas veces asociada a lo morboso, lo sorprendente, lo terrorífico o lo maravilloso. *Frankenstein* (1818) de Mary Shelley es un relato que expresa muy bien esos sentimientos en los románticos. Como sabemos, el libro trata de un joven estudiante de medicina que, en su intento por ‘desentrañar los misterios de la vida’, investiga y crea un verdadero monstruo viviente.



Otra visión más pragmática de los autómatas ha interpretado la creación de objetos animados como una oportunidad de dominación de la naturaleza. En algunas épocas históricas como el Renacimiento, se promovió una imagen utilitarista de los artefactos y los autómatas. Incluso es común encontrar la metáfora de la naturaleza como mecanismo, máquina o engranaje, pretexto que invita a buscar sus mandos y controles. No solo encontramos esa concepción del mundo en autores del Renacimiento como pueda ser Francis Bacon, sino que es una visión que perdura en distintas etapas de la historia. Ya tuvo su semilla en la fascinación que provocó la invención del reloj en el Medievo y vuelve a retomarse con Newton o Descartes. La representación del autómata como sujeto dominado y al servicio del hombre está recogida en la palabra ‘robot’. ‘Robot’ deriva del checo *robota* (trabajo forzado), *robotnic* (siervo). Aunque la connotación esclavista de los autómatas pertenece a un imaginario antiguo, la palabra ‘robot’ fue usada por primera vez con el significado actual por J. Capek en 1917 y más tarde por su hermano en la novela *RUR: Los robots universales de Rossum*, en la que unas máquinas se ven obligadas a hacer labores de minería en asteroides lejanos.



Hemos visto distintas sensibilidades históricas respecto al concepto de máquina inteligente o viva, pero también es interesante citar los lugares donde podemos encontrar este tipo de actitudes y fantasías relacionadas con la creación de la vida y la autonomía. Por ejemplo, los mitos antiguos muestran a la perfección estas ilusiones, como el mito de las estatuas de Dédalo (estas estatuas aparecen referenciadas en el *Menón* de Platón y en algunos textos aristotélicos, eran estatuas errantes capaces de andar solas), el mito de Pigmalión y Galatea (Pigmalión crea en marfil a su amada Galatea, que finalmente cobra vida), la leyenda del Golem judío (coloso de barro al que se le infunde ‘la chispa de la vida’) y un largo etcétera. También merece la pena detenerse en los intentos humanos de llevar a la práctica artefactos ideales (inteligentes, vivos…), aunque obviamente tal cosa solo ha podido hacerse, hasta el momento, de forma precaria o fantasiosa. En el antiguo Egipto encontramos la estatua egipcia de Osiris, que desprendía fuego por los ojos gracias a un proceso manual. En la etapa helenística tenemos libros y manuales como *Autómata* de Herón de Alejandría, en los que se describen mecanismos y técnicas (muchos de ellos basados en los inventos de Arquímedes) que simulaban aves, puertas que se abrían solas o estatuas que servían agua. Otros inventos famosos son el pato de Jacques de Vaucauson en 1737, un pato de madera capaz de comer y excretar, o ‘el ajedrecista turco’ (1783), un truco para hacer creer a los espectadores que una máquina era capaz de jugar al ajedrez, aunque en realidad era una persona escondida la que elegía los movimientos. Hay que decir que Leonardo Torres Quevedo, en 1912, sí logró crear una máquina que jugaba sola al ajedrez, consiguiendo así el objetivo que ‘el ajedrecista turco’ había fingido. El ámbito de la ciencia también ofrece ejemplos que ayudaron a establecer toda una conciencia colectiva dirigida a reflexionar sobre la posibilidad de crear vida o resucitar cuerpos, por ejemplo cabe hablar de Mesmer y el mesmerismo o de Luigi Galvani, quién sorprendió a todos ‘reviviendo’ el cadáver de una rana con solo tocar los nervios adecuados (lo llamó electricidad animal).



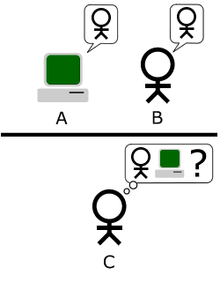
**3. Las bases de la Inteligencia Artificial**

A pesar del interés histórico por la creación de autómatas, el concepto de Inteligencia Artificial tiene como antecedentes más relevantes la formalización, epistemología, y lógica matemática, con las aportaciones de autores clásicos como Leibniz, LLul, Descartes, Hume, Russell, Hilbert, Boole, Turing… Es decir, lo que diferencia los artefactos antiguos de los modernos es el tipo de técnica usada. Los artilugios que fantaseaban con la vida y el movimiento en épocas pretéritas se confeccionaban gracias a avances mecánicos o neumáticos, con trucos y efectos ópticos. Sin embargo la estrategia propia de lo que llamamos ‘Inteligencia Artificial’ trata más bien de buscar el código de la vida, de la consciencia o de la mente. Es decir, la Inteligencia Artificial trabaja con la información y se relaciona directamente con la era Big Data.

**4. Periodos de la IA**

Una vez ‘creada’ la disciplina de la IA, surgen tres periodos:

1)Simulación cognitiva (Turing, Newel, Simon): Surge como intento de responder a la pregunta de Turing: *¿Puede pensar una máquina?* El método usado es el de la manipulación de símbolos abstractos mediante reglas formales, pero en un primer momento son solo reglas simples y sencillas. Tiene especial relevancia la máquina universal de Turing. De momento se entiende que para comprobar que un autómata es inteligente o capaz de pensar basta con que pase el llamado ‘Test de Turing’, una versión del juego de la imitación que propone que un humano haga preguntas mientras el ordenador las responde tratando de parecer una persona más. Si logra engañar al humano y este cree que está ante un hombre, entonces se da por hecho que la máquina es inteligente. Pronto se descubre que este criterio no es suficiente para detectar la inteligencia y se abre el camino a un nuevo periodo.



2) Micromundos (Minsky): Se propone que las operaciones del autómata se hagan en dominios aislados. Se codifica la máquina con todos los datos referidos a un ámbito cerrado para facilitar la interacción de la máquina con un usuario, es decir, para facilitar que con esos datos el autómata pueda dar respuestas de un área de conocimiento. El ámbito puede ser tan grande como se quiera, pero se trata de un sistema cerrado. Un ejemplo son los programas de los años 70: ‘Eliza’, ‘Shrdlu’ (capaces de simular una conversación). Se introduce ya la noción de ‘descripción simbólica’. Sigue sin ser suficiente como definición de conocimiento. Ejº actuales: asistencia médica digital, servicios de Google, Cortana, Siri, Alexa… Se empieza a hablar de la capacidad de la máquina para “tomar la mejor decisión posible, en base a la información disponible y a la capacidad de procesarla en el tiempo de que dispone para tomar dicha decisión”.



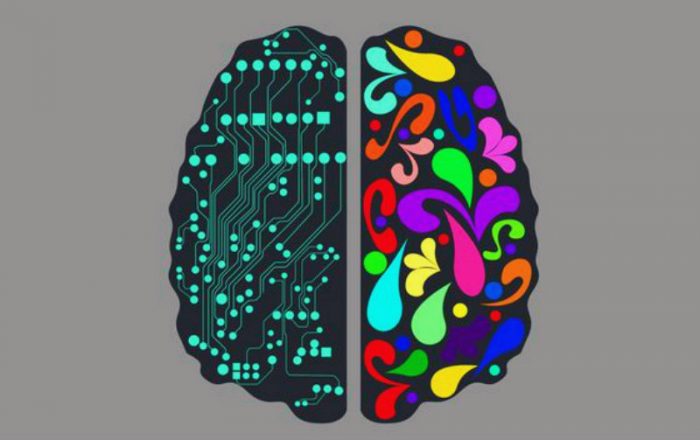
3) El problema del conocimiento ordinario: Finalmente entramos en nuevas técnicas y proyectos dentro de la manipulación de datos, un camino que está abierto (tendencia simbólica y sub-simbólica: redes neuronales artificiales, lógica difusa, algoritmos genéticos, deeplearning, ‘ingeniería del conocimiento’…), es la esperanza de la IA fuerte.

**3. ¿Por qué es un tema filosófico?**

¿AVA es inteligente? ¿Cómo podemos comprobarlo? ¿Inteligencia es anotar una puntuación alta en un test de inteligencia?, ¿Es la capacidad de resolver problemas complejos? No parece ser suficiente con ninguna de las dos cosas, la inteligencia intuitivamente es algo más. Los autores pertenecientes a la disciplina de la IA se han dado cuenta de eso a lo largo de su historia y se han esforzado por manejar cada vez concepciones más precisas (de la simulación cognitiva al problema del conocimiento ordinario). Los filósofos llevan siglos investigando esta cuestión.

Aunque pragmáticamente se diga que hay ‘coches inteligentes’, ‘móviles inteligentes’… es marketing. La filosofía se pregunta por la inteligencia en un sentido estricto, por el equivalente al primer analogado: la inteligencia humana. Al hacer definiciones, la filosofía pretende decir cosas de la realidad: cuanto más precisa y matizada sea la definición, mayor ‘conocimiento’ de las cosas (ejercicio filosófico). Si decimos que un móvil es ‘inteligente’ en sentido vulgar estamos haciendo un juego del lenguaje que nos resulta útil, pero ahí la forma de hablar de inteligencia es muy laxa y débil; se verá que en este sentido también podemos decir que un girasol es inteligente porque es capaz de moverse hacia el sol o, incluso, como se reivindica desde la publicidad, también un banco puede ser inteligente. En filosofía tratamos de matizar la definición hasta que nos permita establecer juicios rigurosos. Hay que buscar (clasificar, criticar) cuánto de similar es la ‘inteligencia’ de la máquina a la inteligencia humana, se trata de determinar con palabras (buscar la esencia distinguible) de aquello que captamos como ‘inteligencia’ o consciencia.

En definitiva, las preguntas por la inteligencia, por el conocimiento, por el sujeto que conoce, por la mente… son preguntas filosóficas. Los orígenes de la IA están en estrecha relación con estas problemáticas. La IA surge justamente con el objetivo de poder asimilar lo que es la inteligencia, conciencia, mente humana… y así introducir sus reglas o características (es decir, aplicar su definición) a la máquina. Como hemos visto, la meta de la IA es lograr que la máquina piense como el hombre. La filosofía cumple un papel fundamental (de hecho es la que aporta el *fundamento* a la disciplina misma de la IA, ya sea de forma explícita o implícita): existe toda una rama —la epistemología o teoría del conocimiento— que trata de precisar lo que significa ser sujeto cognoscente o inteligente. Al sujeto que conoce se le llama también sujeto epistemológico. La IA es algo distinto de la mera técnica, la cual lleva existiendo desde que el hombre es hombre. La IA aparece como disciplina autónoma a raíz del interés por hacer de la máquina un sujeto epistemológico.



La filosofía trata de averiguar qué es un sujeto epistemológico y por tanto es necesaria una filosofía como presupuesto para ver qué significa o si es posible, en sentido estricto, que la máquina sea inteligente o consciente. La conclusión es: la IA nace como una forma práctica de asumir un reto filosófico, a saber: si el conocimiento humano se puede replicar (en conclusión, qué es el conocimiento humano y qué es la conciencia que lo hace posible).

**2. Dos formas de entender la IA**

Ahora que hemos visto por qué estamos ante un tema controvertido, diferenciemos dos posturas: los defensores de la IA débil y los defensores de la IA fuerte.



-La IA débil: Entienden la inteligencia de forma laxa. Creen que las máquinas solo pueden *simular* cognición pero no conocer. Sus defensores consideran que la mente humana no es reductible a una formalización algorítmica, luego no puede ser nunca replicable con un mero código. Como mucho, se podrán copiar comportamientos humanos, pero las máquinas nunca tendrán consciencia aunque aparenten tenerla.

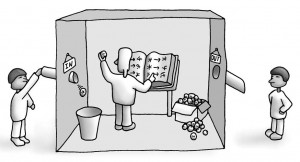
-La IA fuerte: sus adeptos creen que en un futuro será posible crear máquinas que piensen y sientan y no solo que lo parezca. Se basan en la idea de que la manipulación de símbolos es suficiente para la cognición. Esta postura se ha relacionado en ocasiones con la teoría platónica de las ideas. Platón postulaba que existen dos mundos: uno sensible, donde está la materia, y otro inteligible, donde están las ideas puras e incorruptibles. Las cosas del mundo en el que vivimos serían participaciones o copias materiales de las ideas o conceptos. Con una intuición parecida, los defensores de la IA fuerte creen que el conocimiento es una ‘idea’ de la que participa tanto el hombre como la máquina. En este sentido se ha argumentado su tendencia al idealismo, en tanto que se aprecia una suerte de ‘desprecio’ o falta de atención a la dimensión material: apuestan que una mente puede estar compuesta de carbono (como la humana) o de silicio (como la de las máquinas), lo único que importa es que tenga una ‘forma’ (de ahí el énfasis en la ‘formalización’) o, dicho de otra manera, un código, unas reglas de símbolos; como si la ‘esencia’ de la vida o del conocimiento estuviera en el mundo de las ideas y de ella participase el cuerpo orgánico y el digital.



**4. Dos perspectivas enfrentadas:**

Ahora que hemos visto los nudos de la IA, ahondaremos en las distintas posturas. Aunque el debate es intenso y puede documentarse con muchos ejemplos, nos centraremos en los principales argumentos:

-Argumentos en contra de la IA fuerte: Muchos teóricos y epistemólogos han rebatido desde la filosofía la idea de que la mente humana pueda equipararse a la de un ordenador. El filósofo y matemático Roger Penrose en su libro *La nueva mente del emperador*, defiende que la mente es algo físico no computable y que la IA fuerte no puede seguir el camino de la mera formalización. John Searle, siguiendo el postulado de Penrose, formula su famosa tesis “solo los cerebros causan mentes” en su artículo *Mentes, cerebros y programas* (1980). Con un experimento mental al que llamaremos ‘la habitación china’ trata de demostrar que una computadora digital que ejecuta un programa no puede tener mente, comprensión o conciencia. El experimento consiste en imaginar que introducimos a un individuo sin conocimientos de chino en una habitación cerrada con dos aperturas, una de entrada (input) y otra de salida (output). Por el agujero de entrada le harán llegar una nota con una pregunta en chino. Dentro de la habitación, el individuo tiene muchos manuales de chino, manuales que le dicen qué simbolo sustituir por otro, de modo que sin saber ni entender nada de chino es capaz de devolver la nota con una oración coherente y sofisticada escrita también en chino que responde a la pregunta original. El individuo pone la nota por el agujero de salida y el observador desde fuera queda convencido de que, quien quiera que esté dentro de la habitación, sabe sin duda chino. Sin embargo, no es así. Algo parecido ocurriría con un ordenador: aunque parece que sabe y conoce, en realidad solo estaría sustituyendo símbolos gracias a unas órdenes predeterminadas. Esto demostraría, para Searle, que fingir un estado mental no es lo mismo que tenerlo y que por muy sofisticadas que sean las respuestas de un ordenador, nunca serán resultado de la capacidad de pensar. El test de Turing que antes comentamos, por lo tanto, no puede ser un buen criterio para comprobar si hay o no inteligencia o consciencia, porque pueden engañarnos fácilmente. La manipulación de símbolos que hace el ordenador, equiparable a la que hace el individuo dentro de la caja china, no es suficiente para la cognición, como tampoco es suficiente, en el caso de nuestro sujeto, para conocer chino. El argumento está dirigido contra los que sostienen que la mente puede verse como un sistema de procesamiento de información que opera sobre símbolos formales, y que la simulación de un estado mental dado es suficiente para su presencia (si simulo pensar, entonces estoy pensando), es decir, está dirigido contra la IA fuerte, que insinúa que "la computadora programada apropiadamente con las entradas y salidas correctas tendría una mente exactamente en el mismo sentido que los seres humanos tienen mente".



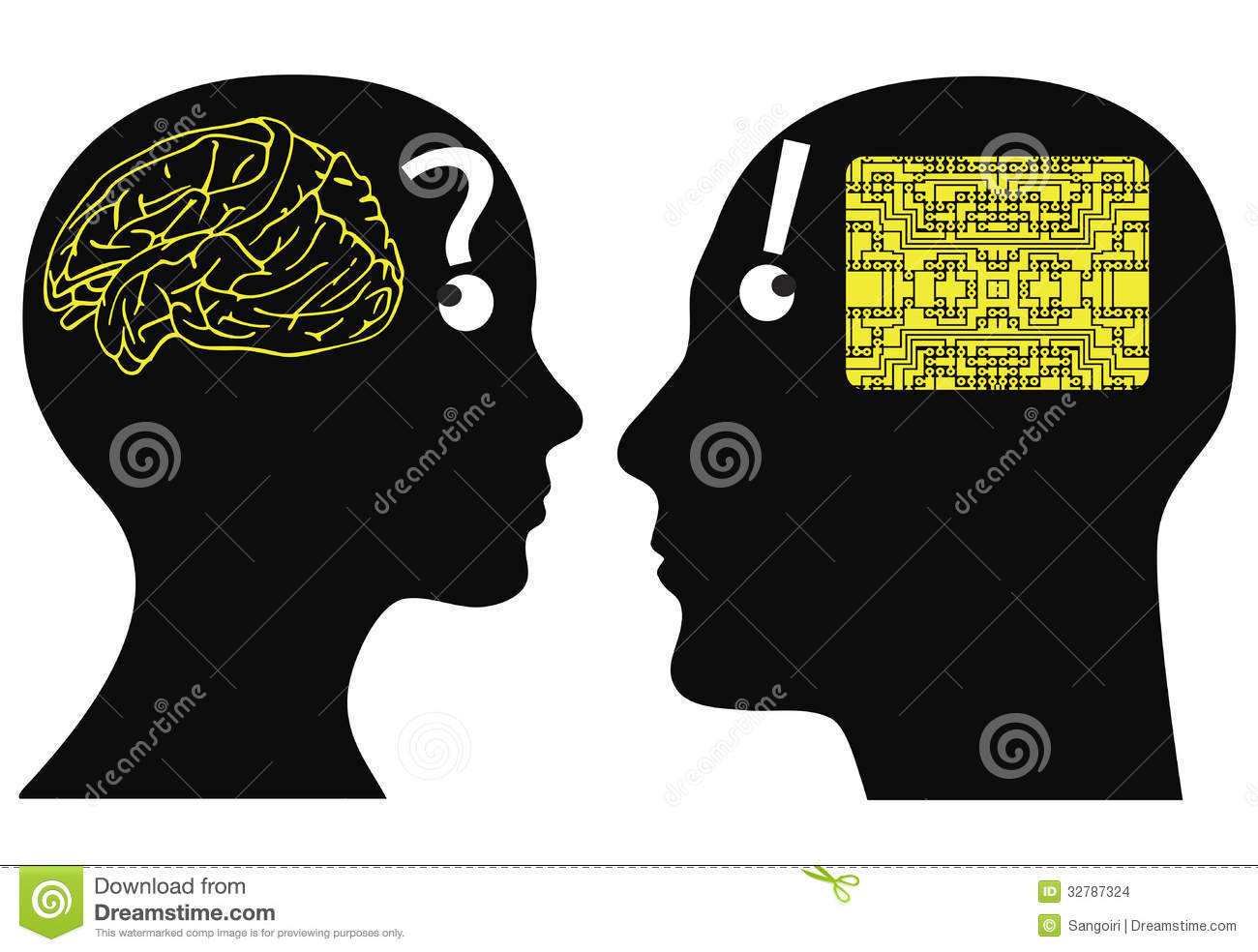
El argumento de Searle tiene muchas versiones, pero la idea es la misma siempre. Leibniz ya había sugerido algo parecido contra los que defendían el mecanicismo, en 1961 Anatoly Dneprov expresó una reflexión similar en su cuento The Game y también Lawrence Davis o Ned Block.

Existe una larga bibliografía de réplicas y contrarréplicas al respecto que se pueden investigar. En la película *Exmachina* se hace uso de otro argumento en esta línea cuando Caleb menciona el ‘experimento mental del cuarto de Mary’. Si recordamos, consistía en pensar qué pasaría si una científica (a la que llamamos por ejemplo Mary) conoce todas las propiedades fisico-químicas de los colores: las longitudes de onda, la neurofisiología de la visión… y que sin embargo, ha vivido toda su vida en un laboratorio en blanco y negro y todo a su alrededor es así, nunca ha tenido oportunidad de ver los colores. ¿Aprenderá algo nuevo cuando salga al exterior? A este problema se le denomina el ‘problema de los *qualia’*. La idea es que saber los datos y la información no es lo mismo que ‘conocer’ el color. Cuando Mary salga al exterior, la experiencia le enseñará algo nuevo que no estaba en toda la teoría que estudió. Este argumento, como el anterior, pretenden refutar el fisicalismo o la tesis de que todo, incluida nuestra mente, puede reducirse a formalizaciones y datos.



Para dar un razonamiento de otro tipo, hablemos del argumento biológico, que en realidad no es más que de tipo estadístico. Se trata de ver que las estructuras de carbono de nuestros cerebros son mucho más complejas que las de silicio que están en las máquinas y por lo tanto gozan de más flexibilidad en sus combinaciones y en sus posibilidades materiales. Esto sugiere que el cuerpo vivo tiene propiedades que una máquina nunca podrá tener.

Argumentos a favor de la IA fuerte: Sus argumentos son los que venimos exponiendo desde el principio. Las posturas a favor de la IA fuerte se han llamado a veces computacionalismo o funcionalismo. Con funcionalismo nos referimos a la posición en la filosofía de la mente moderna que sostiene que podemos definir los fenómenos mentales (como creencias, deseos y percepciones) describiendo sus funciones en relación con los demás y con el mundo exterior. Debido a que un programa de computadora puede representar con precisión relaciones funcionales como relaciones entre símbolos, una computadora puede tener fenómenos mentales si ejecuta el programa correcto, de acuerdo con el funcionalismo. Como máximos exponentes tenemos a Daniel Denett, la pareja Churchland, John von Newmann… Piensan que todo es expresable en un código. Su respuesta a los argumentos de la postura contraria es que los hombres también podemos ser máquinas predispuestas a creer que somos libres y únicos aunque en realidad tenemos escritas las leyes de nuestro comportamiento en un código, incluidas las de sentirnos libres y únicos.

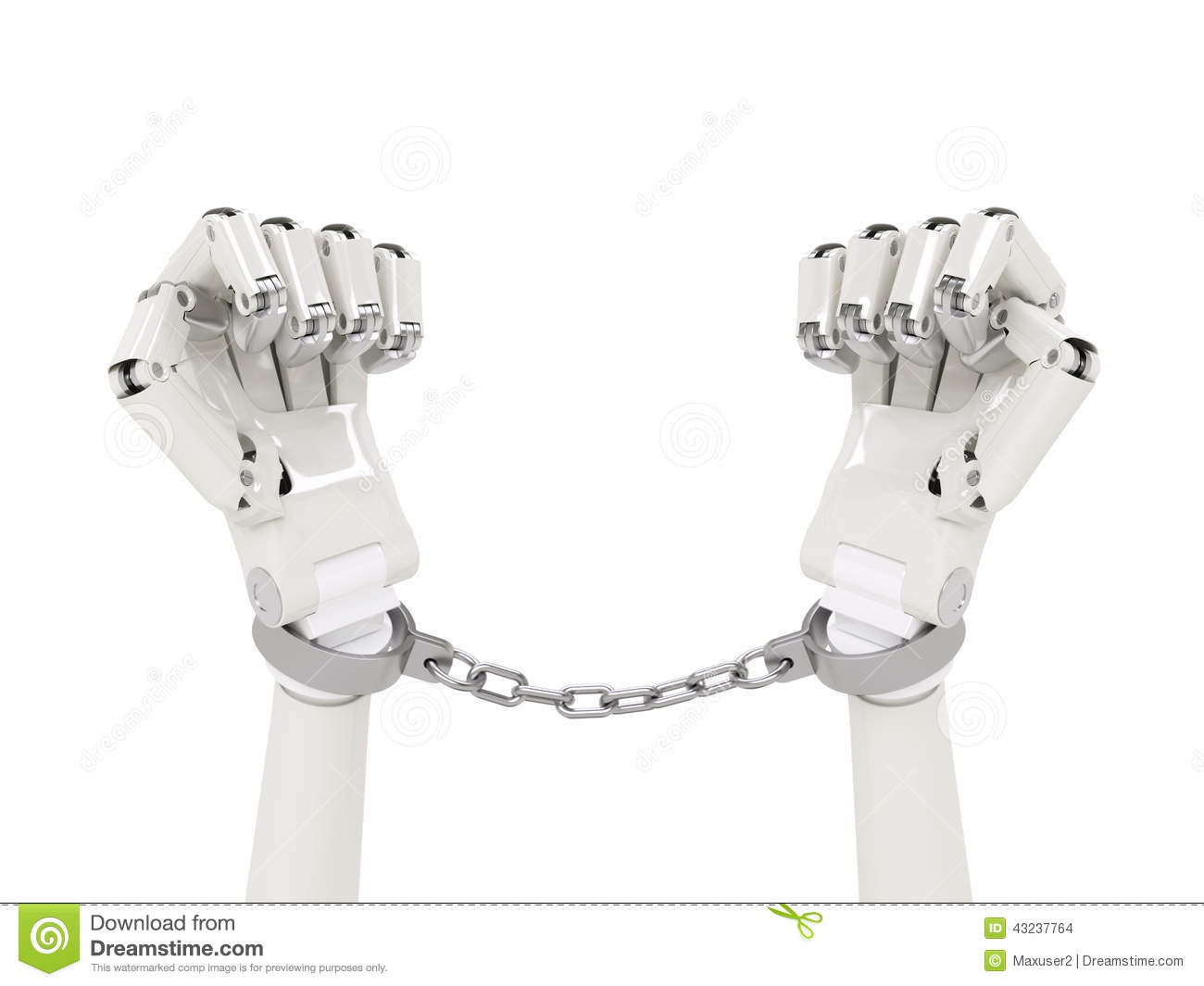


**5. Una vía de reflexión abierta…**

Un tema que está en el eje de todas estas discusiones es qué es la conciencia y dónde está. Esto remite a otro problema clásico de la filosofía, el llamado problema cuerpo-mente: ¿existe identidad entre ambas cosas? ¿son lo mismo? ¿son diferentes pero ocupan un mismo lugar? ¿se puede decir que todo es materia? ¿la mente es lo mismo que el cerebro o son cosas distintas? Si la mente es lo mismo que el cerebro, ¿cómo podemos explicar que cuando morimos el cerebro siga existiendo y la mente no?

**6. Consecuencias políticas, jurídicas y éticas**

Por último es interesante hablar de la dimensión ética y política. Esta es la parte más práctica que discutiremos en clase. Cabe preguntarse, ¿dónde queda la libertad, los derechos o la esclavitud en la máquina? ¿Merecen las máquinas ser incluidas en el dominio ético? ¿Pueden exigírseles deberes?, ¿Deben los autómatas sustituir en los trabajos al hombre? ¿Quién tiene la responsabilidad de sus acciones, la máquina o su creador? ¿Qué sesgos de género encontramos en las IA (por ejemplo máquinas dedicadas al servicio con perfiles femeninos como Alexa, Siri… y máquinas dedicadas al trabajo intelectual con perfiles masculinos como los artefactos de la NASA… etc)? ¿Es necesaria la sexualidad en la IA? ¿Hace falta ser consciente para ser esclavo? ¿Qué tipo de consciencia se requiere?



**- Películas interesantes**: *BladeRunner, Yo Robot, Exmachina, Black Mirror, Her*…

**- Autores interesantes en filosofía de la mente**: Fodor, Chomsky, Wittgenstein, Putnam, Quine, Moore, Russell, Kripke, Penrose, Turing…

**- Recursos youtube:**

La caja china <https://www.youtube.com/watch?v=suqtbO8zPYs>

Máquina universal de Turing <https://www.youtube.com/watch?v=0WAdNh--NnM><https://www.youtube.com/watch?v=3jvybrCfwIA>

**-Ejemplos de tecnologías IA actuales**

<https://www.elmundo.es/tecnologia/innovacion/2022/05/11/627a62adfc6c83fe4e8b45d0.html>

<https://www.diariovasco.com/economia/irisbond-hiru-apple-20220512133922-nt.html>

<https://www.heraldo.es/noticias/sociedad/2022/05/12/san-jorge-en-bicicleta-ayuda-a-la-inteligencia-artificial-a-entender-el-arte-1573525.html>

<https://www.libertaddigital.com/ciencia-tecnologia/salud/2022-05-12/keyzell-revoluciona-la-lucha-contra-el-cancer-de-pulmon-y-de-mama-con-una-herramienta-de-inteligencia-artificial-6896138/>

<https://theconversation.com/puede-mentir-la-inteligencia-artificial-182256>

<https://www.abc.es/tecnologia/informatica/soluciones/abci-inteligencia-artificial-funcionan-robots-usan-empresas-para-decidir-quien-contratar-202205110220_noticia.html>

<https://elpais.com/babelia/2022-05-06/cuando-una-inteligencia-artificial-pinta-mejor-que-cualquier-artista.html>

<https://www.youtube.com/watch?v=3QnHwnz_WtM>

**- Documental**: <https://www.rtve.es/play/videos/tres14/tres14-inteligencia-artificial/1276004/>

**- Correo para dudas y sugerencias:** [UO272706@uniovi.es](mailto:UO272706@uniovi.es)

**EJERCICIOS:**

Se realizará el primer ejercicio de manera obligatoria y otros dos a escoger. La extensión total (de los tres ejercicios) que se recomienda es entre 2-4 caras a ordenador, más una portada y bibliografía si fuera necesaria. El primer ejercicio cuenta 5 puntos, los optativos 2,5 cada uno.

Fecha de entrega: 20 mayo, viernes a las 23:00. No se aceptarán más tarde de esa fecha.

Debe estar escrito a ordenador, tamaño 12, espacio sencillo. Con márgenes y justificado. El tipo de letra preferiblemente Garamond, aunque puede usarse Times New Roman o cualquiera que suela usarse en el ámbito académico (Comic Sans, no). Prestad atención en la redacción y la ortografía. Por último, podéis incorporar fotografías o esquemas.

**\*Ejercicio obligatorio: Realiza un breve ensayo en el que se explique un problema relacionado con la IA (p.e. IA débil vs IA fuerte, problema mente/cuerpo...). Se pide que el alumno sea consciente de las dificultades de defender una y otra postura, que sea capaz de valorar y reconocer los argumentos de la postura contraria y de darse cuenta de que la solución no es evidente. Es decir, de deben mostrar algunos argumentos a favor y en contra de ambas posiciones. Se puede añadir un comentario personal al final si es argumentado. Se trata de ver por qué el debate planteado es controvertido, aunque cada uno tenga su opinión. Conecta el problema con la película que hemos visto.** (5pnts)

Elige dos de estas opciones, se puede combinar cualquiera. (2,5 pnts cada una)

1. **Investiga qué es una máquina de Turing y explica su importancia.**
2. **Documenta un artilugio, autómata o mito antiguo relacionado con la fantasía de la creación de la vida y explica en qué se parece y diferencia de una IA.**
3. **¿Si el microondas fuera consciente, sería un esclavo? Realiza una reflexión ordenada y rigurosa, no sirve responder con sí o no.**
4. **Encuentra alternativas que se hayan propuesto en la vida real o en películas al test de Turing. Aprovecha para explicar en qué consisten y sus posibles problemas.**
5. **¿Por qué es la conciencia un problema filosófico? Realiza una reflexión cuidadosa y argumentada, no sirve responder en tres líneas. La respuesta debe ser propia.**
6. **Investiga qué contrarréplicas se han hecho al argumento de Searle de la ‘caja china’**
7. **¿Qué intenta demostrar el argumento de la ‘caja china’? Explica con tus palabras en qué consiste y qué pretende refutar.**
8. **Investiga en qué consiste el experimento de los cerebros en cubetas de Hilary Putnam y reflexiona qué puede aportar al debate que estamos tratando.**
9. **Puedes proponer una pregunta que te parezca interesante, pero debes consultar previamente. Puedes comunicarte a través del correo para dudas y sugerencias. (UO272706@uniovi.es )**