Fundamentos de Inteligencia Artificial

Tema 1

Manuel Ramos Cabrer

Tema 1: Introducción

- Concepto de Inteligencia Artificial
- Perspectiva histórica
- Relación con otras disciplinas
- Paradigmas y ámbitos de aplicación
- Conceptos de sistema basado en conocimiento y sistema experto
- Ingeniería del conocimiento
- Perspectiva integradora: Agentes inteligentes
- Estado del arte

Inteligencia Artificial:

- "Rama de las ciencias de la computación que trata del diseño de sistemas inteligentes, es decir, sistemas que presentan características que normalmente asociamos con la inteligencia humana: comprensión del lenguaje, razonamiento, aprendizaje, resolución de problemas, etc".

(Barr y Feigenbaum, 1981)

- "El estudio de cómo hacer que los ordenadores realicen tareas cognoscitivas que, por ahora, las personas hacen mejor".

(Rich y Knight, 1991)

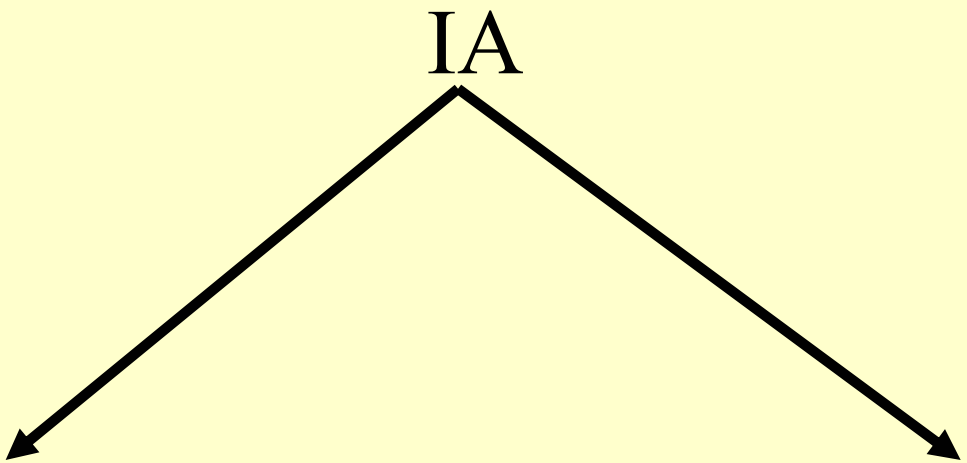
- Definición desde el punto de vista ingenieril: definición en base a las técnicas utilizadas.

Clasificación de definiciones:

- Chariak y McDermot: énfasis en *pensamiento*
  criterio: inteligencia ideal, o *racionalidad*
- Rich y Knight: énfasis en *comportamiento*
  criterio: inteligencia *humana*

	Inteligencia humana	Racionalidad
Razonamiento	Sistemas que piensan como humanos: *Enfoque cognoscitivo*	Sistemas que piensan racionalmente; *Enfoque lógico*
Comportamiento	Sistemas que actúan como humanos: *Enfoque "test de Turing"*	Sistemas que actúan racionalmente: *Enfoque de agentes racionales*

Inteligencia Artificial: dualidad



"INTERFERTILIZACIÓN"

Ejemplos:

⟶ Modelo de procesador cognoscitivo

⟵ Modelo de red semántica

Evolución histórica: Prehistoria

- Autómatas
  - Grecia (época alejandrina):
    - **Herón**: actores artificiales que representaban la Guerra de Troya**.**
    - Mitología: Dios **Talos**, robot hijo de Dédalo que defendía Creta.
  - Romanos y bárbaros:
    Ninguna aportación interesante.
  - Edad Media:
    - **Mayordomo** de San Alberto Magno.
    - **León** de Leonardo DaVinci.
    - **"Ars Magna"** de Ramón Llull (1232 - 1315) : Intento de mecanización del razonamiento.
  - Siglo XVIII:
    - **Droz** (padre e hijo): androides de tamaño natural:
      - Niño que escribía mensajes de hasta 40 caracteres.
      - Niño que realizaba dibujos sobre un papel.
      - Mujer que tocaba el órgano real y movía la cabeza.
    - **Vaucouson**:
      - Tocador de flauta.
      - Pato que movía las alas, emitía sonidos, comía y hacía la digestión.

Evolución histórica: Prehistoria

- Máquinas automáticas
  - Pascal
  - Leibnitz: Sistema binario.
  - Boole: Pretendía un álgebra para representar el conocimiento..
  - Primeros fraudes: Von Kempelen ("Mazel Chess Automaton")
  - 1943: Punto de inflexión:
    - Aparecen tres artículos sobre "Cibernética":
      - Wiener, Rosemblueth y Bigelow (MIT).
      - McCulloch (Illinois) y Pitts (MIT).
      - Clairk (Cambridge).
    - A partir de este momento se usa el lenguaje psicológico en la descripción de máquinas de manera habitual.
  - Años 50: se pretenden llevar estas ideas a la práctica ya que se considera que el avance de la técnica ya es suficiente:
    - Newell, Shaw y Simon: trabajan en el "General Problem Solver" (GPS).
    - Samuel: trabaja en una máquina de damas que aprendía de la experiencia.
    - Minsky: Geometría plana: trabaja en una maquina para razonar sobre figuras
    - Test de Turing (1950)
  - CONFERENCIA DE DARMOUTH (Verano de 1956):
    - Se define una rama del saber a la que ponen el nombre de "Inteligencia Artificial"

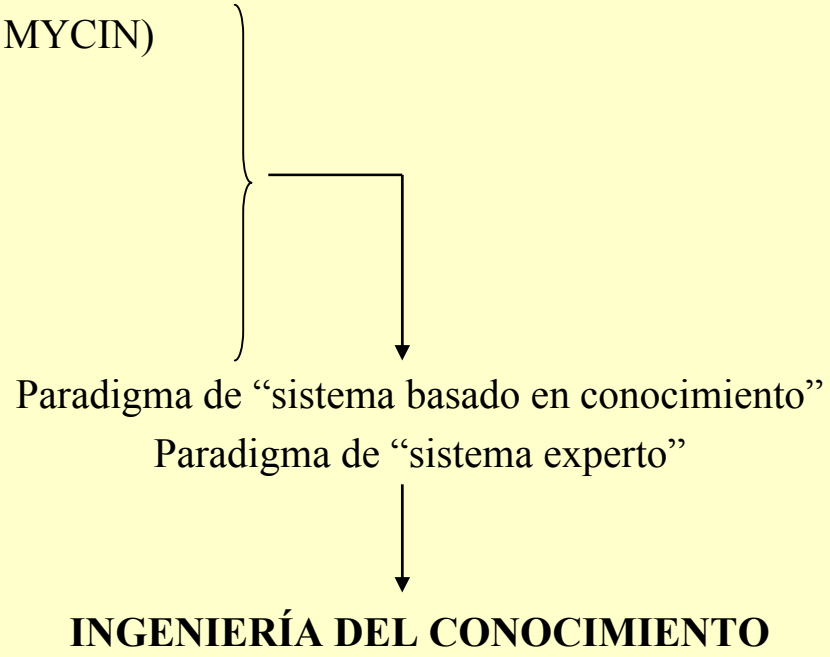
Evolución histórica: Años 60

Se intenta llevar a cabo los objetivos planteados en la conferencia de Darmouth:

- Áreas de aplicación
  - Traducción automática
  - Juegos
  - Reconocimiento de formas
  - Integración simbólica
  - Demostración de teoremas
  - Resolución general de problemas (GPS)
- Técnicas
  - Búsqueda heurística
  - Procesamiento simbólico (LISP)

Evolución histórica: Años 70

- Trabajos
  - Stanford (DENDRAL, MYCIN)
  - M.I.T.
  - Carnegie - Mellon
  - Cambridge
  - Edimburgo
  - ...



- Críticas a la IA
  - Hubert Dreyfus: “Alquimia e IA” (1972)
  - Informe Lightill (1973)
  - John Weizenbaum: “Potencia del ordenador y razón humana (1976)

Evolución histórica: Años 80

- 1981: Conferencia de Tokyo
  - MITI → Robots
  - MITI → Superordenadores
  - MITI → Quinta generación
- Reacciones en EE.UU.
  - MCC (Microelectronics and Computer technology Corporation)
  - SCS (Strategic Computing and Survivability)
  - MCNC (Microelectronics Center of Nort Carolina)
  - Proyecto “Guerra de las galaxias”
- Reacciones en Europa.
  - Programas ALVEY, ESPRIT, EUREKA, ...

Áreas de aplicación de la IA:

- Lenguaje Natural
- Robótica
- Visión artificial
- Reconocimiento de formas
- Sistemas expertos
- Aprendizaje

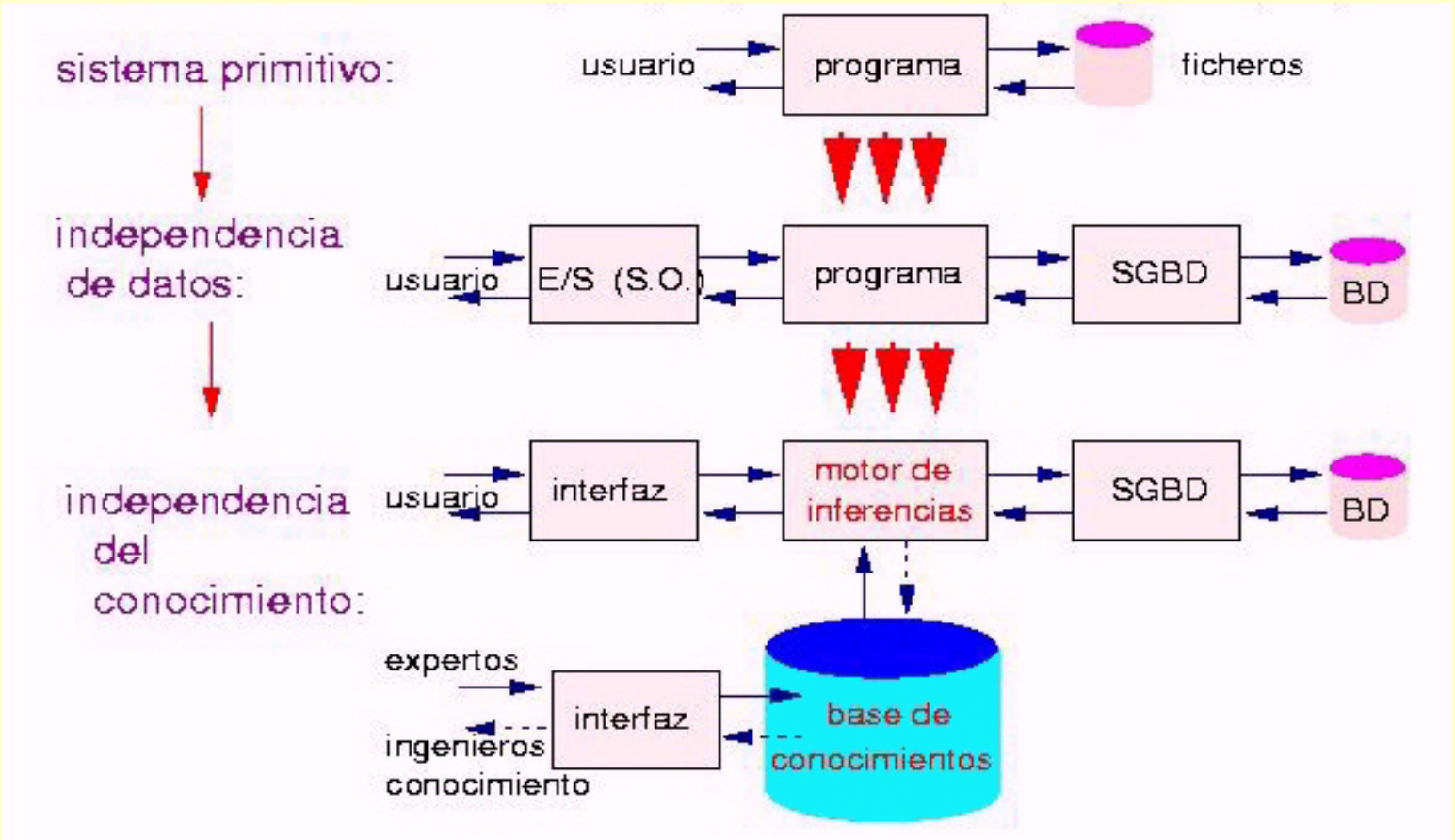
Tratamiento heurístico
de información simbólica

Paradigmas de la Inteligencia Artificial:

- **PROCESAMIENTO SIMBÓLICO:**
  problemas de naturaleza no numérica.
- **BÚSQUEDA HEURÍSTICA:**
  hacer tratable la complejidad del proceso
- **SISTEMAS EXPERTOS:**
  visión conductista, funcional.
- **SISTEMAS BASADOS EN CONOCIMIENTO:**
  visión cognoscitiva, estructural
- **AGENTES INTELIGENTES:**
  "sociedades" de componentes con "personalidad"

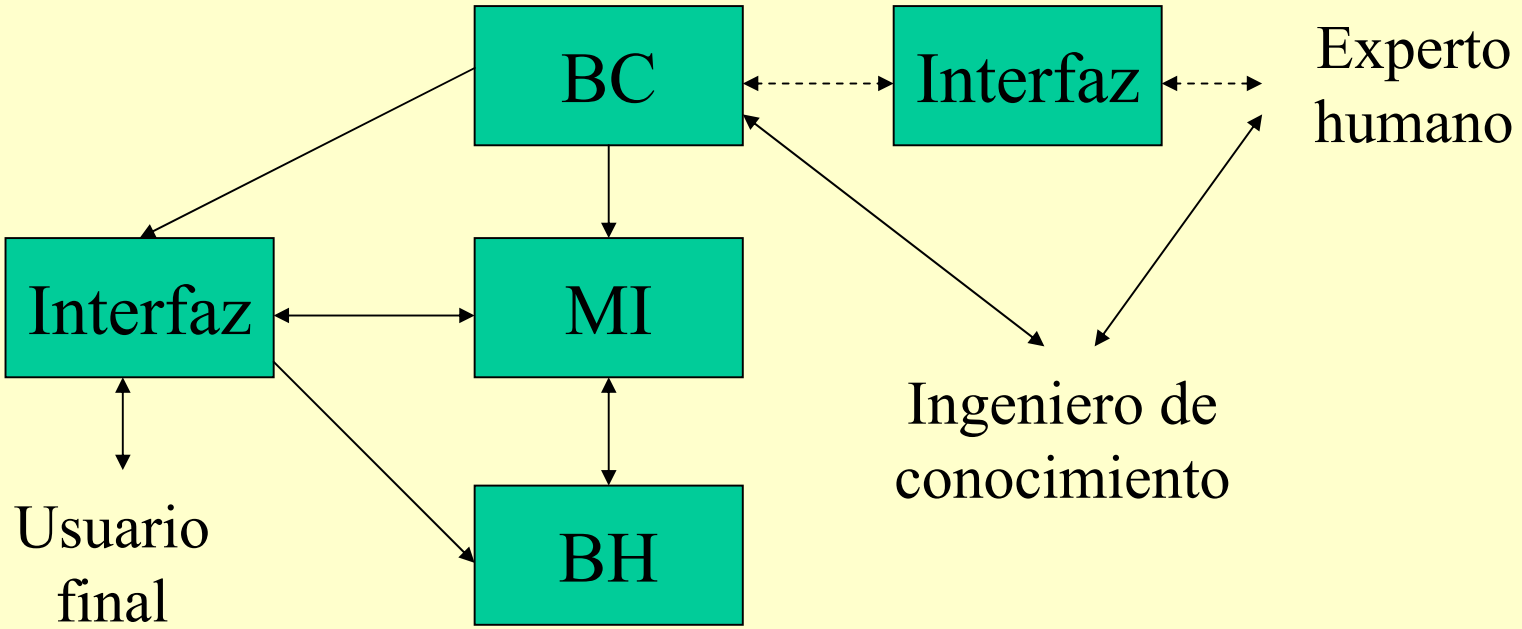
Sistema basado en conocimiento:

Evolución de modelos estructurales de sistemas software

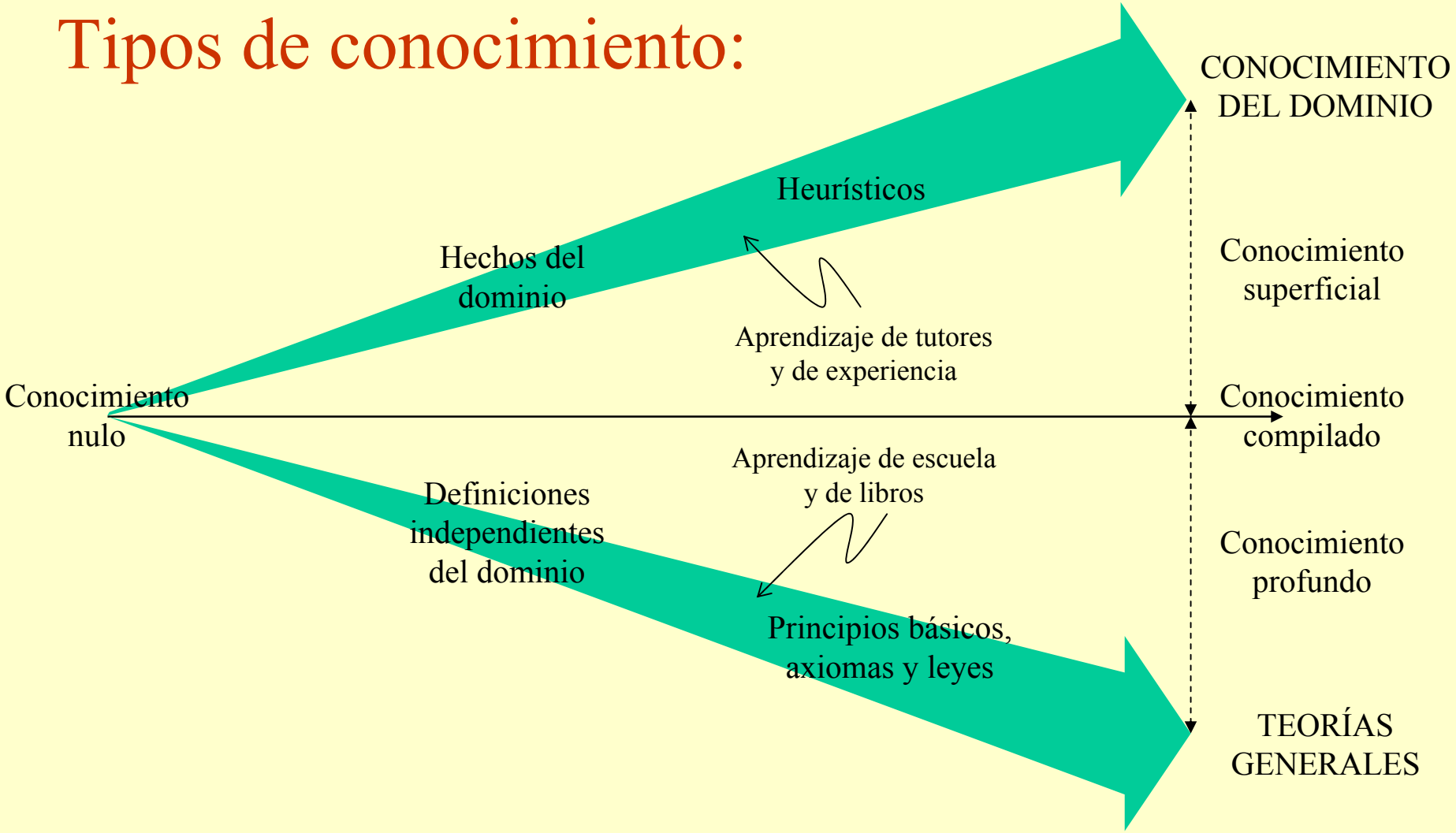


Sistema basado en conocimiento:

Arquitectura básica:



Tipos de conocimiento:



(De Harmon y King, 1985)

Sistema experto:

- "Incorporación en un ordenador ... de la pericia de un experto de modo que el sistema pueda dar **consejos inteligentes** o tomar **decisiones inteligentes** y **justificar** su razonamiento".

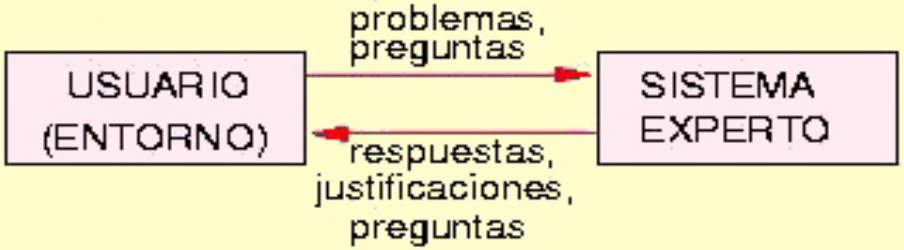
(British Computer Society)

- Sistema con capacidad para:
  - Problemas muy difíciles.
  - Razonamiento heurístico.
  - Interacción fácil.
  - Descripciones simbólicas.
  - Conocimiento impreciso.
  - Hipótesis alternativas.
  - Explicación de preguntas.
  - Justificación de conclusiones.

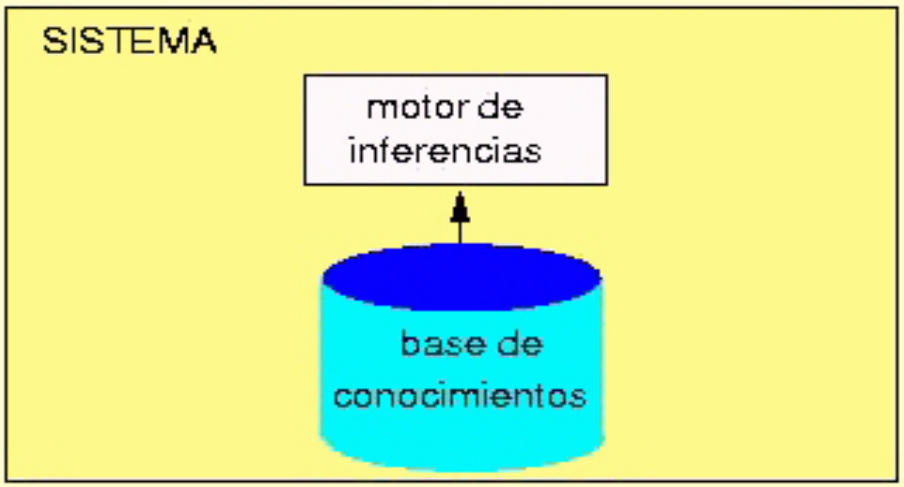
(Hayes-Roth, 1984)

SS.BB.C. vs. SS.EE.:

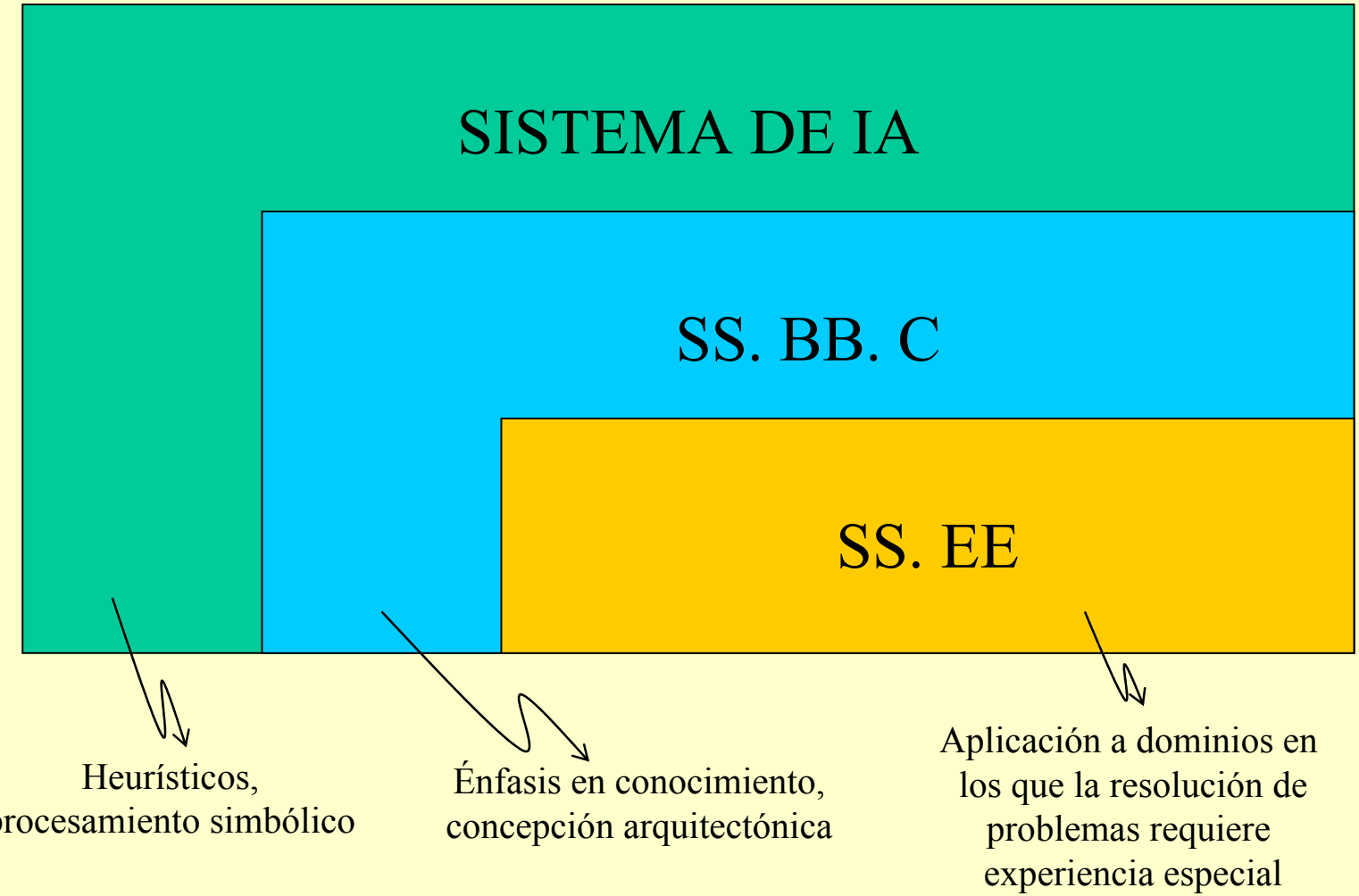
SE: punto de vista funcional (conductista)



SBC: punto de vista estructural (cognoscitivo)



I.A., SS.BB.C. y SS.EE.:



Áreas susceptibles de aplicación de los SBC:

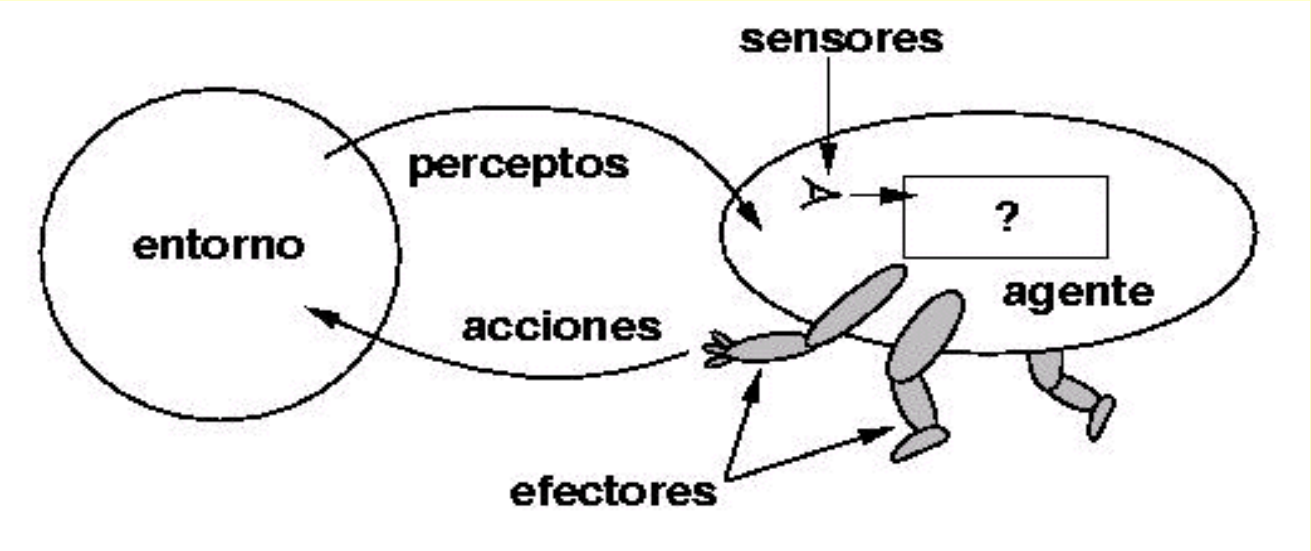
- Punto de vista de los **conocimientos** aplicados:
  - Son muchos y se pueden explicitar.
  - Son evolutivos.
  - Son inciertos e incompletos.
- Punto de vista de la **aplicación** de los conocimientos a la resolución de problemas:
  - Procedimientos más heurísticos que algorítmicos.
  - Procesamiento simbólico domina sobre el numérico.
  - Incertidumbre en los métodos.
- Punto de vista de la **rentabilidad**:
  - Inversiones fuertes.
  - Pocos expertos humanos.
  - Coste alto de errores.

Ingeniería del conocimiento:

Diseño y desarrollo de sistemas basados en conocimiento, combinando:

- Métodos de inferencia simbólica.
- Técnicas de representación del conocimiento.
- Técnicas de concepción de interfaces.
- Conocimiento del dominio.
- Técnicas de comunicación humana.

Paradigma de la “agencia”:



Agente: Sistema que actúa (eventualmente, por encargo) para producir ciertos efectos en su entorno.

Agencia: Cualidades (funcionales, estructurales y/o procesales) propias de los agentes.

Modelo funcional de un agente:

```
Public interface Agente {
      Accion getAccion (percepto p)
}
```

Evaluación: Medida del desempeño (*performance)*.

Agente racional ideal: dado un percepto genera la acción que maximiza el desempeño en función de:
(1) La secuencia de perceptos anteriores
(2) El conocimiento incorporado en el agente

Autonomía: El agente es más autónomo en la medida en que su comportamiento se basa más en (1) y menos en el valor inicial de (2)

Otras definiciones de “agente autónomo”:

- **Jennings y Wooldridge (1996)**

  “Realiza su trabajo sin intervención directa de humanos u otros agentes y tiene control sobre sus propias acciones y su estado interno”

- **Franklin y Graesser (1996)**

  “Sistema *situado* en y como parte de un entorno que percibe y sobre el que actúa en el tiempo siguiendo su propia agenda de tal modo que sus actos influyen sobre lo que percibe en el futuro”

Estado del arte:

¿Cuáles de las siguientes tareas puede hacer una máquina hoy en día?

- Jugar una partida aceptable de tenis de mesa
- Conducir por una carretera de montaña llena de curvas
- Conducir por el centro de El Cairo
- Jugar una partida aceptable de cartas
- Descubrir y probar un nuevo teorema matemático
- Escribir una historia divertida
- Dar consejos de manera competente en una especialidad médica
- Traducir de inglés hablado a español hablado en tiempo real