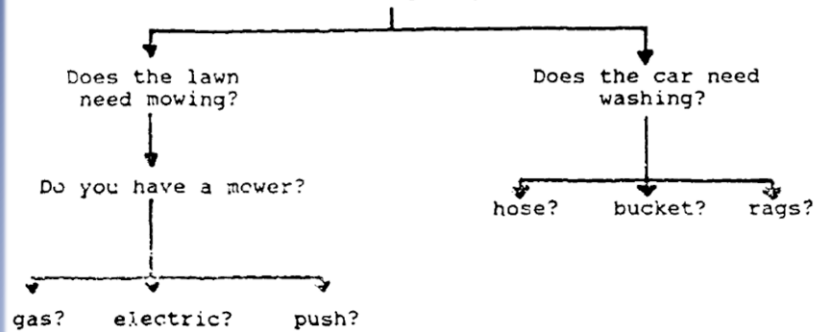


Tema 1

BACKWARD CHAINING

GOAL: Make \$20.00

RULE: If the lawn is shaggy and
the car is dirty and you mow
the lawn and wash the car,
then Dad will give you \$20.00



*** The inference engine will test each rule or ask the user for additional information.

Caracterización de la inteligencia artificial fuerte y débil

1. Introducción a la inteligencia artificial

1.1. ¿Qué es la inteligencia artificial?

La **inteligencia artificial** es un **área de conocimiento compartida** entre la ciencia teórica, la ciencia experimental y la ingeniería, por cuanto por un lado estudia modelos teóricos, por otro los prueba estudiando su evolución dadas diferentes condiciones de contorno y además crea aplicaciones de esos modelos teóricos que tienen usos en la vida cotidiana, en la industria, la medicina, etc...

No existe, curiosamente, una definición universalmente aceptada de lo que es la Inteligencia Artificial o IA o de que características ha de tener un programa inteligente -de hecho, **ni siquiera está claro a que podemos llamar inteligencia cuando salimos del esquema antropocéntrico**-.

Si que podemos hablar de diferentes definiciones de inteligencia artificial en función de la aproximación que se haga a este campo de conocimiento.

Según Russel y Norvig, la inteligencia artificial está dedicada a la construcción de programas con dos dimensiones de objetivos

- La **dualidad conducta-razonamiento**: es decir, que esos programas lleguen a un objetivo determinado -una conducta o comportamiento- sin importar el modo en el que lleguen al mismo, o bien que tengan una manera de razonar que podamos considerar inteligente.
- La **dualidad comportamiento humano - comportamiento ideal**: esto es, comparar el rendimiento del programa con el comportamiento real de las personas o con el de una inteligencia ideal.

A partir de estas dos dimensiones podemos crear diferentes definiciones de lo que es la inteligencia artificial que pasamos a desarrollar

1.2. Definiciones de inteligencia artificial

1.2.1. La conducta como inteligencia: el test de Turing

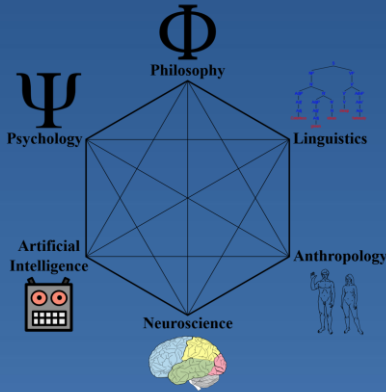
Esta definición de inteligencia es planteada por **Alan Turing**, el padre de la Ciencia de la Computación o Informática. Turing, en su artículo de 1950 "Computing Machinery" publicado en la revista "Mind", en su número 59 -pags 433-460, donde afirma que **si una máquina puede actuar como un humano, se puede considerar como inteligente**. Para ello



Alan Turing, padre de la Informática y creador del concepto de Test de Turing, que permite determinar si una máquina es o no inteligente.

El concepto de test de Turing junto con la contraposición entre el modelo de conducta como inteligencia frente a pensamiento racional como inteligencia ha sido empleado como elemento central en obras de la ciencia ficción prospectiva como "¿Sueñan los androides con ovejas eléctricas?", de Philip K. Dick, que fue llevado al cine por Ridley Scott con el título "Blade Runner" (1982), precursor del movimiento ciberpunk





La ciencia cognitiva nace de la combinación de diversas ciencias y áreas de conocimiento tales como la filosofía, la inteligencia artificial o la lingüística.

desarrolló el llamado **test de Turing**, que consiste en comunicar a un ser humano, a través de un terminal, con un ente situado en una habitación contigua, de modo que si en la interacción entre el ente y el humano este último no puede distinguir entre las respuestas de su interlocutor y las que daría una persona, si dicho interlocutor es una máquina podemos considerarla inteligente.

1.2.2. El razonamiento como inteligencia: la ciencia cognitiva

Según este enfoque no importa tanto la conducta generada por el programa como que **el razonamiento que lleva a la conducta siga patrones humanos**. Newell y Simon son los pioneros de esta línea de trabajo. Es un concepto que semeja superficialmente al de Turing pero se diferencia en algo fundamental y es que **tener un comportamiento aparentemente humano no garantiza que el mecanismo que lleva a dicho comportamiento sea semejante al pensamiento humano** – en la película “Blade Runner” adaptación de la novela “¿Sueñan los androides con ovejas eléctricas?”, dicha contraposición se teatraliza en la figura de los test de identificación de replicantes, androides con aspecto humano-. **La ciencia cognitiva trata precisamente de determinar como pensamos los humanos de cara a poder replicar los mecanismos cerebrales.**

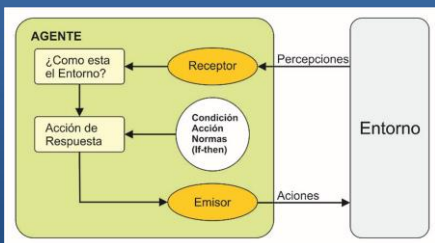
1.2.3. La racionalidad como inteligencia: la lógica

De acuerdo con este punto de vista, la inteligencia se demuestra cuando usando las leyes de la lógica se puede inferir nuevo conocimiento a partir de la información almacenada

1.2.4. El comportamiento racional como inteligencia: los agentes

Para este planteamiento, lo importante es **llegar a conseguir la conducta buscada a partir de aquellos datos que se consideran ciertos, no tanto a través de las inferencias lógicas sino también de otros procedimientos**. En este modelo cobra importancia el concepto de **agente**, el cual para ser inteligente **debe basar su comportamiento en el pensamiento racional y en las conclusiones obtenidas a partir de este, teniendo en cuenta las limitaciones de tiempo e información disponible.**

Estos agentes **colaboran** entre ellos en lo que se llama siste-



Esquema del comportamiento de un agente inteligente.

mas multiagente, los cuales pueden ser todos del mismo tipo o de tipos diferentes – por ejemplo el modelo predador – presa, que busca el equilibrio entre depredadores y presas en un nicho ecológico definido por el programador-.

2. Modelos de inteligencia artificial

2.1. Introducción

En el tema 1 del módulo de Modelos de Inteligencia Artificial se ha hablado de los cuatro tipos de inteligencia artificial que definieron Russell y Norvig en los años 90, así como de los diferentes modelos de sistemas inteligentes -símbolos físicos, modelos biológicos, etc...- sin embargo existen otras clasificaciones que nos permiten definir diferentes modelos de inteligencia artificial.

En función de los tipos de tareas que nos ayudan a resolver los modelos de IA podemos distinguir entre inteligencia artificial fuerte e inteligencia artificial débil

2.2. Inteligencia artificial fuerte

2.2.1. Concepto

La *Inteligencia artificial fuerte* o *general* nos permite resolver problemas que requieren una inteligencia humana promedio o mayor, analizando el problema y determinando el curso de acción más adecuado. Una IA fuerte puede pensar de forma racional o comunicarse en lenguajes humanos. Es un objetivo más que un modelo realmente existente – y el primer teorema de incompletitud de Godel afirma su imposibilidad-, dado que no se ha conseguido alcanzar ese nivel de inteligencia con las técnicas de IA empleadas hasta ahora.

Entre los campos de aplicación actual de la IA fuerte están los asistentes robóticos o los robots humanoides

La aplicación de la IA fuerte al hardware requeriría contar con la robótica y la mecatrónica

2.2.2. Ventajas e inconvenientes

En la actualidad aún no se ha conseguido que ningún sistema de IA se considere como IA fuerte.

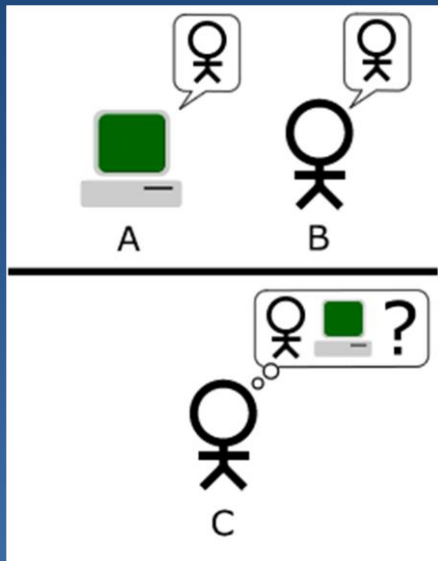
Para Turing había una serie de características propias de los seres humanos, como el humor, tener conciencia de si misma o las emociones, que no pueden ser replicadas por una máquina, si bien se ha conseguido que aprendan de la experiencia y corrijan sus errores.



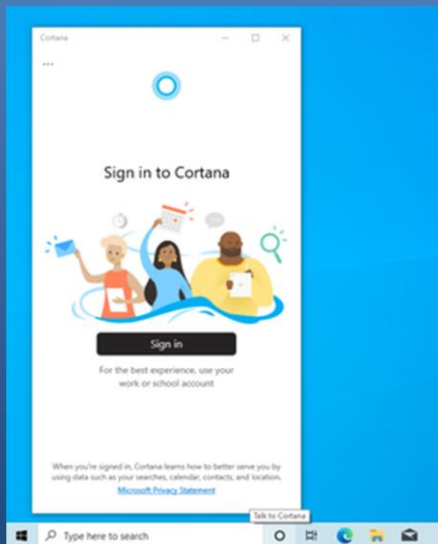
Brent Spiner, actor estadounidense conocido por su interpretación del teniente-comandante Data, un androide con una IA fuerte, en el universo Star Trek



Hal 9000, una IA fuerte consciente de si misma, cuyo comportamiento causa la muerte de la tripulación de la Discovery en la obra de Arthur C. Clarke, 2001, una odisea del espacio, llevada al cine por Stanley Kubrick en 1968



Esquema de un test de Turing, que permitiría identificar una IA como equivalente a un humano.



Cortana es un ejemplo de IA débil

La capacidad de comportarse como un humano requiere del cumplimiento de múltiples conjuntos de reglas que son demasiado complejas de implementar para un sistema software de IA.

2.3. Inteligencia artificial débil

2.3.1. Concepto

La Inteligencia artificial **débil** o **reducida** nos permite resolver **problemas sencillos**, dentro de un rango limitado. Un ejemplo sería un asistente virtual, o cualquier sistema capaz de procesar grandes cantidades de datos y tomar decisiones en **apoyo al ser humano**, automatizando procesos o realizando tareas de segmentación y clasificación de textos o imágenes,

2.3.2. Ventajas e inconvenientes. Aplicaciones

Entre las ventajas de la IA débil está su funcionalidad para trabajos físicos y cognitivos, lo que permite **sustituir la mano de obra** en muchos procesos tales como los sistemas de control de un coche, la detección de spam, los algoritmos de web semántica o los buscadores.

Su inconveniente es que estos sistemas funcionan en un ámbito concreto -los juegos de mesa, el procesamiento del lenguaje-, pero **no pueden actuar fuera de su campo de trabajo**.

2.4. Otras clasificaciones de la inteligencia artificial: inteligencia artificial clásica vs conexionista

Por otro lado podemos distinguir entre inteligencia artificial **clásica** o convencional e inteligencia artificial **conexionista** o **computacional** en función de los procesos que están detrás de las técnicas empleadas en cada caso

- La **inteligencia artificial clásica** o **convencional**, **almacena grandes cantidades de información** en memoria y a partir de las entradas de datos del exterior y la información almacenada obtiene soluciones.
- Un ejemplo serían los sistemas expertos, que permiten obtener soluciones a los problemas presentados si estos encajan con la información almacenada previamente pero no son capaces de innovar ni aprender. Emplean la lógica proposicional, así como la **estadística**.

Entre ellos están las heurísticas de búsqueda para la resolución de problemas, el razonamiento basado en casos, los sistemas expertos y las redes bayesianas.

La *inteligencia artificial conexionista* o *computacional* emplea modelos de aprendizaje interactivo imitando el funcionamiento de la naturaleza -del cerebro humano, por ejemplo-, que permite *resolver problemas más complejos en los que no hay información suficiente*. Aquí entran las técnicas de *Aprendizaje automático*, como las máquinas de vectores soporte, *las redes neuronales*, los modelos de Markov, las técnicas de regresión y clasificación, el aprendizaje supervisado y no supervisado, así como técnicas de IA propiamente dicha como los sistemas difusos y los algoritmos genéticos.

A lo largo de este manual nos centraremos en las técnicas de inteligencia artificial más clásicas, si bien al final estudiaremos también modelos de IA que suponen una transición desde el modelo convencional al conexionista, como son los algoritmos genéticos y la lógica difusa, y seguiremos con el estudio de las técnicas de procesamiento del lenguaje natural. Las técnicas propias del aprendizaje automático se estudiarán en el módulo de Sistemas de Aprendizaje Automático.