

INTELIGENCIA ARTIFICIAL 06 de febrero

Artificial Intelligence a guide to intelligent systems

1.1 the story of artificial intelligence, or from the "dark ages" to knowledge-based systems

La inteligencia artificial como ciencia fue fundada por 3 generaciones de investigadores.

El primer trabajo reconocido en el campo de la IA fue presentado por Warren McCulloch y Walter Pitts en 1943. McCulloch tenía dos títulos en psicología y medicina de la universidad de Columbia. Su

investigación del sistema nervioso central resultó en la primera gran contribución a la IA: un modelo de neuronas del cerebro.

McCulloch y Pitts propusieron un modelo de un modelo artificial de redes neuronales, en donde cada neurona fue postulada como un estado binario.

Ellos demostraron que su modelo de red neuronal era un equivalente de la máquina de Turing, y probaron que cualquier función calculable puede ser calculada por algunas redes de neuronas conectadas.

El modelo de redes neuronales simula tanto lo teórico y el trabajo experimental para modelar el cerebro en el laboratorio. Aunque demostraron que trabajar con el modelo binario de neuronas no es el correcto. Ahora sabemos que una neurona tiene unas altas características no lineales y no pueden ser consideradas como 2 simples estados.

El segundo "padre" de la IA McCulloch después de Alan Turing, crea la piedra angular de la computación neuronal y redes neuronales artificiales (ANN).

Después del declive en los 70, el campo de las ANN volvió a los finales de los 80.

El tercer fundador de la IA fue John von Neumann, su colega y amigo de Alan Turing. Durante la segunda guerra mundial, él jugó un rol importante

en el proyecto Manhattan, también se convirtió en un asesor de la ENIAC para diseñar una variable electrónica discreta de una computadora automática, una máquina de programa almacenable. Otro de los investigadores de primera generación fue Claude Shannon, él compartió la idea de Alan Turing en la posibilidad de una máquina inteligente, en 1950 publica un papel de un juego de ajedrez donde un juego simple de ajedrez implicaba alrededor de 10^{120} movimientos posibles. Shannon demostró que se necesitan usar heurística en la búsqueda de la solución.

Otro padre de la IA fue John McCarthy, en 1956 él trajo junto con otros investigadores redes neuronales artificiales y la teoría automática. Su trabajo fue financiado por la IBM. Este trabajo dio nacimiento a una nueva ciencia llamada Inteligencia Artificial.

1.1.2 The rise of artificial Intelligence, or the era of great expectations (1956-late 1960's)

En los primeros años de la IA estaban caracterizados por un gran entusiasmo, grandes ideas y mucho suceso limitado.

John McCarthy definió el lenguaje de alto nivel (LISP), en 1958 presentó un papel llamado "programas con sentido común" en este papel propuso un programa llamado "tomador de decisiones" para buscar la solución para los problemas generales del mundo. Más importante el programa fue diseñado para aceptar nuevos axiomas o en otras palabras nuevo conocimiento basado en un sistema incorporando los principios centrales de la representación del conocimiento y razonamiento.

Otro organizador del taller de Dartmouth fue Marvin Minsky también se movió a la MIT, él se centró en el desarrollo de un ~~outlook~~ ^{análisis} antiológico en representación del conocimiento y razonamiento. El trabajo de redes neuronales de McCulloch y Pitts continúa, se mejoraron métodos de aprendizaje y Frank Rosenblatt probó el teorema de percepción convergente, demostrando que su algoritmo de conocimiento puede ajustarse a la intensa conexión de perceptrón. Uno de los proyectos más ambiciosos de la era de grandes expectativas era el GPS, el cual fueron Allen Newell y Herbert Simon quienes propusieron el programa. El GPS fue probablemente el primer intento para separar el método de solución de problemas, estaba basada en la técnica ahora referido como análisis medio-fin. Newell y Simon postularon que el problema puede ser resuelto puede ser definido en estados. El análisis anterior fue usado para determinar la diferencia entre el estado actual y el estado deseable del problema, y para escoger y aplicar operadores para alcanzar el estado de éxito. Si el estado de éxito puede no obtenerse inmediatamente, un nuevo estado cercano del éxito puede establecerse, y el procedimiento se repetirá hasta que el estado de éxito sea alcanzado.

Sin embargo, el GPS no logró resolver problemas complicados, el programa se basó en lógica formal y por lo tanto podía generar un número infinito de operadores posibles, lo cual es inherentemente ineficiente. La cantidad de tiempo de computadora y

memoria que requería el GPS para resolver problemas del mundo real llevó al abandono del proyecto.

1.2.3 Unfulfilled promises, or the impact of reality (late 1960s - early 1970s)

En los mediados los 50, los investigadores de IA hicieron promesas de construir una máquina de inteligencia de todo propósito en la escala del conocimiento humano, y para exceder la inteligencia humana para el 2000. En 1970, sin embargo, ellos comprendieron que esos reclamos eran muy optimistas. A pesar de que algunos programas de IA podían demostrar un nivel de inteligencia de la máquina en uno o dos problemas de juego, casi ningún proyecto de IA podría abordar una selección más amplia de tareas o problemas más difíciles del mundo real.

Las dificultades principales de la IA en los finales de los 60 fueron:

- Los primeros programas de IA tenían poco o ningún conocimiento sobre el dominio de un problema específico, ya que los investigadores estaban desarrollando métodos generales para clases amplias de problemas. Utilizaban estrategias de búsqueda, probando combinaciones de pasos pequeños hasta encontrar la correcta. Aunque funcionaba para problemas simples, la idea de escalar estos programas para resolver problemas más grandes resultó ser incorrecta.
- Los problemas fáciles o tratables pueden resolverse en tiempo polinómico, mientras que los problemas difíciles o intratables requieren tiempos exponenciales.

en función del tamaño del problema. La teoría de NP-Complete demostró la existencia de problemas NP (no deterministas polinómicos) que son NP-completos, siendo los más difíciles en esta categoría.

Muchos problemas que la IA intentó abordar eran demasiado amplios y difíciles. En la década de 1970 los proyectos de traducción financiados por gobiernos de EE.UU. y el Reino Unido fueron cancelados debido a la dificultad de resultados significativos en la investigación de IA.

En la misma década de 1970, se hizo evidente que el dominio de problemas para las máquinas inteligentes debía ser lo suficientemente restringido. Los métodos generales y débiles utilizados anteriormente resultaron insuficientes, llevando a la necesidad de resolver casos típicos en áreas mediante grandes pasos de razonamiento.

El programa DENDRAL, desarrollado en Stanford, fue un ejemplo clave de esta aproximación. Se utilizó para analizar químicos, enfrentándose al desafío de determinar la estructura molecular del suelo marciano. El equipo desarrolló un sistema basado en conocimientos específicos, marcando un cambio paradigmático en la IA. DENDRAL originó la metodología de sistemas expertos y la ingeniería del conocimiento, destacando la importancia de mapear conocimientos teóricos a reglas específicas. Posteriormente, el proyecto MYCIN se enfocó en el diagnóstico médico de enfermedades sanguíneas infecciosas, demostrando que las computadoras podían igualar a expertos humanos en áreas específicas.

A pesar del éxito, los sistemas expertos representaron desafíos, como su limitación a dominios estrechos, falta de flexibilidad, capacidad de explicación limitadas y dificultades en verificación y validación. A pesar de esto, las aplicaciones exitosas de sistemas expertos crecieron en la década de 1980, mostrando su madurez tecnológica.

Aunque han mostrado su valor en diversas áreas, los sistemas expertos tienen limitaciones en su capacidad de aprendizaje, flexibilidad y validación. A pesar de los desafíos han sido fundamentales en aplicaciones importantes.

A mediados de la década de 1980, los investigadores y expertos descubrieron que construir un sistema experto requería más que simplemente comprar un sistema de razonamiento o una carcasa de sistema experto y agregar suficiente reglas. La desilusión sobre la aplicabilidad de la tecnología de sistemas expertos incluso llevó a predicciones de un "invierno" en la financiación severamente reducida para proyectos de IA. Ante eso, los investigadores de IA decidieron volver a examinar las redes neuronales.

A finales de la década de 1960, ya se habían formulado la mayoría de ideas y conceptos básicos necesarios para la computación neuronal. Sin embargo, solo a mediados de la década de 1980 surgió la solución, impulsada por avances y financieros, así como por la necesidad de procesamiento de información similar al cerebro.

En la década de 1980, la necesidad de

procesamiento de información similar al cerebro, junto con los avances en tecnología informática y el progreso en neurociencia, llevaron a un resurgimiento dramático de las redes neuronales. Se desarrollaron importantes contribuciones a la teoría y el diseño, como la teoría de resonancia adaptativa de Grossberg y las redes de retroalimentación de Hopfield. Sin embargo, el avance real ocurrió en 1986 con el algoritmo de aprendizaje de retropropagación, reinventado por Rumelhart y McClelland, marcando el renacimiento de las redes neuronales.

Las redes neuronales artificiales han evolucionado desde los modelos iniciales de McCulloch y Pitts hasta convertirse en un campo interdisciplinario con raíces en la neurociencia, psicología, matemáticas e ingeniería. A pesar de los desafíos, como las limitaciones de los sistemas expertos, las redes neuronales continúan desarrollándose tanto en teoría como en aplicaciones prácticas, siendo las obras de Hopfield y Rumelhart y McClelland las más significativas e influyentes en el renacimiento de las redes neuronales en la década de 1980.

La inteligencia natural es un producto de la evolución, y al simularla, se esperaba descubrir cómo los sistemas vivos alcanzaban la inteligencia de alto nivel. La evolución biológica implica competencia por la supervivencia, donde las especies más aptas tienen mayores posibilidades de reproducirse y transmitir sus genes. La

computación evolutiva o IA se basa en modelos computacionales de selección natural y genética, utilizando algoritmos genéticos, estrategias evolutivas y programación genética.

Los algoritmos genéticos, introducidos por John Holland en la década de 1970, manipulan "cromosomas" artificiales mediante operaciones genéticas como selección, cruzamiento y mutación. Las estrategias evolutivas, propuestas por Ingo Rechenberg y Hans-Paul Schwefel en la década de 1960, se centran en la optimización de parámetros mediante cambios aleatorios inspirados en la mutación natural.

Ambos enfoques resuelven una amplia gama de problemas complejos, proporcionando soluciones robustas para búsquedas y optimizaciones no lineales. La programación genética, promovida por John Koza en la década de 1990, aplica el modelo genético al desarrollo de programas informáticos como soluciones a problemas, superando el desafío de hacer que las computadoras resuelvan problemas sin programación explícita.

Los algoritmos genéticos, estrategias evolutivas y programación genética son áreas en rápido crecimiento en la IA, con un gran potencial para abordar problemas complejos y no lineales.

La tecnología en redes neuronales ofrecen una interacción más natural con el mundo real en comparación con los sistemas

basados en razonamiento simbólico. Las redes neuronales pueden aprender, adaptarse a cambios en el entorno del problema, establecer patrones en situaciones donde no se conocen reglas y manejar información difusa o incompleta. Aunque carecen de instalaciones de explicaciones y a menudo actúan como una caja negra, las redes neuronales y los sistemas expertos complementan en lugar de competir.

Los sistemas expertos clásicos son efectivos para aplicaciones en sistemas cerrados con entradas precisas y salidas lógicas. Utilizan conocimientos expertos en forma de reglas y, si es necesario, pueden interactuar con el usuario. Sin embargo, la limitación radica en que los expertos humanos no siempre pueden expresar su conocimiento en términos de reglas o explicar su razonamiento, lo que puede conducir al fracaso del sistema. La computación neuronal se puede utilizar para extraer conocimiento oculto en grandes conjuntos de datos y obtener reglas para sistemas expertos.

Otra tecnología importante es la lógica difusa, que maneja conocimiento y datos vagos, imprecisos e inciertos. A diferencia de los métodos basados en probabilidad, la lógica difusa se centra en valores difusos que capturan el significado de las palabras, el razonamiento humano y la toma de decisiones. Se introdujo en 1965, y aunque inicialmente fue lenta en ser aceptada, se ha utilizado con éxito en diversos

electrodomésticos y productos japoneses
desde 1987.