

Introducción a la Ingeniería del Conocimiento

Juan Luis Castro Peña

Preguntas a responder del Tema

- ¿Qué es un Sistema Basado en el Conocimiento?
- ¿Para qué problemas es adecuada esta metodología?
- ¿Cómo se desarrolla un SBC?
- ¿Cuál es la tarea de un ingeniero del conocimiento?
- ¿Cuál es la relación de la IC con las otras asignaturas de la rama?

La sociedad del conocimiento

- Cumbre de la UE en Lisboa (2000)
- Conocimiento para obtener beneficios económicos y sociales
- Conocimiento como un valor en sí mismo.
- Renovado interés por analizar y manipular el conocimiento

Datos, Información y Conocimiento

- Datos: mínima unidad semántica, por sí solos irrelevantes, no dicen nada sobre el por qué de las cosas y no útiles para la acción.
- Información: datos procesados y que tienen un significado (relevancia, propósito y contexto).
 - Contextualizando
 - Calculando
 - Categorizando:
 - Condensando:
- Conocimiento: experiencia, valores, información y *know-how* que sirve para la incorporación de nuevos hechos e información, es útil para la acción.
 - Comparación.
 - Predicción.
 - Conexiones.

¿Qué es conocimiento?

- Datos: ... --- ...
- Información: S.O.S. Significado
- Conocimiento: Operación de salvamento.

Propósito y competencia.

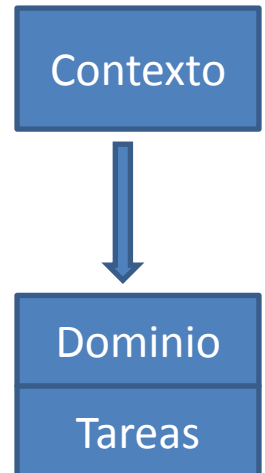
Capacidad para generar una reacción

Específico de un problema complejo.

Fuente vaga e incompleta.

No solución directa y clara

Intuición, experiencia, no solo libros.



Sistemas Expertos (SE)

- Sistemas que utiliza **conocimiento experto** (proporcionado por un humano) para resolver un **problema complejo** :
 - Suelen interaccionar con el usuario durante la resolución del problema
 - Justifican la solución
 - Conocimiento revisable y adaptable
- Pretende comportarse como un experto

Sistemas basados en el Conocimiento

- La ingeniería de conocimiento produce SBC.
- SBC: Sistema que usa conocimiento específico del dominio del problema.
- Sistema Experto: Tareas que requieren **razonamiento** humano.
- Conocimiento representado explícitamente de forma separada (**Base de Conocimientos**).
- Funcionamiento no algorítmico, suele incluir **estrategias y estructuración** (metaconocimiento)

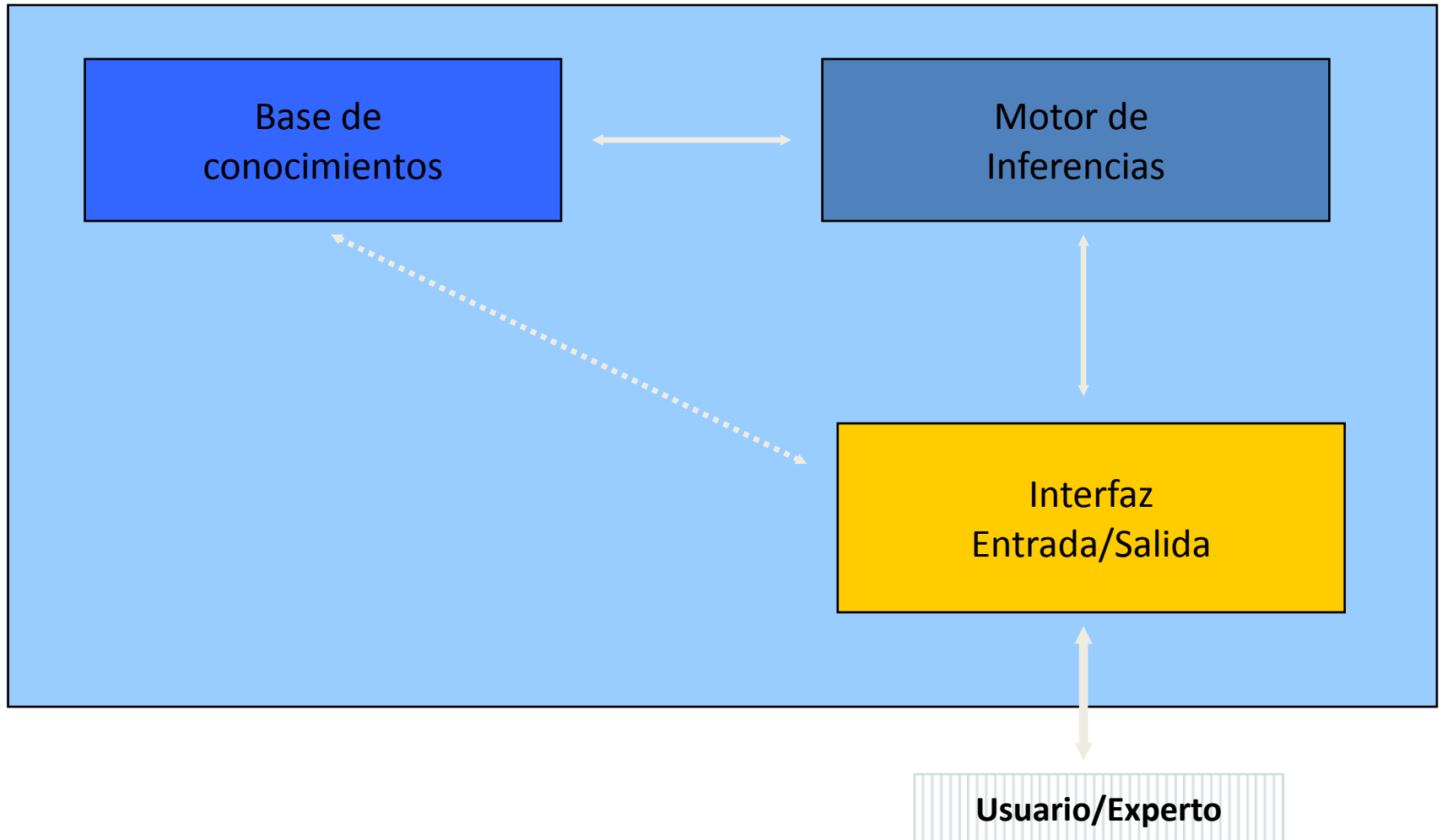
Sistemas basados en el Conocimiento

- Sistema software capaz de soportar la representación explícita del conocimiento de un dominio dado específico y de explotarlo a través de los mecanismos apropiados de razonamiento para proporcionar un comportamiento de nivel alto en la resolución de problemas

Problemas adecuados para SBC

- Problemas poco estructurados en los que nos podemos encontrar:
 - requisitos subjetivos,
 - entradas inconsistentes, incompletas o con incertidumbre,
 - que no pueden ser resueltos aplicando algoritmos clásicos o la investigación operativa.
 - Se dispone de fuentes de conocimiento

Esquema de un Sistema basado en el Conocimiento



Áreas de aplicación

Ejemplos de áreas de aplicación:

- Configuración
- Diagnóstico
- Enseñanza
- Interpretación
- Planificación
- Predicción
- Control

Ventajas

- Disponibilidad:
 - El SBC está disponible para cualquier hardware de cómputo adecuado y coste reducido
- Permanencia:
 - El SBC funciona permanentemente,
- Experiencia múltiple:
 - El conocimiento de varios especialistas puede estar disponible para trabajar simultánea y continuamente en un problema.
 - El nivel de experiencia combinada de muchos SBC puede exceder el de un solo especialista humano.

Ventajas

- Respuestas no subjetivas
 - Un SBC ofrece respuestas sólidas, completas y sin emociones en todo momento.
- Explicación del razonamiento
 - Un SBC puede explicar clara y detalladamente el razonamiento que conduce a una conclusión
- Adaptable y mejorable sin necesidad de IC
 - Solo hay que modificar el conocimiento que es algo editable por cualquier usuario

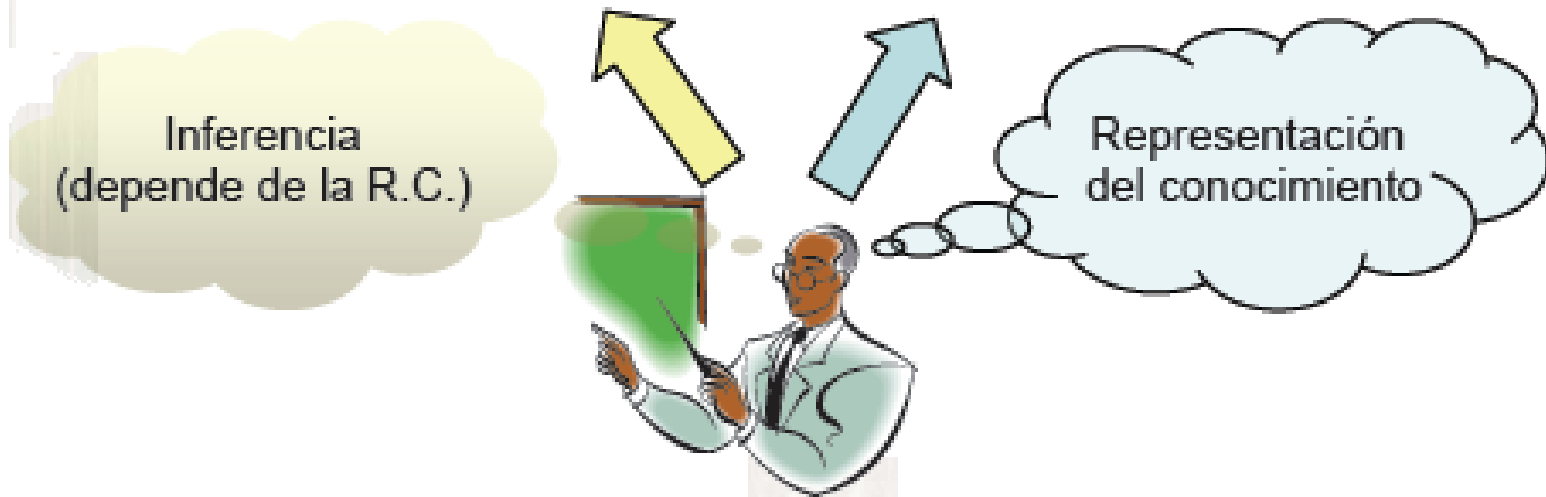
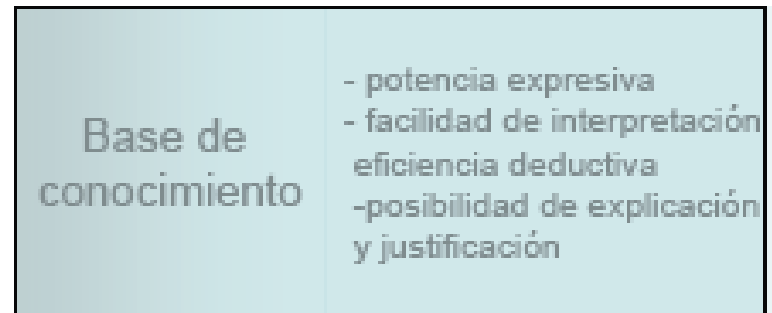
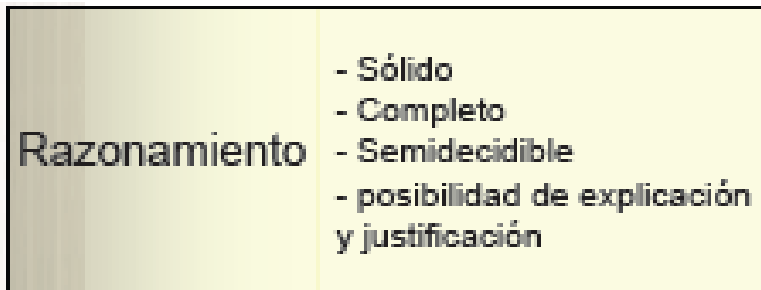
Ingeniería de Conocimiento

- Proceso de adquirir, estructurar, formalizar y hacer operativos un conjunto de conocimientos en un programa (SBC) que resuelva una tarea compleja adecuadamente.
- Importante por:
 - Conocimiento tiene valor por sí mismo y sobrevive a implementaciones.
 - Los errores en el conocimiento son decisivos
 - Facilita escalabilidad y mantenimiento.

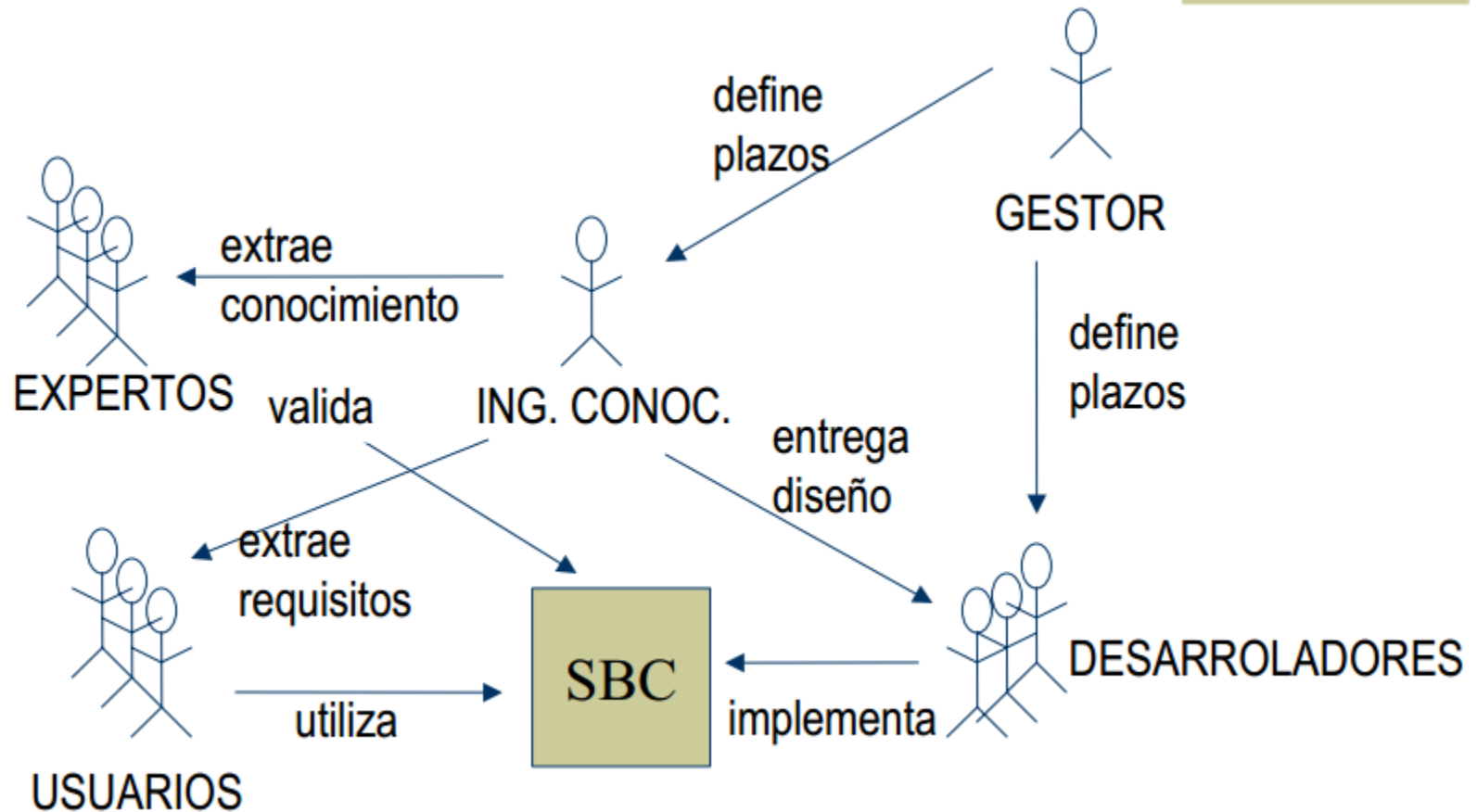
Problemas abordados por la IC

- La **adquisición del conocimiento** y cómo “almacenar” el conocimiento humano mediante una representación abstracta efectiva.
- La **representación del conocimiento** en términos de una estructura de datos que una máquina pueda procesar.
- Los **sistemas de razonamiento** o cómo hacer uso de esas estructuras abstractas para generar información útil en el contexto de un caso específico.

Tareas del ingeniero de conocimiento



Actores del desarrollo de un SBC



Tareas del ingeniero de conocimiento

- Los algoritmos de inferencia son altamente independientes del problema, pero dependientes del formalismo de representación del conocimiento.
- La Base de Conocimiento suele ser totalmente dependiente del problema:
 - Investigar el dominio del problema.
 - Aprender qué conceptos son importantes en tal dominio.
 - Obtener una representación formal de tales conceptos, y cómo se relacionan.
 - Decidir un formalismo de representación.
 - Reutilizar conocimiento.

Dificultades para el desarrollo de un SBC

- Conocimiento utiliza variables complejas (a veces mas que el resto del problema)
- Conocimiento difícil de representar
- Adquirir conocimiento proceso arduo y difícil
 - Conocimiento incompleto
 - Conocimiento presupuesto
 - Conocimiento inconsistente
- Sistema de inferencia no reproduce lo que se espera

Problemas a evitar

- Ingeniero del conocimiento se mete a experto
- Experto se mete a ingeniero del conocimiento
- Experto no entiende bien el objetivo (cree que el IC es un programador a su servicio)
- IC cambia SBC por algoritmo

Ciclo de vida tradicional para el desarrollo de una BC

- 1) **Identificar** la tarea, análisis de viabilidad e impacto
- 2) **Adquirir conocimiento** (con experto y consultas documentación)
- 3) **Conceptualizar**: Estructurar conocimiento en conceptos y tareas, crear una ontología del dominio (modelo conceptual)
- 4) **Formalizar** el conocimiento general acerca del dominio (modelo formal)
- 5) **Implementar** formalización (con desarrollador)
- 6) **Verificar y Validar** funcionamiento esperado (con usuario y experto)

Ejemplos históricos

- **DENDRAL 1965:** identificación estructura moléculas orgánicas
- **MYCIN 1970:** identificar infecciones en sangre, introduce Factores de Certeza
- **PROSPECTOR 1974:** prospección de minerales, introduce nuevos métodos para tratamiento incertidumbre
- **XCON 1980:** asistir a la compra de sistemas de computación VAX

Precedentes y Evolución

- Los 60: Sistemas de propósito general (GPS)
- Los 70: Sistemas basados en reglas (Mycin)
- Los 80: Adolescencia de metodología (KADS)
- Los 90: Madurez de metodología (CommonKADS)
- Este siglo: Conceptualización del Conocimiento (Ontologías)

Esquemas de Sistemas Inteligentes

Problema sin solución algorítmica eficiente



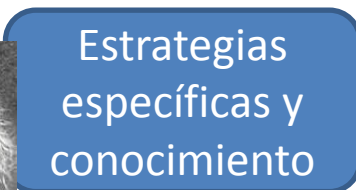
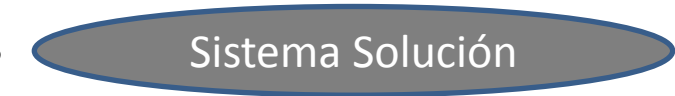
Aprendizaje Automático

Adquisición de conocimiento

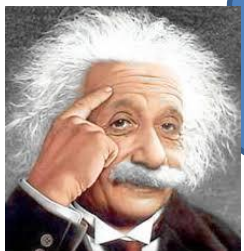


Metaheurísticas

Técnicas de sistemas inteligentes



Ingeniería del conocimiento



Tareas de la semana

- Describir el conocimiento propio para clasificar un email como SPAM
- Plantearse cual sería la mayor dificultad para hacer un SBC para decidir que un correo es SPAM
- Leer el documento Tutorial Clips que está en decsai