
Introducción a la Inteligencia Artificial (IA)

Introducción

- J.McCarthy: “It is the science and engineering of making intelligent machines, especially intelligent computer programs [...] using computers to understand human intelligence”
 - » Entender y construir agentes inteligentes
 - » Ver:
<http://www.aaai.org/AITopics/html/reference.html#online>
- Orígenes
 - » Desde la antigüedad: estudio de la memoria, el aprendizaje, el razonamiento, etc.
 - » Aproximaciones a la IA con la 2ª Guerra Mundial.
 - » Denominación *Inteligencia Artificial*: J. McCarthy en conferencia Dartmouth de 1956 J. McCarthy, M. L. Minsky, N. Rochester, and C.E. Shannon.
- Disciplinas relacionadas con Inteligencia Artificial:
 - » Percepción (visión, PLN), razonamiento lógico, resolución de problemas, demostración automática, aprendizaje, representación de conocimiento, etc.

Historia de la IA, I

- Gestación de la IA (1943-1956).
 - » Neuronas artificiales. McCulloch y Pitts (1943).
 - » Programas de ajedrez. Shannon y Turing (1951). Leonardo Torres Quevedo (1914).
 - » General Problem Solver (Puzzles).
 - » Geometry Theorem Prover. IBM, 1959.
 - » LISP. Creado por John McCarthy en 1958. MIT.
 - » Algoritmo de resolución de Robinson (1965).
 - » Mundo de bloques:
 - Visión.
 - Propagación de Restricciones.
 - Lenguaje Natural.
 - Planificación.

Historia de la IA, II

- Entusiasmo (1952-1969)
 - » M.Minsky: Búsqueda heurística. Aprendizaje (1968)
 - » T.Winograd: Comprensión de LN (1972)
 - » P.Winston: Mundo de bloques (1970/75)
- Realismo (1966-1974)
 - » Intratabilidad computacional de las soluciones
 - » Sistemas basados en el conocimiento (1969-1979).
 - DENDRAL (1969): Buchanan et al. Estructura molecular.
 - SHRDLU (1971): Winograd.
 - MYCIN (1974): Buchanan et al. Diagnóstico médico. Incertidumbre.
- IA moderna
 - » IA en la industria (1980-1988).
 - Sistemas Expertos comerciales (década de los 90)
 - Minería de datos (década de los 90).
 - Recuperación de Información (Salton, Spärk-Jones).
 - » Redes neuronales
 - Explotación de métodos teóricos de mediados de siglo.
 - Mecánica estadística (Hopfield), modelos de memoria de Rummelhart, redes topológicas (Kohonen).
 - » Aprendizaje Automático
 - » Representación del conocimiento. Tecnologías.
 - Redes Bayesianas (Pearl 1988)
 - Agentes (SOAR Newell et al. 1987)
 - Inteligencia distribuida / colectiva, Redes Sociales, Web Crawling, , Robots (mascotas, autónomos (DARPA)), domótica, etc.

Campos relacionados con IA, I

- Filosofía (desde 428 a.C.).
 - » Teorías de razonamiento, aprendizaje, etc
 - » Platón, Sócrates, Aristóteles
 - » Descartes (s. XVI). Distinción entre “mente” y “materia”.
 - » Leibniz (s. XVII). Materialismo.
 - » Francis Bacon (s. XVI). Empirismo.
 - » Bertran Russell (s. XIX). Positivismo lógico.
- Matemáticas (desde 800).
 - » Teorías formales de Lógica, Probabilidad, Teoría de la Decisión.
 - » Noción de algoritmo de Al-khowarazmi (matemático árabe, s. IX)
 - » Boole (1815)
 - » Frege (1848). Lógica de primer orden.
 - » Hilbert (1862). En 1900, presentó “23 problemas para el siglo XX”.

Campos relacionados con IA, II

- » Gödel (1906).
 - Teorema de incompletitud (1931).
 - En Lógica de Primer Orden con aritmética inductiva hay fórmulas ciertas no demostrables.
- Psicología (desde 1879).
 - » Teorías para estudiar la mente. Psicología cognitiva.
 - Craik (1943). Etapas de un agente basado en el conocimiento:
 - Transformación del estímulo a una representación interna.
 - Derivación de las representaciones internas.
 - Traducción de las representaciones internas a acciones.
- Lingüística (desde 1957).
 - Teorías acerca de la estructura y significado del lenguaje.
- Informática (desde 1940).
 - Herramientas para hacer realidad IA

Propósito de la IA

	SER HUMANO	SER RACIONAL
PENSAR como	2. Máquinas con mente	3. Modelos para procesos naturales
ACTUAR como	1. Androides	4. Diseño de máquinas con comportamiento inteligente

Definición de IA, I

- 1) Sistemas que actúan como el hombre.
 - » Test de Turing. *Computing machinery and intelligence*, Mind, 59, 433:460, 1950.
 - » Comportamiento inteligente
 - Incapacidad de diferenciar entre respuestas del ordenador y repuestas humanas.
 - Supone:
 - Procesamiento del Lenguaje Natural.
 - Representación del Conocimiento.
 - Razonamiento Automático.
 - Aprendizaje Automático.
 - » Test total de Turing.
 - Se permitiría interacción física entre persona y ordenador.
 - Supone:
 - Visión (para percibir objetos).
 - Robótica (para mover objetos).

Definición de IA, II

- 2) Sistemas que piensan como el hombre.
 - » Teorías de funcionamiento de la mente humana:
 - Campos de visión
 - Lenguaje Natural
 - Aprendizaje
 - » Se buscan modelos de IA compatibles con técnicas experimentales en Psicología.

Definición de IA, III

- 3) Sistemas que piensan racionalmente.
 - » Desarrollo de Lógica Formal a finales del siglo XIX y principios del XX.
 - » Trata de crear sistemas inteligentes utilizando la Lógica Formal.
 - Del estilo de los silogismos de Aristóteles:
 - “Sócrates es un hombre.
 - Todos los hombres son mortales,
 - luego Sócrates es mortal”
 - » Hacia 1965, había programas que resolvían problemas formulados en Lógica (supuestos memoria y tiempos suficientes).
 - » Inconvenientes:
 - Necesaria una representación del conocimiento informal (o difuso). Uso de probabilidades.
 - Explosión combinatoria de posibilidades.

Definición de IA, IV

- 4) Sistemas que actúan racionalmente.
 - » Uso de agentes: percepción + actuación.
 - » Se necesita resolver situaciones, que el pensamiento racional no puede por sí solo hacer:
 - Acciones reflejas: “retirar la mano del fuego”.
 - » El estudio de IA como agentes racionales tiene dos ventajas:
 - Es más general que el “pensamiento racional”.
 - Es más cercano al método científico que el comportamiento y el pensamiento humanos.

Agentes, I

- Agente: percepción + actuación.
 - » Agente racional: persigue éxito en sus objetivos
 - » Ejemplos de agentes racionales: descripciones PAGE
 - Sistema de diagnosis médicas.
 - Percepciones: síntomas, respuestas pacientes
 - Acciones: tratamientos, pruebas, preguntas
 - Objetivos: la salud del paciente, minimización de costes
 - Entorno: el paciente, el hospital
 - Controlador de una refinería.
 - Percepciones: temperatura, presión
 - Acciones: abrir, cerrar válvulas, ajustar temperatura
 - Objetivos: maximizar la pureza, seguridad
 - Entorno: una refinería

Agentes, II

- Un tutor interactivo de inglés.
 - Percepciones: palabras escritas
 - Acciones: crear ejercicios, sugerencias, correcciones
 - Objetivos: maximizar las calificaciones del alumno
 - Entorno: un conjunto de estudiantes
- Un taxista automático.
 - Percepciones: cámaras, velocímetro, micrófono
 - Acciones: mover el volante, acelerar, frenar, hablar al pasajero
 - Objetivo: seguridad, rapidez, legalidad, comodidad
 - Entorno: carreteras, autopistas

Agente simple

- Algoritmo:

función agente-simple (percepción)

“return” una acción

memoria ;;;; variable global

memoria = actualiza-memoria(memoria, percepción)

acción = elige-mejor-acción(memoria)

memoria = actualiza-memoria (memoria, acción)

“return” acción

- Observaciones:

- » Se debería mantener una secuencia de percepciones en memoria (no siempre es posible)
- » Se deberían tener criterios externos de éxito (la medida del desempeño no es parte de la descripción del agente)

Agente de búsqueda en tablas

- Implementa un *mapa ideal de comportamiento*
- Dificultad en conseguir que el agente razone
- Algoritmo:

función agente-de-búsqueda-en-tablas (percepción)

“return” una acción

.....
,,,,,,,,,,,,,

variables globales

percepciones ;lista inicialmente vacía

tabla ;indexada por lista de
percepciones

añade-percepción-al-final-de-

percepciones(percepción, percepciones)

acción = busca(percepciones, tabla)

“return” acción

- Inconvenientes:
 - » Tabla enorme (y difícil de generar).
 - » Agente sin autonomía (si el entorno cambia, la tabla también debería hacerlo).

Agente reflejo simple

- Es imposible construir una tabla de búsqueda para cualquier secuencia de percepciones.
- Incorporan el uso de reglas:
 - » Si “coche-de-enfrente-frena” entonces “frenar”
 - (resume información y especifica comportamiento)
- Algoritmo:
 - función agente-reflejo-simple (percepción)
 - “return” una acción
 - reglas ;;;; variable global
 - ;;; se genera descripción abstracta de la percepción
 - estado = interpreta (percepción)
 - ;;; selecciona regla que cumple descripción estado
 - regla = selecciona-regla (estado, reglas)
 - acción=aplica-regla(regla)
 - “return” acción
- El uso de reglas ayuda en la toma de decisiones del agente

Un agente reflejo con estado

- Se mantiene un estado del mundo (configurado por las percepciones).

- Algoritmo:

funcion agente-reflejo-con-estado (percepción)

“returns” una acción

estado, reglas ;;;variables globales

;;; se actualiza la descripción interna del estado que

;;; mantiene el agente

estado = actualiza-estado (estado, percepción)

regla = selecciona-regla (estado, reglas)

acción=aplica-regla(regla)

estado = actualiza-estado (estado, acción)

“return” acción

Otros agentes

- Agente basado en el objetivo
 - » La información del objetivo es relevante
 - » Combinación de información sobre objetivo con información de sus posibles acciones
 - Ejemplo: En una intersección, es posible “izquierda”, “derecha” o “recto”, pero importa el objetivo (no es únicamente un agente reflejo).
 - » Búsqueda y Planificación (campos de IA).
 - Caso simple: Acción implica objetivo
 - Caso complejo: Varias acciones implican objetivo
- Agentes basados en utilidad:
 - » Utilidad(estado)=grado de preferencia o satisfacción para un estado. N° real.
 - » Permiten tomar decisiones racionales en 2 situaciones
 - Cuando hay varios objetivos en conflicto
 - Cuando hay varios objetivos alcanzables pero ninguno con certidumbre
 - » Los agentes basados en el objetivo dan una solución pero puede ser un camino no muy satisfactorio.
 - » Ejemplo: Programas de juego. *Dilema del prisionero*.

Tipos de entorno

- **Accesible/no-accesible:**
 - » ¿el agente tiene acceso al estado completo del entorno?
 - Si: no necesita mantener un estado interno del mundo. Uso de sensores. (aspiradora)
 - No: (*8-puzzle*)
- **Determinista/no-determinista:**
 - » ¿el estado siguiente del entorno está completamente determinado por el estado actual y la cadena de acciones elegidas por el agente?
 - (D: *8-puzzle*; ND: control refinería)
- **Episódico/no-episódico**
 - » Episodios: procesos percepción-acción independientes en el tiempo
 - (E: robot selector de componentes; NE: ajedrez)
- **Estático/dinámico/semidinámico:**
 - » ¿el entorno puede cambiar mientras el agente decide?
 - Semidinámico: estático + penalización por retraso en la toma de decisión (ajedrez, análisis imágenes)
- **Discreto/continuo**
 - » ¿número finito de percepciones/acciones?
 - (D: ajedrez; C: robot taxista)

Sistemas multiagentes

- Inteligencia artificial distribuida (IAD) :
 - » Parte de la Inteligencia Artificial
 - » Objetivo: estudiar los comportamientos inteligentes colectivos que son producto de la cooperación de diversas entidades denominadas agentes.
- En un sistema o entorno con múltiples agentes, se necesita que entre ellos exista:
 - » Coordinación
 - » Comunicación
 - » Negociación
 - » Etc.
- **Sistema Multiagente:** MAS, *Multiagent System*.
 - » Objetivo: estudiar la coordinación de la conducta inteligente entre un conjunto de agentes inteligentes autónomos.

Sistemas multiagentes

- Aplicaciones de sistemas multiagentes y de la IAD:
 - » Comercio electrónico.
 - » Optimización de procesos de producción industrial.
 - » Análisis de los procesos de negocio entre empresas.
 - » Monitorización y administración de redes de telecomunicación en tiempo real.
 - » Investigación de los aspectos sociales de la inteligencia y simulación de fenómenos sociales complejos.