

# INSTITUTO TECNOLÓGICO DE MORELIA

INTELIGENCIA ARTIFICIAL

ENSAYO CUATRO CAPITULOS:

INTRODUCCION A LA INTELIGENCIA ARTIFICIAL

ALUMNO:

JULIO ALEJANDRO YESCAS TINOCO

13/11/2023



## Introducción

La comprensión integral de los diversos temas relacionados con la Inteligencia Artificial (IA) exige sumergirse en el desarrollo evolutivo que ha caracterizado este campo a lo largo de los años. Este proceso de evolución ha implicado la exploración y presentación de diversas perspectivas, cada una contribuyendo a la construcción de un conocimiento profundo que confiere significado a la propia noción de inteligencia artificial. Aunque comúnmente se asocia la IA con la vanguardia tecnológica actual, su conceptualización y desarrollo se remontan a la década de 1956, revelando así un linaje histórico que ha alimentado su crecimiento y expansión en el ámbito científico y tecnológico. Explorar la historia y el contexto de la inteligencia artificial no solo arroja luz sobre los avances contemporáneos, sino que también proporciona una base sólida para contextualizar y apreciar la continua evolución de esta disciplina fascinante. En este viaje a través del tiempo, se develan los hitos y contribuciones que han dado forma a la inteligencia artificial, transformándola de un concepto novedoso a un pilar fundamental en la intersección entre la capacidad humana y la innovación tecnológica.

## Desarrollo

El texto comienza destacando la falta de una definición clara de inteligencia artificial y presenta varias definiciones extraídas de libros de texto. Estas definiciones se dividen en dos categorías: aquellas que se centran en procesos mentales y razonamiento, y aquellas que se centran en la conducta. También se discute la diferencia entre enfoques centrados en humanos y enfoques centrados en la racionalidad.

Cuatro Enfoques de la Inteligencia Artificial:

el enfoque centrado en el comportamiento humano, el enfoque del modelo cognitivo, el enfoque de las "leyes del pensamiento" y el enfoque del agente racional. Estos enfoques representan diferentes perspectivas para abordar la inteligencia artificial.

1. Comportamiento Humano: Enfoque de la Prueba de Turing:
  - Presenta la Prueba de Turing como una definición operacional de inteligencia.
  - Detalla las capacidades necesarias para que una máquina supere la prueba, como el procesamiento de lenguaje natural, la representación del conocimiento y el aprendizaje automático.
2. Pensar como un Humano: Enfoque del Modelo Cognitivo:
  - Explora la necesidad de comprender cómo piensan los humanos para crear programas que imiten el pensamiento humano.
  - Destaca la ciencia cognitiva como un campo interdisciplinario que combina modelos computacionales de IA con experimentos psicológicos.
3. Pensamiento Racional: Enfoque de las "Leyes del Pensamiento":
  - Remonta el enfoque lógico de Aristóteles y la lógica formal como base para la inteligencia artificial.
  - Señala obstáculos en la transformación de conocimiento informal en notación lógica y las limitaciones prácticas de resolver problemas de manera lógica.
4. Actuar de Forma Racional: Enfoque del Agente Racional:
  - Define un agente como algo que razona y destaca las características distintivas de los agentes informáticos.
  - Describe un agente racional como aquel que actúa para lograr el mejor resultado o el mejor resultado esperado, considerando la percepción y adaptación al entorno.

También se reconoce la inevitable limitación de lograr la racionalidad perfecta en entornos complejos, lo que lleva a la consideración de la racionalidad limitada.

La revisión histórica se centra en cuestiones clave relacionadas con la filosofía, la matemática, la economía y la neurociencia.

#### Filosofía:

Se aborda la historia de la filosofía desde Aristóteles hasta la actualidad, destacando preguntas fundamentales sobre reglas formales, la generación de inteligencia mental, el origen del conocimiento y la relación entre conocimiento y acción.

#### Matemáticas:

Se presenta la evolución de la lógica formal desde los filósofos griegos hasta George Boole y Gottlob Frege. Se destaca el desarrollo de la teoría de algoritmos y la importancia de la lógica en la representación del conocimiento.

Se explora el concepto de NP-completitud y la intratabilidad de ciertos problemas, junto con la contribución de Alan Turing y su máquina para entender la computabilidad.

#### Economía:

La economía se introduce como la ciencia de estudiar cómo las personas toman decisiones para maximizar su bienestar económico. Se menciona el trabajo de Adam Smith y la teoría de juegos de von Neumann y Morgenstern.

Se destaca la teoría de la decisión y la investigación operativa en el contexto de la inteligencia artificial.

#### Neurociencia:

Se aborda la evolución del estudio del sistema neurológico, desde los primeros indicios de la conexión entre el cerebro y el pensamiento hasta los estudios más modernos.

Se menciona el papel clave de científicos como Broca, Golgi y Cajal en la identificación de áreas cerebrales específicas y la comprensión de la estructura neuronal.

En general, el resumen proporciona una visión panorámica de cómo estas disciplinas han influido en el desarrollo de la IA, desde las preguntas filosóficas fundamentales hasta los avances matemáticos, económicos y neurocientíficos.

El fragmento destaca diversas aplicaciones contemporáneas de la inteligencia artificial (IA). Aquí se presentan algunas de las capacidades actuales:

**Planificación autónoma:** El programa de la NASA Agente Remoto realiza planificación autónoma a bordo de una nave espacial, generando planes y monitoreando operaciones.

**Juegos:** Deep Blue de IBM derrotó al campeón mundial de ajedrez Garry Kasparov, demostrando capacidades de juego avanzadas.

**Control autónomo:** El sistema de visión por computadora ALVINN fue entrenado para dirigir un automóvil de forma autónoma, mostrando control de dirección en un largo trayecto.

**Diagnóstico médico:** Programas de diagnóstico médico basados en análisis probabilista han alcanzado niveles comparables a médicos expertos en algunas áreas.

**Planificación logística:** Durante la crisis del Golfo Pérsico de 1991, la herramienta DART, basada en técnicas de planificación de IA, automatizó la planificación logística de transporte para las fuerzas de Estados Unidos.

**Robótica en cirugía:** Se utilizan asistentes robóticos en operaciones de microcirugía, como el sistema HipNav que crea modelos tridimensionales y guía el implante de prótesis de cadera.

**Procesamiento de lenguaje y resolución de problemas:** El programa informático PROVERB resuelve crucigramas mejor que la mayoría de los humanos, utilizando restricciones en programas de relleno de palabras y una amplia base de datos.

Estos ejemplos ilustran cómo la IA se ha integrado en diversas áreas de la vida cotidiana, desde la exploración espacial hasta la medicina y los juegos, demostrando su impacto en la sociedad actual.

Este fragmento del texto aborda el concepto de agente racional y su aplicación en la inteligencia artificial. Aquí se destacan algunos puntos clave:

**Agente Racional:** Se introduce el concepto de agente racional como central en la perspectiva de la inteligencia artificial. Un agente racional es aquel que se comporta tan bien como puede, tomando decisiones que maximizan su medida de rendimiento.

**Agentes y su Entorno:** Se define un agente como cualquier cosa capaz de percibir su entorno mediante sensores y actuar en ese entorno utilizando actuadores. Se ilustra la idea de percepción y secuencia de percepciones en la toma de decisiones del agente.

**Medidas de Rendimiento:** Se discute la importancia de las medidas de rendimiento, que determinan el éxito en el comportamiento del agente. Estas medidas deben ser objetivas y adaptadas al entorno. Se menciona que las medidas de utilidad deben reflejar lo que se quiere lograr en el entorno.

**Racionalidad:** La racionalidad de un agente en un momento dado depende de la medida de rendimiento, el conocimiento del medio, las acciones disponibles y la secuencia de percepciones. Un agente racional toma acciones que maximizan su medida de rendimiento basándose en la información disponible.

Ejemplo de la Aspiradora: Se utiliza un ejemplo de una aspiradora como agente para ilustrar el concepto de racionalidad. Se consideran diferentes medidas de rendimiento y condiciones del entorno para evaluar la racionalidad del agente.

Este fragmento proporciona una base conceptual para entender cómo se define y evalúa la racionalidad en la inteligencia artificial, destacando la importancia de la medida de rendimiento y la adaptación al entorno.

Racionalidad vs. Omnisciencia:

Ser racional no implica ser omnisciente. Un agente racional toma decisiones basadas en su conocimiento actual, sin conocer el resultado futuro con certeza.

La perfección implica conocer el resultado real, mientras que la racionalidad se centra en maximizar el rendimiento esperado.

Ejemplo de Racionalidad en la Práctica:

Se proporciona un ejemplo de cruzar la calle donde la racionalidad se basa en la información disponible en ese momento, sin conocimiento futuro.

Recopilación de Información y Exploración:

La recopilación de información es esencial para la racionalidad. Los agentes racionales no solo recopilan información, sino que también aprenden lo máximo posible de sus percepciones.

Se menciona la exploración, destacando la importancia de modificar las acciones para recopilar información adicional cuando sea necesario.

Aprendizaje y Autonomía:

La definición de racionalidad implica que un agente racional debe aprender y adaptarse a medida que adquiere experiencia.

La autonomía es crucial. Un agente racional debe depender cada vez menos del conocimiento inicial proporcionado por su diseñador a medida que adquiere experiencia.

Especificación del Entorno de Trabajo (REAS):

Se introduce el concepto de entorno de trabajo (REAS: Rendimiento, Entorno, Actuadores, Sensores) como la especificación esencial para diseñar un agente racional.

Se utiliza el ejemplo de un taxi automático para ilustrar cómo se especifica el entorno de trabajo.

Ejemplos de Tipos de Agentes:

Se presentan ejemplos de diferentes tipos de agentes, como un sistema de diagnóstico médico, un sistema de análisis de imágenes de satélites, un robot para la selección de componentes, un controlador de una refinería y un tutor de inglés interactivo.

Cada tipo de agente tiene medidas de rendimiento, un entorno específico, actuadores y sensores particulares.

En general, la lectura proporciona una base sólida para comprender los principios fundamentales de la racionalidad en agentes inteligentes y cómo se aplican en diversos entornos.

Hipótesis de la IA Débil y Fuerte: Se mencionan dos hipótesis: la IA débil, que sugiere que las máquinas pueden actuar inteligentemente en ciertos contextos, y la IA fuerte, que afirma que las máquinas pueden tener una inteligencia general comparable a la humana.

Desafíos Filosóficos: Se discute la idea de abordar la IA desde el "computacionalismo" y se cuestiona si este enfoque puede realmente producir resultados duraderos. Se sugiere desviar los esfuerzos hacia enfoques distintos.

Definición de IA: La definición de IA se presenta como la búsqueda del mejor programa agente en una arquitectura dada. Se argumenta que, por definición, la IA es posible, ya que para cualquier arquitectura digital habrá programas que puedan ser enumerados y probados.

Cuestionamientos Filosóficos sobre el Pensamiento de las Máquinas: La discusión se centra en la pregunta "¿Pueden pensar las máquinas?" y se compara con preguntas como "¿Pueden volar las máquinas?" y "¿Pueden nadar las máquinas?" Se destaca la falta de una definición clara de lo que significa que una máquina "piense".

Test de Turing: Se menciona el Test de Turing propuesto por Alan Turing como un medio para evaluar la inteligencia de una máquina mediante la capacidad de mantener una conversación convincente. Se señala que, hasta el momento, ninguna máquina ha superado el criterio del 30 por ciento frente a jueces con conocimiento.

Argumento de Incapacidad: Se presenta el "argumento de incapacidad", que afirma que hay acciones que las máquinas nunca podrán realizar. Se discute cómo los computadores han demostrado habilidades en diversos campos, pero se reconoce que aún hay tareas desafiantes.

Objeción Matemática: Se aborda el argumento basado en el teorema de la incompletitud de Gödel y la afirmación de que las máquinas son mentalmente inferiores a los humanos. Se cuestiona la aplicabilidad de este teorema a los computadores y se destaca la inconsistencia humana.

Argumento de la Informalidad: Se discute el "argumento de la informalidad del comportamiento" propuesto por Turing y se critica la idea de que todo comportamiento

inteligente pueda capturarse mediante un conjunto de reglas lógicas. Se menciona la propuesta de Dreyfus sobre el conocimiento holístico y se señalan problemas con esta visión.

Críticas a GOF AI: Se menciona la crítica de Dreyfus a la "Good Old-Fashioned AI" (IA muy anticuada) y se discuten los problemas que él plantea, destacando cómo muchos de estos problemas han sido abordados en el diseño actual de agentes inteligentes.



## Conclusión

En conjunto, los textos proporcionados ofrecen una visión integral y reflexiva sobre la naturaleza, las posibilidades y los desafíos de la inteligencia artificial (IA). Desde perspectivas filosóficas hasta debates sobre la viabilidad de las máquinas para pensar, los argumentos exploran diversas facetas de este campo en rápido desarrollo.

Se destaca la distinción entre las hipótesis de IA débil y fuerte, abordando la capacidad de las máquinas para actuar inteligentemente en contextos específicos o para alcanzar una inteligencia general equiparable a la humana. Además, se examinan críticamente los fundamentos filosóficos de la IA, incluidos los desafíos asociados con el "computacionalismo" y la necesidad de explorar enfoques alternativos.

La definición de IA como la búsqueda del mejor programa agente en una arquitectura dada se presenta como una cuestión esencial, con el reconocimiento de que la viabilidad de la IA puede depender de cómo se defina. La pregunta central sobre si las máquinas pueden pensar se compara con cuestiones de vuelo o natación, destacando la necesidad de una definición clara y consensuada.

El Test de Turing emerge como un criterio para evaluar la inteligencia de las máquinas, y se subraya que, hasta el momento, ninguna ha superado el umbral del 30 por ciento frente a jueces con conocimiento. Se abordan objeciones, como el argumento de incapacidad y la objeción matemática basada en el teorema de la incompletitud de Gödel, desafiando las afirmaciones sobre las limitaciones fundamentales de las máquinas.

En última instancia, aunque se plantean cuestionamientos significativos y se reconocen los límites actuales, el progreso en áreas como el procesamiento neuronal, la visión activa y los enfoques de aprendizaje ofrece una perspectiva optimista sobre el potencial continuo de la IA. El diálogo filosófico y técnico en estos textos refleja la complejidad y la fascinación que rodean a la inteligencia artificial.