



LA INTELIGENCIA ARTIFICIAL COMO UNA CIENCIA DE LA COMPUTACIÓN



Yuniesky Coca Bergolla
Miguel Llivina Lavigne



LA INTELIGENCIA ARTIFICIAL COMO UNA CIENCIA DE LA COMPUTACIÓN

Yuniesky Coca Bergolla
Miguel Llivina Lavigne



EDITORIAL EDUCACIÓN CUBANA

La Habana, noviembre de 2021

Esta obra está bajo una Licencia Creative Commons
Atribución-NoComercial-CompartirIgual 4.0 Internacional.



© Yuniesky Coca Bergolla, 2021

© UNESCO, 2021

Diseño y Edición:

DI. Ismael Adán Quesada Chow

ISBN: 978-959-18-1341-1

Sello Editor EDUCACIÓN CUBANA

Dirección de Ciencia y Técnica - MINED

Calle 17 y O. Vedado. La Habana, Cuba.



ÍNDICE

INTRODUCCIÓN.....	5
LOS CUATRO ENFOQUES DE LA INTELIGENCIA ARTIFICIAL	9
■ El enfoque del comportamiento como humano	10
■ El enfoque de pensamiento como humano.....	12
■ El enfoque del pensamiento racional.....	14
■ El enfoque del comportamiento racional.....	16
ASUMIENDO UNA DEFINICIÓN.....	19
EL DESARROLLO DE AGENTES RACIONALES	25
■ Entorno de trabajo REAS.....	29
CONSIDERACIONES FINALES.....	41
GLOSARIO DE TÉRMINOS.....	43
REFERENCIAS.....	45



INTRODUCCIÓN

Este libro es una publicación científica, didáctica, metodológica y educativa, destinada a maestros, profesores, directores, metodólogos, inspectores, supervisores, técnicos, trabajadores de los distintos niveles educativos y otros profesionales de los organismos, instituciones y organizaciones gubernamentales y no gubernamentales en la República de Cuba y República Dominicana. Es el segundo de cuatro libros que tienen como objetivo principal propiciar la incorporación de conceptos e ideas generales sobre Inteligencia Artificial en las escuelas, familias y comunidades. Este segundo libro va dirigido a presentar los enfoques en que se ha trabajado y definido la Inteligencia Artificial.

Se presenta un primer capítulo donde se exponen los cuatro puntos de vista principales que se han asumido para definir el término Inteligencia Artificial. En el segundo capítulo se presentan algunos criterios sobre las temáticas específicas que estudia dentro de la Ciencia de la Computación y se presenta una definición en base a estos criterios. En el tercer capítulo se aborda la IA desde el desarrollo de agentes racionales. Este es un paradigma de programación nacido indisolublemente relacionado con la IA, de ahí que se presenten las

ideas generales que lo sustentan. Finalmente, se pone a disposición de los lectores un Glosario de Términos y la Bibliografía.

Como antecedentes de esta publicación se elaboró en la Universidad de las Ciencias Informáticas (UCI) un documento base, el cual fue distribuido a especialistas, profesores, investigadores y otros profesionales relacionados con la Inteligencia Artificial. En él se brindan las ideas generales de posibles temáticas a abordar en edades tempranas, las cuales pudieran ser formalizadas en la educación general o tratadas como temas complementarios en diversos espacios. De igual forma, como parte del IV Taller Internacional de Enseñanza de las Ciencias Informáticas en el marco de la IV Conferencia Científica Internacional de la Universidad de las Ciencias Informáticas UCIENCIA 2021, se efectuó un panel de expertos donde participaron como panelistas, reconocidos profesores e investigadores de IA en Cuba. Todos los criterios fueron tomados en cuenta para enriquecer las ideas que aquí se exponen.

Aprovechamos la ocasión para expresar nuestro agradecimiento a todos los que aportaron opiniones, valoraciones y sugerencias para mejorar esta obra. Sobre todo al grupo de investigación de Inteligencia Artificial de la Universidad de las Ciencias Informáticas, al grupo de debate sobre ética y robótica de la Oficina Regional de la UNESCO en La Habana, integrado por desarrolladores y otros especialistas de los joven club, universidades y la propia oficina. De igual forma al colectivo de la disciplina Inteligencia Computacional de la carrera Ingeniería en Ciencias Informáticas de la UCI.

Esta obra pretende servir de base para contribuciones teóricas y prácticas mediante el desarrollo de actividades educativas curriculares y no curriculares para las instituciones educativas, escuelas, universidades, familias, comunidades y la sociedad en general, así como para las instituciones y organizaciones gubernamentales y no

gubernamentales; y de todos los sectores e instituciones de la sociedad. No se trata de enseñar a desarrollar aplicaciones de Inteligencia Artificial, sino, de contribuir al desarrollo del pensamiento computacional y nociones básicas necesarias para incentivar el estudio y profundización en IA con un posicionamiento ético, crítico y creativo, desde edades tempranas. Finalmente los exhortamos a leer, estudiar y profundizar en los aspectos que se abordan, para contribuir entre todos al desarrollo de la Inteligencia Artificial como componente esencial para la informatización de la sociedad.

LOS CUATRO ENFOQUES DE LA INTELIGENCIA ARTIFICIAL

La Inteligencia Artificial (IA) es apasionante y misteriosa. Como Ciencia joven que es, reúne criterios diferentes de sus científicos, principalmente sobre sus objetivos y núcleos de conocimientos que aborda. Son varias las definiciones elaboradas del término Inteligencia Artificial, desde diversos puntos de vista. Russell y Norvig (2004) presentan definiciones extraídas de ocho libros de texto, las cuales se identifican con uno de cuatro enfoques generales (Figura 1).

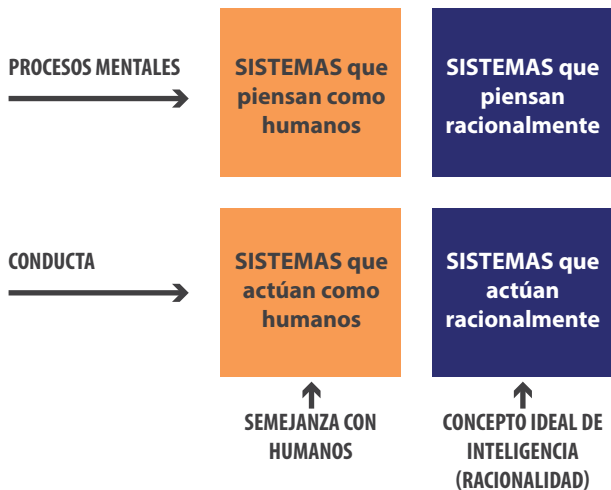


Figura 1. Enfoques en las definiciones de la Inteligencia Artificial.

Los dos enfoques que aparecen en la parte superior se refieren a procesos mentales y al razonamiento, mientras que los de la parte inferior aluden a la conducta. Los enfoques de la izquierda miden el éxito en términos de la fidelidad en la forma que se desempeñan los humanos, mientras que los de la derecha toman como referencia un concepto ideal de inteligencia, la racionalidad. Se considera que un sistema es racional si hace lo correcto, en función de su conocimiento.

A lo largo de la historia se han seguido los cuatro enfoques mencionados. Cada uno de ellos ha tenido defensores y detractores. Sobre todo porque el enfoque centrado en el comportamiento humano debe ser una ciencia empírica, que incluya hipótesis y confirmaciones mediante experimentos. Mientras, el enfoque racional implica una combinación de matemáticas e ingeniería. Sin embargo, más que ignorarse, cada uno a ayudado al otro en su propio desarrollo.

■ El enfoque del comportamiento como humano —

La **Prueba de Turing**, propuesta por Alan Turing (1950), se diseñó para proporcionar una definición operacional y satisfactoria de inteligencia. En vez de presentar una lista de cualidades necesarias para obtener inteligencia artificialmente, sugirió una prueba basada en la incapacidad de diferenciar entre entidades inteligentes indiscutibles y seres humanos. El computador supera la prueba si un evaluador humano no es capaz de distinguir si las respuestas, a una serie de preguntas planteadas, son de una persona o no. En la actualidad, podemos decir que desarrollar un sistema para que supere la prueba requiere un trabajo considerable. El sistema debería poseer las siguientes capacidades (Russell & Norvig, 2004):

- **Procesamiento de lenguaje natural** que le permita comunicarse satisfactoriamente en un idioma.
- **Representación del conocimiento** para almacenar lo que se conoce o siente.
- **Razonamiento automático** para utilizar la información almacenada para responder a preguntas y extraer nuevas conclusiones.
- **Aprendizaje automático** para adaptarse a nuevas circunstancias y para detectar y extrapolar patrones.

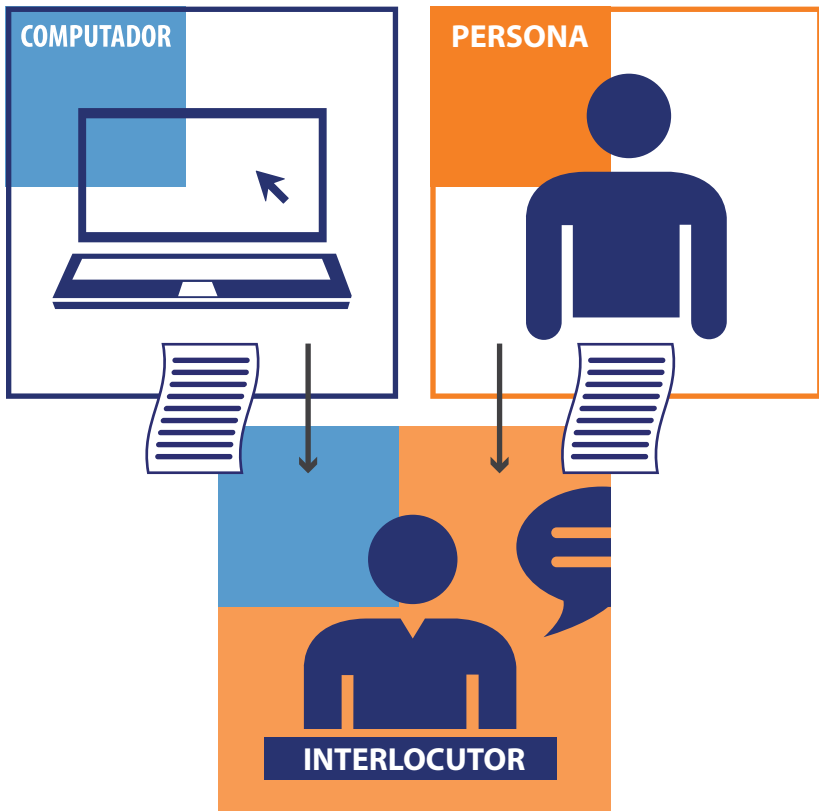


Figura 2. Prueba de Turing.

La Prueba de Turing evitó deliberadamente la interacción física directa entre el evaluador y el sistema, dado que para medir la inteligencia es innecesario simular físicamente a una persona. Sin embargo, la llamada Prueba Global de Turing incluye una señal de vídeo que permite al evaluador valorar la capacidad de percepción del evaluado, y también le da la oportunidad al evaluador de pasar objetos físicos a través de una ventana. Para superar la Prueba Global de Turing el sistema debe contar con (Russell & Norvig, 2004):

- **Visión computacional** para percibir objetos.
- **Robótica** para manipular y mover objetos.

Turing merece ser reconocido por diseñar una prueba que se conserva vigente después de más de 70 años. Muchos investigadores del campo de la IA no han dedicado esfuerzo a la evaluación de sus sistemas con la Prueba de Turing. La IA se mueve hacia aspectos más prácticos, donde es más importante el estudio de los principios en los que se basa la inteligencia que duplicar un humano.

El enfoque de pensamiento como humano

Para decir que un sistema piensa como un humano, es necesario contar con un mecanismo para determinar cómo piensan los humanos. Es necesario conocer el funcionamiento de las mentes humanas. Se podría intentar atrapar nuestros propios pensamientos o realizar ciertos experimentos psicológicos. Si en algún momento se contara con una teoría lo suficientemente precisa sobre cómo trabaja la mente, se podrá expresar esa teoría en la forma de un programa de computador. Mientras, se hace imposible modelar el pensamiento humano.

Si los datos de entrada y salida de un programa y los tiempos de reacción son similares a los de un humano, existe la evidencia de que algunos de los mecanismos del programa se pueden comparar con los que utilizan los seres humanos. Allen Newell y Herbert Simon, desarrollaron el Sistema de Resolución General de Problemas (SRGP), este programa resuelve correctamente los problemas propuestos. Pero ellos fueron más allá, les interesaba seguir la pista de las etapas del proceso de razonamiento y compararlas con las seguidas por humanos a los que se les enfrentó a los mismos problemas. Los resultados muestran semejanzas, pero muy limitadas. En el campo interdisciplinario de la ciencia cognitiva convergen modelos computacionales de IA y técnicas experimentales de psicología, intentando elaborar teorías precisas y verificables sobre el funcionamiento de la mente humana (Russell & Norvig, 2004).

En los comienzos de la IA algunos podrían argumentar que un algoritmo resolvía adecuadamente una tarea y que por tanto era un buen modelo de representación humana, o viceversa. Actualmente se reconocen diferencias entre la IA y la ciencia cognitiva, lo cual ha permitido que ambas disciplinas se desarrollen más rápidamente. Los dos campos continúan alimentándose entre sí, especialmente en las áreas de la visión y el lenguaje natural. En particular, el campo de la visión ha avanzado recientemente con la ayuda de una propuesta integrada que tiene en cuenta la evidencia neurofisiológica y los modelos computacionales (Russell & Norvig, 2004).

Lo más cercano en nuestros días a este enfoque, es el gran desarrollo que han tenido las redes neuronales artificiales (Figura 3). Algunos autores optimistas hablan de lograr en un futuro la verdadera Inteligencia Artificial mediante el estudio y simulación del cerebro humano con modelos similares a las actuales redes neuronales.

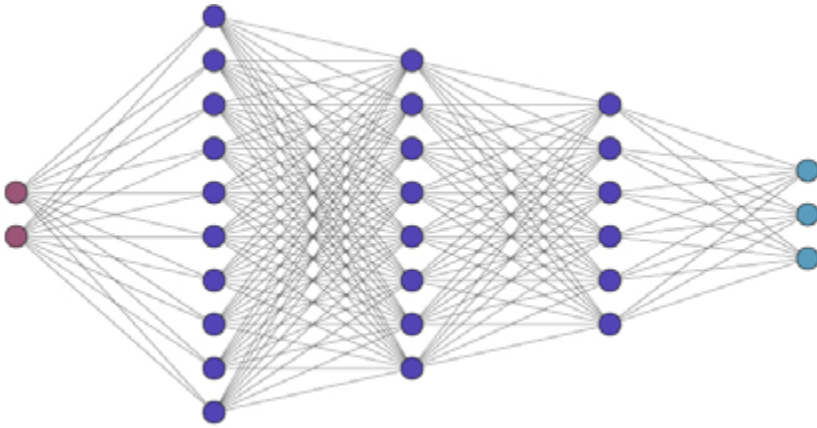


Figura 3. Modelo más cercano al pensamiento como humano: Las redes neuronales artificiales.

El enfoque del pensamiento racional

El filósofo griego Aristóteles fue uno de los primeros en intentar codificar una manera correcta de pensar, un proceso de razonamiento irrefutable. Sus silogismos son esquemas de estructuras de argumentación mediante las que siempre se llega a conclusiones correctas si se parte de premisas correctas (por ejemplo: Sócrates es un hombre; todos los hombres son mortales; por lo tanto, Sócrates es mortal). Estas leyes de pensamiento supuestamente gobiernan la manera de operar de la mente; su estudio fue el inicio de la lógica y es base de la IA actual.

Estudiosos de la lógica desarrollaron, en el siglo XIX, una notación precisa para definir sentencias sobre todo tipo de elementos del mundo y especificar relaciones entre ellos. Estas ideas son más abarcadoras que la notación aritmética común. Ya en 1965 existían

programas que, en principio, resolvían cualquier problema resoluble descrito en notación lógica. La llamada tradición logista dentro del campo de la IA trata de construir sistemas inteligentes a partir de estos programas (Figura 4).



Figura 4. La idea de sistemas con pensamiento racional.

Este enfoque presenta dos obstáculos. No es fácil transformar conocimiento informal y expresarlo en los términos formales que requieren de notación lógica, particularmente cuando el conocimiento que se tiene es incompleto. En segundo lugar, hay una gran diferencia entre poder resolver un problema en teoría y poder hacerlo en la práctica. Incluso problemas con apenas una docena de datos pueden agotar los recursos computacionales de cualquier computador a menos que cuente con alguna directiva sobre los pasos de razonamiento que hay que llevar a cabo primero. Aunque los dos obstáculos están presentes en los intentos de construir sistemas de razonamiento computacional, surgieron por primera vez en la tradición lógica.

El enfoque del comportamiento racional

En este enfoque aparece el término agente, el cual ha ganado tanto desarrollo que se considera hoy un paradigma dentro de la IA. Un agente puede considerarse como un ente que razona. Pero de los agentes informáticos se espera que tengan otros atributos que los distinguan de los programas convencionales, como que estén dotados de controles autónomos, que perciban su entorno, que persistan durante un período de tiempo prolongado, que se adapten a los cambios y que sean capaces de alcanzar objetivos diferentes. Un agente racional es aquel que actúa con la intención de alcanzar el mejor resultado o, cuando hay incertidumbre, el mejor resultado esperado.

En el caso del enfoque de la IA según las leyes del pensamiento, todo el énfasis se pone en hacer inferencias correctas. La obtención de estas inferencias correctas puede, a veces, formar parte de lo que se considera un agente racional, ya que una manera racional de actuar es llegar a la conclusión lógica. Si una acción dada permite alcanzar un objetivo, hay que llevar a cabo dicha acción. Sin embargo, el efectuar una inferencia correcta no depende siempre de la racionalidad. Existen situaciones para las que no hay nada correcto que hacer y en las que hay que tomar una decisión.

Existen también formas de actuar racionalmente que no implican realizar inferencias. Retirar la mano de un objeto caliente es un acto reflejo mucho más eficiente que una respuesta lenta llevada a cabo tras una deliberación cuidadosa.

Todas las habilidades que se necesitan en la Prueba de Turing deben permitir emprender acciones racionales. Por lo tanto, es necesario contar con la capacidad para representar el conocimiento y razonar basándonos en él. Ello permitirá alcanzar decisiones correctas en una amplia gama de situaciones. Es necesario ser capaz de ge-

nerar sentencias comprensibles en lenguaje natural. El enunciado de tales oraciones permite a los agentes desenvolverse en una sociedad compleja. El aprendizaje no se lleva a cabo por sabiduría únicamente, sino que hay que profundizar en el conocimiento de cómo funciona el mundo para facilitar la concepción de estrategias para manejarse en el entorno de mejor manera. La percepción visual es necesaria para poder tener una idea mejor de lo que una acción puede llegar a representar, por ejemplo, ver una deliciosa comida contribuirá a tomar la decisión de acercarnos a ella.

Estudiar la IA desde el enfoque del diseño de un agente racional ofrece al menos dos ventajas. Primero, es más general que el enfoque que proporcionan las leyes del pensamiento. Efectuar inferencias correctas es sólo uno de los mecanismos existentes para garantizar la racionalidad. Segundo, es más afín a la forma en la que se ha producido el avance científico, ya que la norma de la racionalidad está claramente definida y es de aplicación general. Por el contrario, la conducta humana se adapta bien a un entorno específico, y en parte, es producto de un proceso evolutivo complejo, en gran medida desconocido, que aún está lejos de llevarnos a la perfección. Por tanto, se hace más práctico centrarse en los principios generales que rigen a los agentes racionales y en los elementos necesarios para construirlos (Figura 5).

Obtener una racionalidad perfecta, es decir, hacer siempre lo correcto, no es posible en entornos complejos. La demanda computacional que esto implica es demasiado grande, pero se reconoce que es un buen punto de partida para el análisis. Esta posición simplifica el problema y proporciona el escenario base adecuado sobre el que se asientan las bases de la IA.

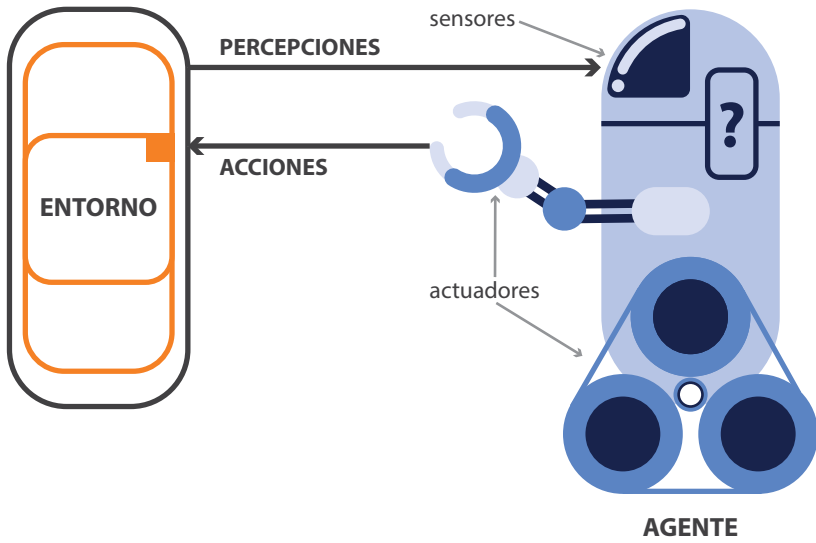


Figura 5. Los agentes racionales pueden ir desde sistemas computacionales básicos hasta sofisticados robots.



ASUMIENDO UNA DEFINICIÓN

Más que la existencia de una definición formal y acabada de qué se entiende por Inteligencia Artificial, lo importante es conocer que la misma es una disciplina que proporciona métodos para resolver problemas para los cuales el enfoque algorítmico tradicional de la computación no es suficiente. La insuficiencia puede estar dada por dos razones fundamentales (Bello, García, García, & Reynoso, 2002). Primero la carencia de un algoritmo para resolver el problema. Por ejemplo, diagnosticar una enfermedad, jugar ajedrez, elaborar el plan de acción de un robot, etc. Segundo, aunque exista un algoritmo para resolver el problema, cuando la dimensión del mismo crece, ya ese algoritmo no es computacionalmente aplicable. Por ejemplo, el problema del viajero vendedor.

Con los elementos vistos hasta el momento, podemos, a modo de debate, proponer algunos criterios que formalizan una definición de la Inteligencia Artificial como ciencia derivada de la Ciencia de la Computación. Primeramente, identificamos los núcleos básicos de conocimiento de la IA. Podemos afirmar que la representación del conocimiento juega un papel esencial en formalizar los problemas de la IA. El razonamiento, desde las búsquedas e inferencias, garantiza encontrar las soluciones racionales a partir del conocimiento

del agente. La incertidumbre es inherente al tipo de problemas que resuelve la IA. El aprendizaje, por su parte, es el nivel superior de inteligencia, desde el punto de vista que la hemos abordado. De ahí que identifiquemos cuatro núcleos básicos de conocimiento de la IA: representación del conocimiento, razonamiento, tratamiento de la incertidumbre y aprendizaje.

Un acercamiento a la IA desde esta perspectiva (Coca, 2021) la define como la ciencia de la computación encargada de aplicar métodos de representación, procesamiento y extracción de conocimiento, mediante programación multiparadigma, en el desarrollo de sistemas informáticos con comportamiento racional. Especificando en la definición los núcleos de conocimiento antes expuestos, podemos afirmar que:

la Inteligencia Artificial es la ciencia de la computación encargada de aplicar métodos de representación del conocimiento, razonamiento, tratamiento de la incertidumbre y aprendizaje, en la concepción de sistemas informáticos con comportamiento racional.

Se puede identificar como **objeto de estudio**: El desarrollo de sistemas con comportamiento racional. Los cuatro núcleos de conocimiento presentados pueden considerarse como categorías que incluyen tareas generales sobre las que los científicos desarrollan sus investigaciones de manera sistemática y que no son obvios, es decir, requieren aplicar métodos científicos para su solución. La Tabla 1 muestra una muy simplificada relación de problemas relacionados a los cuatro núcleos de conocimiento o categorías de la Inteligencia Artificial.

Tabla 1. Categorías y algunos problemas a resolver por la IA

CATEGORÍAS DE LA IA	PROBLEMA
Representación del conocimiento	Representación del conocimiento tácito Representación del lenguaje natural
Tratamiento de la incertidumbre	Modelos para la toma de decisiones computacionales
Razonamiento	Búsqueda en espacios de estados Inferencia en sistemas basados en el conocimiento
Aprendizaje automático	Minería de datos Autonomía de agentes

El razonamiento, visto desde el punto de vista computacional como las vías para lograr un comportamiento racional, se considera el centro de los cuatro núcleos. La representación del conocimiento se realiza para facilitar el razonamiento. El tratamiento de la incertidumbre se materializa en el razonamiento, aunque se toma en cuenta desde la representación del conocimiento y se expresa en el aprendizaje. En tanto el aprendizaje se materializa mediante el razonamiento. Estos criterios refuerzan la idea de que la división de los términos se realiza para facilitar su estudio. Su estrecha relación hace imposible exponer alguno sin abordar los demás.

La IA tiene un método particular para resolver sus problemas: la utilización de heurísticas. La IA se encarga de definir y utilizar heurísticas tanto en el razonamiento como en el aprendizaje. Las heurísticas son la vía esencial de la IA para dar respuestas a los problemas de su objeto de estudio: El desarrollo de sistemas con comportamiento racional.

Hay tareas mucho más complejas a las cuales se enfrenta la Inteligencia Artificial y que requieren la conjunción de varios o incluso

todos los núcleos de conocimiento, incluyendo otras áreas y ciencias. Algunos de los ejemplos más significativos son (Figura 6):



Figura 6. Núcleos de conocimiento y algunas aplicaciones de la IA.

El procesamiento de lenguaje natural: Aunque algunos autores lo consideran un núcleo básico, la realidad es que para lograr el procesamiento de lenguaje natural es necesario *representar el co-*

nocimiento tanto de elementos gramaticales del idioma como del contenido específico que se analice. Es necesario llevar a cabo un *razonamiento* que incluya la percepción del interlocutor y tomar la decisión de la respuesta a brindar, tomando en cuenta un fuerte *tratamiento de la incertidumbre*. También es necesario el *aprendizaje* a corto plazo, por ejemplo, para llevar el control del tema que se trata, y a largo plazo para obtener el conocimiento que se comunica.

La visión por computador: Necesita *representar el conocimiento* que viene en forma de imágenes estáticas o en secuencias de video. Posteriormente, aplica métodos de *razonamiento* para interpretar las imágenes que sin dudas necesitarán un fuerte *tratamiento de la incertidumbre*. Requiere el *aprendizaje* mínimo de la secuencia de imágenes y del perfeccionamiento del análisis a realizar. Los algoritmos de aprendizaje automático son esenciales en el propio análisis de las imágenes y su interpretación.

También es evidente la relación de la IA con otras ramas de la computación o incluso otras ciencias para lograr los resultados esperados en visión por computador. Se reconoce su relación con los gráficos por computador y la física, pero también deben tomarse en cuenta aspectos de la medicina, la biología e incluso la psicología.

Los videojuegos: Esta es un área de mucho interés, incluso en el mercado. Es uno de los espacios donde mejor se representan los núcleos de conocimiento de la IA. Tanto es así que son varias las experiencias en el mundo que utilizan videojuegos como parte del proceso de enseñanza-aprendizaje de asignaturas de IA.

La *representación del conocimiento* tanto del juego en sí como de los personajes, es esencial para la inteligencia en el juego. El *razonamiento*, conocido como la lógica del juego, es el núcleo del videojuego, el cual mantendrá un *tratamiento de la incertidumbre*, no tanto para lograr los resultados más precisos como en otras tareas

de la IA, sino para ser más realistas, más parecidos a lo que pasaría en la vida real o simulada. *El aprendizaje*, de igual forma, permite mayor realismo en el juego, sobre todo cuando logra una adecuada adaptación al jugador.

Los videojuegos guardan relación con otras áreas como los gráficos por computador, la física y las artes, tanto en lo relativo a la música como a la imagen.

La robótica: Esta es la más compleja de las tareas de la IA. Quizás por eso, muchas veces se tiende a decirle IA a los robots. A la robótica se ajustan las mismas necesidades de los videojuegos o cualquier sistema de visualización gráfica. Asumiendo que solo cambia el entorno, se pasa de uno gráfico simulado a uno real, un personaje en un videojuego es como un robot en la vida real. A esto se le adiciona la visión por computador, el procesamiento de lenguaje natural, el movimiento, el tacto, entre otras muchas que requieren áreas como la mecánica, la electrónica y muchas otras ciencias e ingenierías.

EL DESARROLLO DE AGENTES RACIONALES

La definición del término agente es muy controvertida, son varios los autores que han dado su propia interpretación y definición de agentes. Russell y Norvig (1995) en su libro *Artificial Intelligence. A modern approach* realizan uno de los primeros pronunciamientos sobre agentes. Los autores se referían a un agente como una entidad que percibe y actúa sobre un entorno. A pesar de ser una definición muy sencilla encierra la esencia del concepto. Un poco más elaborada es la definición de agente como un sistema computacional, situado en un entorno específico, que es capaz de actuar de forma autónoma y flexible, para cumplir sus propios objetivos (Martínez, 2012).

Lo abarcador de este concepto lo hace ser reconocido como un nuevo paradigma de programación, una nueva forma de ver el mundo para ser llevado al lenguaje de la computación. Reconocido dentro de la rama de la Inteligencia artificial, ya ha sido generalizado a una amplia gama de aspectos de la sociedad.

Asumiendo a un humano como agente, podemos decir que recibe información del entorno a partir de sus sentidos y órganos como los ojos, oídos, manos, etc. El humano actúa en el entorno utilizando manos, piernas, boca y otras partes del cuerpo. Un agente no humano puede recibir pulsaciones del teclado, archivos de información,

paquetes vía red a modo de entradas sensoriales y actúa sobre el medio con mensajes en el monitor, escribiendo ficheros, enviando paquetes por la red, etc. Se asume que cada agente puede percibir sus propias acciones, pero no siempre sus efectos.

El término percepción se utiliza para indicar que el agente puede recibir entradas en cualquier instante. La secuencia de percepciones de un agente refleja el historial completo de lo que el agente ha recibido. En general, un agente tomará una decisión en un momento dado dependiendo de la secuencia completa de percepciones hasta ese instante. Si se puede especificar qué decisión tomará un agente para cada una de las posibles secuencias de percepciones, entonces se habrá explicado más o menos todo lo que se puede decir de un agente. En términos matemáticos se puede decir que el comportamiento del agente viene dado por la función del agente que proyecta una percepción dada, en una acción.

La función que describe el comportamiento de un agente se puede presentar en forma de tabla; en la mayoría de los casos esta tabla sería muy grande, infinita, a menos que se limite el tamaño de la secuencia de percepciones que se quiera considerar. Dado un agente, con el que se quiera experimentar, se puede, en principio, construir esta tabla teniendo en cuenta todas las secuencias de percepción y determinando qué acción lleva a cabo el agente en respuesta. La tabla es, por supuesto, una caracterización externa del agente.

Muchas veces se asume indistintamente el término sistema o el término agente. Un sistema computacional puede verse como un agente, lo cual se defiende desde el propio paradigma de programación para agentes. Siguiendo este paradigma, todo sistema se modela desde una visión de agentes. Cuando un sistema está formado por varios agentes, es reconocido como un sistema multiagente. Con la intención de ser más generales, en este libro, se tratan los agentes

como entidades computacionales que pueden ser sistemas independientes o estar incluidos en sistemas computacionales desarrollados siguiendo cualquier paradigma de programación. De ahí que se asume el término sistema (computacional) como más general.

Con estas ideas se puede definir que:

un agente es una entidad computacional que se comporta a partir de maximizar una medida de rendimiento cuyo valor depende de las percepciones realizadas de su entorno.

Son varias las características que se le atribuyen a los agentes, sin embargo hay tres que se reconocen como esenciales para que un agente inteligente pueda alcanzar sus objetivos (Coca, 2009) (Martínez, 2012):

Reactivo: El agente debe ser capaz de percibir su entorno y actuar en consecuencia a los cambios que ocurren en el mismo.

Pro-activo: El agente debe ser capaz de comportarse siguiendo sus propios objetivos.

Social: El agente debe ser capaz de comunicarse con otros agentes o con los humanos mediante algún tipo de comunicación.

Partiendo de que el objeto de estudio de la Inteligencia Artificial es el desarrollo de sistemas con comportamiento racional, se puede abordar la racionalidad del agente desde cuatro puntos de vista (Russell & Norvig, 2004):

- La medida de rendimiento, que define el criterio de éxito.
- El conocimiento acumulado por el agente, del medio en el que habita.
- Las acciones que el agente puede llevar a cabo.

- La secuencia de percepciones del agente hasta un momento determinado.

Con estos criterios se puede decir que un agente racional en cada posible secuencia de percepciones, seleccionará la acción que espera maximice su medida de rendimiento, basándose en las evidencias aportadas por la secuencia de percepciones y en el conocimiento que posee.

Estos elementos están estrechamente ligados a las cuestiones éticas de la IA. El desarrollo de agentes debe incluir los consensos internacionales sobre los aspectos éticos. Desde las percepciones que realiza, pasando por el conocimiento que acumula, la medida de rendimiento que asume y las acciones que lleva a cabo, todas dependen e influyen en el comportamiento ético de los sistemas inteligentes.

En correspondencia con la definición de IA, es necesario tener cuidado al distinguir entre racionalidad y omnisciencia. Un agente omnisciente conoce el resultado de su acción y actúa de acuerdo con él; sin embargo, en realidad la omnisciencia no es posible. Adaptando un ejemplo de Russell y Norvig (2004), suponga que está paseando por el Paseo del Prado de La Habana y ve un amigo en el muro del malecón. El semáforo de automóviles está en rojo y no tiene ningún compromiso en ese momento, entonces, actuando racionalmente, comenzará a cruzar la avenida del malecón. Al mismo tiempo, a 10 km de altura, se desprende la puerta de un avión, y antes de que termine de cruzar al otro lado de la calle fue aplastado.

Este ejemplo muestra que la racionalidad no es lo mismo que la perfección. La racionalidad maximiza el rendimiento esperado, mientras la perfección maximiza el resultado real. Resulta imposible diseñar un agente que siempre lleve a cabo, de forma sucesiva, las mejores acciones después de un acontecimiento. Las acciones racio-

nales dependen sólo de la secuencia de percepción hasta el momento del análisis por el agente.

Entorno de trabajo REAS

En la discusión de la racionalidad de un agente es útil especificar las medidas de rendimiento, el entorno, los actuadores y sensores del agente. A esto se le denomina entorno de trabajo y se conoce por sus siglas REAS (Rendimiento, Entorno, Actuadores, Sensores). En el diseño de un agente, el primer paso debe ser siempre especificar el entorno de trabajo de la forma más completa posible.

El rango de los entornos de trabajo en los que se utilizan técnicas de IA es obviamente muy grande (Tabla 2). Sin embargo, se puede identificar un pequeño número de dimensiones en las que categorizar estos entornos (Russell & Norvig, 2004). Estas dimensiones determinan, hasta cierto punto, el diseño más adecuado para el agente y la utilización de cada una de las familias principales de técnicas en la implementación del agente (Tabla 3).

Tabla 2. Ejemplos de agentes. Definición REAS

TIPO DE AGENTES	MEDIDAS DE RENDIMIENTO	ENTORNO	ACTUADORES	SENSORES
Taxista	Seguridad, rapidez, confortabilidad, beneficio	Ciudad (Carreteras, tráfico, peatones, clientes)	Dirección, acelerador, freno, luces, bocina, visualizador	Cámaras, GPS, relojes, receptor de sonido, teclado

TIPO DE AGENTES	MEDIDAS DE RENDIMIENTO	ENTORNO	ACTUADORES	SENSORES
Sistema de diagnóstico médico	Pacientes sanos, reducir costos	Pacientes, hospital, personal	Resultados, indicaciones	Teclado para la entrada de síntomas, conclusiones, respuestas de pacientes
Sistema de análisis de imágenes de satélites	Categorización de imagen correcta	Conexión con el satélite en órbita	Categorización de imagen correcta	Matriz de pixels de colores
Controlador de una refinería	Maximizar la pureza, producción y seguridad	Refinería, operadores	Válvulas, bombas, calentadores, monitores	Temperatura, presión, sensores químicos
Tutor de inglés interactivo	Maximizar la puntuación de los estudiantes en los exámenes	Conjunto de estudiantes, agencia examinadora	Visualizar los ejercicios, sugerencias, correcciones	Teclado de entrada

Si los sensores del agente le proporcionan acceso al estado completo del medio en cada momento, entonces se dice que el entorno de trabajo es **totalmente observable**. En entornos totalmente observables el agente no necesita mantener ningún estado interno para saber qué sucede en el mundo. Un entorno puede ser **parcialmente observable** debido al ruido y a la existencia de sensores poco exactos o porque los sensores no reciben información de parte del sistema. Por ejemplo, un taxi automatizado no puede saber qué están pensando otros conductores. Este análisis nos lleva a la convicción de que en la mayoría de los problemas del mundo real el entorno es parcialmente observable.

Si el siguiente estado del entorno está totalmente determinado por el estado actual y la acción ejecutada por el agente, entonces se

dice que el entorno es **determinista**; de otra forma es **estocástico**. En principio, un agente no se tiene que preocupar de la incertidumbre en un medio totalmente observable y determinista. Un agente taxista es claramente estocástico, ya que no se puede predecir el comportamiento del tráfico exactamente; más aún, una rueda se puede reventar y un motor se puede romper sin previo aviso. Si el medio es determinista, excepto para las acciones de otros agentes, decimos que el medio es estratégico.

Tabla 3. Dimensiones para analizar los entornos de trabajo de los agentes

CRITERIO	RESUMEN
Observable Parcialmente observable	Es observable cuando los sensores proporcionan acceso al estado completo del entorno en cada momento.
Determinista Estocástico	Si el siguiente estado del entorno está totalmente determinado por el estado actual y la acción ejecutada por el agente, entonces es determinista.
Episódico Secuencial	Si la experiencia del agente se divide en episodios atómicos, hablamos de un entorno episódico. Cada episodio consiste en la percepción del agente y la realización de una única acción posterior. El siguiente episodio no depende de las acciones que se realizaron en episodios previos.
Estático Dinámico	Si el entorno puede cambiar cuando el agente está razonando, el entorno es dinámico. En el estático el agente no necesita estar pendiente del mundo mientras está tomando una decisión sobre una acción, ni necesita preocuparse sobre el paso del tiempo.
Discreto Continuo	Es discreto cuando hay un conjunto finito de estados distintos del entorno, el tiempo no se maneja de forma continua o las percepciones y acciones del agente están delimitadas a un conjunto específico.
Agente individual Entorno multiagente	Un agente está en un entorno multiagente cuando no es la única entidad que depende del comportamiento de los demás entes.

Si la experiencia del agente se divide en episodios atómicos, hablamos de un entorno **episódico**. Cada episodio consiste en la percepción del agente y la realización de una única acción posterior. Es importante tener en cuenta que el siguiente episodio no depende de las acciones que se realizaron en episodios previos. Muchas tareas de clasificación son episódicas. Por ejemplo, un agente que tenga que seleccionar partes defectuosas en una cadena de montaje basa sus decisiones en la parte que está evaluando en cada momento, sin tener en cuenta decisiones previas. La decisión presente no afecta que la próxima fase sea defectuosa. En entornos **secuenciales**, por otro lado, la decisión presente puede afectar a decisiones futuras. El ajedrez y el taxista son secuenciales: en ambos casos, las acciones que se realizan a corto plazo pueden tener consecuencias a largo plazo. Los medios episódicos son más simples que los secuenciales porque el agente no necesita pensar con tiempo.

Si el entorno puede cambiar cuando el agente está razonando, entonces se dice que el entorno es **dinámico** para el agente; de otra forma se dice que es **estático**. Los entornos estáticos son fáciles de tratar ya que el agente no necesita estar pendiente del mundo mientras está tomando una decisión, ni necesita preocuparse sobre el paso del tiempo. Los medios dinámicos, por el contrario, están preguntando continuamente al agente qué quiere hacer; si no se ha decidido aún, entonces se entiende que ha tomado la decisión de no hacer nada.

El taxista es claramente dinámico: tanto los otros carros como el taxi se están moviendo mientras el algoritmo que guía la conducción indica qué es lo próximo a hacer. En cambio, juegos como los crucigramas son estáticos.

La distinción entre **discreto** y **continuo** se puede aplicar al estado del medio, a la forma en la que se maneja el tiempo y a las percep-

ciones y acciones del agente. Por ejemplo, un medio con estados discretos como el del juego del ajedrez tiene un número finito de estados distintos. El ajedrez tiene un conjunto discreto de percepciones y acciones. El taxista conduciendo define un estado continuo y un problema de tiempo continuo: la velocidad y la ubicación del taxi y de los otros vehículos pasan por un rango de valores continuos de forma suave a lo largo del tiempo. La conducción del taxista es también continua, lleva de forma continua el ángulo de dirección, la presión de los pedales, etc.

Las imágenes captadas por cámaras digitales son discretas, en sentido estricto, pero se tratan típicamente como representaciones continuas de localizaciones e intensidades variables.

La distinción entre el entorno de un **agente individual** y el de un **sistema multiagente** puede parecer suficientemente simple. Por ejemplo, un agente resolviendo un crucigrama por sí mismo está claramente en un entorno de agente individual, mientras que un agente que juega al ajedrez está en un entorno con dos agentes.

La pregunta en este punto es ¿Qué entidades se deben considerar agentes? Asumamos el agente A que debe interactuar con un objeto B. Puede tratarse a B como un objeto con un comportamiento estocástico; por ejemplo como las olas de la playa o las hojas que mueve el viento, o bien puede considerarse un agente. La cuestión está en identificar si el comportamiento de B está descrito por la maximización de una medida de rendimiento cuyo valor depende del comportamiento de A. Por ejemplo, en el ajedrez, la entidad oponente B intenta maximizar su medida de rendimiento, la cual, según las reglas, minimiza la medida de rendimiento del agente A. Por tanto, el ajedrez es un entorno multiagente competitivo. Por otro lado, en el medio definido por el taxista circulando, el evitar colisiones maximiza la medida de rendimiento de todos los agentes, así pues

es un entorno multiagente parcialmente cooperativo. Es también parcialmente competitivo ya que, por ejemplo, sólo un carro puede ocupar una plaza de aparcamiento.

Los problemas en el diseño de agentes que aparecen en los entornos multiagente son a menudo bastante diferentes de los que aparecen en entornos con un único agente. Por ejemplo, la comunicación se considera un comportamiento racional en entornos multiagente.

En un sistema multiagente varios agentes actúan en el mismo entorno para dar cumplimiento a tareas específicas. Para la solución de problemas específicos se puede decir que (Martínez, 2012):

- Cada agente tiene información o capacidades incompletas para resolver el problema.
- Cada agente tiene un limitado punto de vista del problema en su conjunto.
- No existe un control global del sistema.
- Los datos pueden ser descentralizados.
- La computación puede ser asíncrona.

Ya existen instituciones que han manejado normas para la estandarización de los agentes y lograr una comunicación entre ellos. Entre ellas tenemos los estándares FIPA (IEEE, 2005) que tienen reglamentaciones de lenguajes de comunicación, construcción y hasta de áreas de aplicación de los agentes.

Muchas de las aplicaciones de sistemas informáticos son algorítmicas y trabajan con información exacta. Pero la mayoría de las aplicaciones requieren sistemas más complejos, sistemas capaces de relacionarse con un entorno cambiante y con un cierto grado

de incertidumbre. Los agentes y los sistemas multiagentes tienen por tanto que ser capaces de proporcionar soluciones a este tipo de problemas. El modelo BDI fue uno de los primeros y más simples que se ha desarrollado para proporcionar soluciones en entornos dinámicos o inciertos. En estos el agente o los agentes sólo tienen una visión parcial del problema, el acceso a la información está limitado y posiblemente manejen un número limitado de recursos. Las creencias, los deseos, las intenciones, y los planes son una parte fundamental del estado de ese tipo de sistemas.

En el campo de la inteligencia artificial, las creencias representan el conocimiento que se tiene del entorno. Desde un punto de vista informático, son la forma de representar el estado del entorno, por ejemplo el valor de una variable, los valores de una base de datos relacional, o expresiones simbólicas del cálculo de predicados. Las creencias son esenciales en esta arquitectura. En medios dinámicos es necesario mantener información sobre los eventos pasados, al mismo tiempo que se debe permitir su adaptación y evolución, por lo que una buena gestión de estos elementos es vital para construir agentes eficientes.

Los deseos, considerados como los objetivos del agente, son otro componente esencial en esta arquitectura. En términos informáticos, un objetivo puede simplemente ser el valor de una variable, un registro, o una expresión simbólica en alguna lógica. Un objetivo representa un estado final deseado. El software convencional está orientado a la tarea, de forma que cada tarea se ejecuta sin ningún recuerdo de porqué ha comenzado su ejecución. Esto significa que el sistema no puede recuperarse ante fallos automáticamente, a menos que esto sea específicamente codificado por el programador y no puede descubrir ni aprovechar oportunidades que surjan inesperadamente.

En cambio, este modelo está orientado al objetivo, la semántica fundamental de los objetivos, sin tener en cuenta cómo se representa computacionalmente, se reflejaría en una lógica de los deseos.

Para alcanzar los objetivos propuestos, a partir de las creencias existentes es necesario definir un mecanismo de planificación que nos permita identificar las intenciones. En este sentido, tenemos que tener claro que los agentes están inmersos en sistemas dinámicos, en los que en ocasiones será necesario decidir si replanificar o no, ante cambios en el entorno, durante la ejecución del plan inicialmente seleccionado. Los defensores de la teoría de la decisión clásica dicen que siempre se debería replanificar, mientras que los diseñadores de software convencional orientado a la tarea obligarían a continuar hasta el final con el plan inicialmente previsto. Lo cierto es que el sistema necesita comprometerse con los planes y subobjetivos, pero también ser capaz de reconsiderar éstos en los momentos clave. Estos planes dirigidos a lograr un objetivo constituyen las intenciones del agente. Las intenciones son simplemente un conjunto de caminos de ejecución que pueden ser interrumpidos de una forma apropiada al recibir información acerca de cambios en el entorno.

La arquitectura propuesta por Rao y Georgeff (Corchado, 2005) se basa en la construcción de sistemas de agentes cuyas estructuras de datos se corresponden con cada uno de los componentes de los agentes BDI: creencias, deseos e intenciones, como se muestra en la Figura 7. Cada uno de los componentes descritos se guarda aisladamente, o sea, existe una lista de creencias, otra de deseos y otra de intenciones. Además, se trabaja con una secuencia de eventos o acontecimientos ocurridos en el entorno. Cada uno de ellos se implementa como una cola, esto es, mediante una estructura de datos lineal con comportamiento primero en llegar primero en salir (FIFO por sus siglas en inglés).

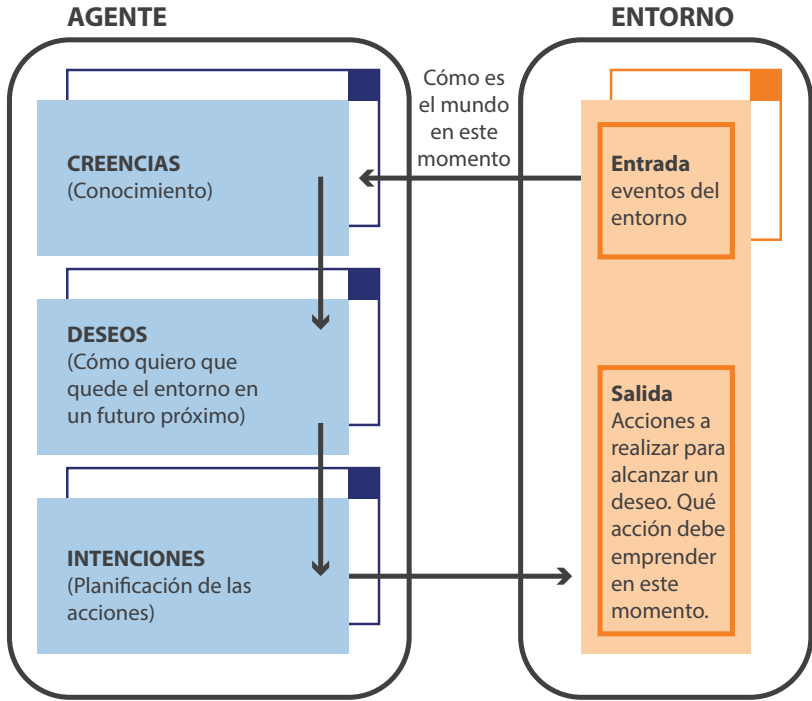


Figura 7. Representación del modelo BDI.

Expongamos un ejemplo de modelado de un agente en un sistema multiagente según BDI.

Agente a modelar: Taxista.

Entorno: Ciudad virtual en 3D.

Otros agentes: otros autos, transeúntes, clientes.

Objetos: Señales de tránsito, obstáculos, asientos en parques o la calle.

Creencias:

- Dimensiones de la calle por donde transita.
- Intersecciones de calles.

- Posición de cada auto en cada instante de tiempo.
- Posición y área de acción de cada señal del tránsito.
- **Deseos (Se relacionan con las medidas de rendimiento):**
- Movimiento realista por el entorno de ciudad.
- Maximizar ganancias.
- Garantizar la seguridad.
- Llegar en el menor tiempo posible.
- Garantizar confort.

Intenciones:

- Respetar cada señal de tránsito.
- No colisionar.
- Decisiones de calles a seguir.
- Decisión de subir o bajar a clientes.

El nivel de detalle con que se modela el agente dependerá de la complejidad que estemos dispuestos a incorporar al agente.

Para la implementación de los agentes es necesario utilizar lenguajes de programación. Pero ¿Existe alguno específico para ello? Uno de los primeros trabajos en lenguajes de agentes fue *Agent-0* (Shoham, 1993) que acuñó el término programación orientada a agentes. Otro trabajo realizado es el JavaLog (Zunino, Berdún, & Amandi, 2001), una unión del Java orientado a objetos con el Prolog, programación lógica. Sin embargo no existe todavía un lenguaje bien establecido para desarrollar orientado a agentes.

Existen varias bibliotecas que sí facilitan el desarrollo desde este paradigma, una de las más reconocidas es JADE (Bellifemine, Poggi, & Rimassa, 1999), una biblioteca desarrollada sobre Java para el desarrollo de sistemas multi-agentes y que sigue el estándar FIPA. De igual forma se han desarrollado varias metodologías (Tran & Low, 2005) siendo una de las más reconocidas la metodología INGENIAS (Gómez, 2002).

Aunque cada arquitectura, metodología, lenguaje o biblioteca específica, tiene sus particularidades para el desarrollo de los agentes, se pueden identificar cuatro tipos básicos de agentes (Russell & Norvig, 2004). Cada uno de ellos con una estructura interna general, que va aumentando en complejidad según el tipo de agente.

Agentes reactivos simples: Realizan una proyección directa de las condiciones del estado actual a las acciones, ignorando el resto de las percepciones históricas.

Agentes reactivos basados en modelos: Mantienen un estado interno que les permite seguir el rastro de aspectos del mundo a partir de la historia percibida y que no son evidentes según las percepciones actuales. Este estado debe reflejar por lo menos alguno de los aspectos no observables del estado actual. Necesitan un método para inferir las propiedades relevantes del mundo a partir de una secuencia de percepciones.

Agentes basados en objetivos: Actúan con la intención de alcanzar sus metas. Guardan información sobre cómo evoluciona el mundo y sobre los resultados de las posibles acciones que puede llevar a cabo. Necesita información sobre su meta de manera que describa las situaciones que son deseables.

Agentes basados en utilidad: Intentan maximizar su función de utilidad. Para ello requieren información de utilidad, que indique lo deseables que son los estados. Es decir, información acción-valor, que indica lo deseables que son las acciones que va a ejecutar y metas que describen las clases de estados que maximizan su utilidad.

Agentes que aprenden: Todos los agentes pueden mejorar su eficacia con la ayuda de mecanismos de aprendizaje. El aprendizaje puede verse como el proceso de modificación de cada componente del agente, lo cual permite a cada componente comportarse más

en consonancia con la información que se recibe, lo que conlleva a mejorar el nivel medio de actuación del agente.



CONSIDERACIONES FINALES

Existe consenso en que la Inteligencia Artificial va dirigida al desarrollo de sistemas que incorporan cierto nivel de conocimiento humano para lograr sus objetivos. Las definiciones se han movido en cuatro direcciones fundamentales: Los sistemas que piensan como humanos, los que actúan como humanos, los que piensan racionalmente y los que se comportan racionalmente. El enfoque de comportamiento racional es el más general, abarca a los demás y prioriza la solución de problemas, sobre la imitación de comportamiento o pensamiento.

Para facilitar el estudio de la IA se propone agrupar sus principales tareas computacionales en cuatro núcleos de conocimiento: La representación del conocimiento, el razonamiento, el tratamiento de la incertidumbre y el aprendizaje. Estos cuatro núcleos de conocimiento van dirigidos a la tarea de construir sistemas con comportamiento racional. Estos sistemas abarcan hoy una gran cantidad de áreas, algunas de ellas tan relacionadas con la IA, como la minería de datos o la robótica, que muchos las asocian directamente con su objeto de estudio.

El desarrollo del paradigma de agentes ha venido a ser un catalizador del desarrollo de la IA. Este paradigma refuerza el objeto

de estudio de la IA en lograr un comportamiento racional en base al conocimiento que se adquiriera del entorno. Como paradigma de programación asume una nueva visión del mundo para el desarrollo de sistemas computacionales, lo cual brinda la posibilidad de desarrollar prácticamente cualquier sistema informático con una visión de IA desde su concepción.



GLOSARIO DE TÉRMINOS

Agente (racional): Entidad computacional que se comporta a partir de maximizar una medida de rendimiento cuyo valor depende de las percepciones realizadas de su entorno.

Aplicaciones (de la IA): Áreas donde se evidencia la utilización práctica de los conocimientos sobre una materia, relacionándose con otras materias científicas o ingenieriles. Las aplicaciones de la IA incluyen otras ramas de las ciencias de la computación, otras ciencias y sectores específicos de la economía y la sociedad. Resaltan en nuestros días el procesamiento de lenguaje natural, la minería de datos, los videojuegos y la robótica.

Aprendizaje (en un agente): Utilizar las percepciones de su entorno para comportarse y mejorar su habilidad para actuar en el futuro.

Aprendizaje no supervisado: El sistema se comporta a partir de patrones de entradas para los que no se especifican los valores de sus salidas.

Aprendizaje reforzado: El sistema actúa a partir de la información que obtiene de su propio comportamiento en el entorno.

Aprendizaje supervisado: El sistema se comporta a partir de la utilización de ejemplos con sus entradas y sus salidas.

Entorno (de un sistema computacional): Conjunto de datos, informaciones, conocimiento, circunstancias, factores, etc., que influyen en el estado o desarrollo de un sistema computacional.

Heurística: Criterio que evalúa la posibilidad de que una búsqueda vaya en la dirección correcta.

Incertidumbre (en IA): Falta de conocimiento sobre el entorno.

Inteligencia Artificial: Es la Ciencia de la Computación encargada de aplicar métodos de representación del conocimiento, razonamiento, tratamiento de la incertidumbre y aprendizaje, en el desarrollo de sistemas informáticos con comportamiento racional.

Núcleo de conocimientos (de la IA): Temáticas básicas en las que se estructura una materia o disciplina para su estudio. Los núcleos estrechamente relacionados de la IA son la Representación del Conocimiento, el Razonamiento, el Tratamiento de la Incertidumbre y el Aprendizaje.

Razonamiento (computacional): Proceso general llevado a cabo por un sistema para comportarse de manera racional a partir del conocimiento que tenga de su entorno.

Representación del conocimiento: Proceso de estructurar el conocimiento mediante una notación suficientemente precisa como para ser utilizado por el sistema en su comportamiento.



REFERENCIAS

- Bellifemine, F., Poggi, A., & Rimassa, G. (1999).** JADE-A FIPA-compliant agent framework. *Proceedings of PAAM*, 97–108. <https://doi.org/10.1145/375735.376120>
- Bello, R. E., García, Z. Z., García, M. M., & Reynoso, A. (2002).** *Aplicaciones de la Inteligencia Artificial* (Primera ed). Guadalajara.
- Coca, Y. (2009).** Agentes inteligentes. Aplicación a la realidad virtual. *Revista Cubana de Ciencias Informáticas*, 3(1–2), 49–54.
- Coca, Y. (2021).** *Modelo de integración de software libre educativo al proceso de enseñanza-aprendizaje de la Inteligencia Artificial*. Universidad de las Ciencias Informáticas.
- Corchado, J. M. (2005).** Modelos y Arquitecturas de Agente. In J. L. Perez de la Cruz (Ed.), *Agentes software y sistemas multi-agente: conceptos, arquitecturas y aplicaciones*. (pp. 1–40).
- Gómez, J. (2002).** Modelado de sistemas multi-agente (Universidad Complutense de Madrid). Retrieved from <http://grasia.fdi.ucm.es/jorge/tesis.pdf>
- IEEE. (2005).** *FIPA Standards Committee. Policies and Procedures*.
- Martínez, Y. (2012).** *A Generic Multi-Agent Reinforcement Learning Approach for Scheduling Problems*. Vrije Universiteit Brussel.

- Russell, S., & Norvig, P. (1995).** *Artificial Intelligence. A modern approach* (1st Edit.). Prentice Hall.
- Russell, S., & Norvig, P. (2004).** *Inteligencia Artificial. Un enfoque moderno* (2da Edició). Madrid: Pearson Educación.
- Shoham, Y. (1993).** Agent-oriented programming. *Artificial Intelligence*, 60(1), 51–92. [https://doi.org/10.1016/0004-3702\(93\)90034-9](https://doi.org/10.1016/0004-3702(93)90034-9)
- Tran, Q. N. N., & Low, G. C. (2005).** Comparison of ten agent-oriented methodologies. In *Agent-Oriented Methodologies* (pp. 341–367). <https://doi.org/10.4018/978-1-59140-581-8.ch012>
- Zunino, A., Berdún, L., & Amandi, A. (2001).** JavaLog : un Lenguaje para la Programación de Agentes. *Revista Iberoamericana de Inteligencia Artificial*, (13), 94–99.

“La Inteligencia Artificial como una ciencia de la computación” es el segundo de cuatro libros que tienen como objetivo principal propiciar la incorporación de conceptos e ideas generales sobre Inteligencia Artificial en las escuelas. Se exponen los cuatro puntos de vista principales que se han asumido para definir el término Inteligencia Artificial. Se presentan algunos criterios sobre las temáticas específicas que estudia y se presenta una definición en base a estos criterios. Se aborda la Inteligencia Artificial desde el desarrollo de agentes racionales, un paradigma de programación nacido indisolublemente relacionado con ella.

Pretendemos que esta obra sirva de base para contribuciones teóricas y prácticas mediante el desarrollo de actividades educativas curriculares y no curriculares para las instituciones educativas, escuelas, universidades, familias, comunidades y la sociedad en general, así como para las instituciones y organizaciones gubernamentales y no gubernamentales; y de todos los sectores e instituciones de la sociedad. No se trata de enseñar a desarrollar aplicaciones de Inteligencia Artificial, sino, de contribuir al desarrollo del pensamiento computacional y nociones básicas necesarias para incentivar el estudio y profundización en IA con un posicionamiento ético, crítico y creativo, desde edades tempranas.

ISBN: 978-959-18-1341-1

