

Tema 2: Adquisición de Conocimiento

Anabel Gómez Ríos

20 de abril de 2016

¿En qué consiste la adquisición del conocimiento?

¿Cuándo se realiza este proceso?

¿De qué fuentes se suele adquirir el conocimiento y en qué medida?

¿En qué consiste el proceso de adquisición del conocimiento?

¿Cómo se realiza la fase del estudio del problema? ¿Qué debemos obtener de esta fase?

¿Qué es la educación del conocimiento? ¿Qué técnicas podemos usar?

¿En qué fases se divide esta parte del proceso? ¿Qué debemos obtener de cada fase?

1. Introducción

La *adquisición de conocimientos (AC)* es el proceso de recolección de información, a partir de cualquier fuente (experto, libros, revistas, informes,...), necesaria para construir un Sistema Basado en Conocimiento. El objetivo es proporcionar a cada etapa la información que se requiere en cada momento del desarrollo. No es un paso concreto en la metodología de desarrollo de un SBC, sino más bien una tarea que se produce en paralelo a todas las etapas de construcción de estos sistemas (identificación, conceptualización, formalización, validación, mantenimiento, ...). Forma parte de cada fase del desarrollo de un SBC, no es un proceso aislado.

1.1. Situación Actual

La adquisición de conocimientos es la tarea más importante en el desarrollo de SBC. Pocas técnicas propias de la IA son aplicables. No existe ningún método completamente automatizado de Adquisición de Conocimiento. Existen unas metodologías para adquirir correctamente el conocimiento.

2. Fuentes de conocimiento

2.1. Personas

1. Expertos, de los que se extrae la mayor parte de los conocimientos.
2. Directivos, de los que se extraen los objetivos del proyecto, el alcance del sistema, el contexto donde será instalado,...

3. Usuarios, de los que se busca comprender el tipo de persona que interactuará con el sistema, sus necesidades, requisitos...

2.2. Documentos

1. Resultados de estudios e investigaciones: datos, estudios, informes, resultados estadísticos, etc.
2. Publicaciones especializadas: suelen contener las versiones más actualizadas.
3. Material utilizado para formación: suele contener conocimientos expuestos de un modo especialmente claro.
4. Casos: Algunas empresas suelen registrar los casos que se les presentan en forma de órdenes de reparación, fichas de clientes o pacientes, estudios o almacenamiento de casos.

2.3. Experiencia

Bien por observación (visitas), que es útil para clarificar las ideas y entender el esquema general del proceso o bien por documentos informales:

1. Notas manuscritas, ayudas de trabajo, etc., que circulan dentro de las organizaciones.
2. Proporciona, a menudo, conocimiento heurístico para resolución de problemas.
3. Suelen reflejar la experiencia de profesionales de la organización a la hora de enfrentarse a ciertos problemas.
4. Proporciona conocimiento semipúblico.

3. Extracción y educción

Se llama extracción del conocimiento cuando la fuente de conocimiento se presenta en forma escrita o registrada y se llama educción de conocimiento cuando los conocimientos se obtienen de humanos.

Proceso de adquisición de conocimiento:

1. Estudio del problema: primeras reuniones y evaluación de viabilidad (de expertos, directivos y usuarios).
2. Extracción de conocimientos (de la documentación).
3. Educción de conocimientos (del experto), que consta de un interrogatorio inicial y de una investigación profunda.

3.1. Estudio del problema

Consiste en determinar los requisitos funcionales del sistema o, en su caso, las necesidades de los usuarios del futuro sistema, o lo que los usuarios esperan del mismo (es decir, entradas, salidas e interacción con el sistema). Hay que introducir al Ingeniero del Conocimiento (IC) en el dominio a un nivel tal que sea capaz de desarrollar un estudio de viabilidad del Sistema Experto (SE) donde se determine si la tarea del experto es tratable o no, mediante la Ingeniería del Conocimiento (esquema general del proceso).

Primeras reuniones: Se buscan conocimientos generales, no de detalle, así como familiarizarse con la terminología del dominio. El ingeniero debe centrarse en el entorno de la tarea y sus usuarios. Hay que dirigir la entrevista a fijar primero las entradas, salidas e interacciones del sistema. Posteriormente hay que conseguir el esquema general del proceso de razonamiento en forma de tareas, subtareas y las relaciones y dependencias entre ellas. Hay que evitar a toda costa que nuestras opiniones e ideas interfieran en los resultados (error muy común que ralentiza esta fase). Una o dos reuniones deben ser suficiente.

Análisis de viabilidad:

1. ¿Es factible obtener los datos que se usan de entrada? ¿Cuánto coste tendría? → Posible modificación: de calculados (demasiado complejo o costoso) a introducidos por el usuario.
2. ¿El esquema general del proceso de razonamiento se puede representar? → Solución más común: sistema modular de tareas.
3. ¿Cada tarea se podría representar como un SBC simple mediante un modelo de representación del conocimiento básico? → Habría que reiterar el proceso para cada tarea que no resulte básica.

3.2. Extracción del conocimiento

Aprender sobre el dominio tanto como sea posible antes de comenzar las sesiones con el experto. Se pretende evitar o, por lo menos, reducir el tiempo que, de otro modo, debería dedicar el experto a fin de iniciar al ingeniero en el tema.

Consulta con el experto: Indique los documentos más importantes a consultar. Indique el material relevante de la colección de manuales. Explique la terminología usada y los conceptos discutidos en la documentación. Proporcione detalles omitidos en los documentos, pero cruciales para un entendimiento global de la tarea. Señale los puntos donde la práctica real difiere de los procedimientos documentales. Explique las anotaciones a mano hechas sobre los documentos de trabajo.

Análisis estructural de textos:

1. Técnicas en las cuales los términos son determinados por el IC en tiempo de ejecución. Ante un determinado texto, establece una serie de términos que deben ser buscados en el texto.

2. Técnicas en las que los términos a buscar están preestablecidos por la técnica y son dependientes del dominio. (Enfermedad, Medicina, Terapia,...)
3. Técnicas en las que los términos a buscar están también preestablecidos, pero son independientes del dominio. (Se define como, Está relacionado con, Es una característica de, etc.)

Términos independientes del dominio: conceptos y relaciones. Estructuras tipo definición para los conceptos y estructuras tipo afirmación relacional para las relaciones. Necesitamos estructuras a buscar y tipo de conocimiento que se obtendrá con cada una (concepto, relación, característica, valor), y un modo de detectar esas estructuras.

Tipos de estructuras a buscar:

1. Definiciones: Introducción de un nuevo concepto en el texto. El criterio puede venir definido en base a distintos criterios (uso, partes que lo componen,...)
2. Afirmaciones: Una afirmación es una frase que establece una verdad. Para el objetivo de extraer conocimientos básicos, las afirmaciones que interesan son aquellas que expresan relaciones entre conceptos.
3. Leyes: Las leyes de un dominio establecen sus principios básicos, así como las reglas que fijan el funcionamiento de objetos del dominio.
4. Procedimientos: Los procedimientos de un dominio establecen los pasos para la resolución de problemas en el dominio. Al igual que en el caso anterior, los conocimientos proporcionados por esta estructura están más allá del objetivo de la extracción de conocimientos a partir de la documentación.

Modo de detectar las estructuras: Búsqueda de patrones, ejemplos y títulos de secciones, subsecciones y remarcados (negrita, cursiva). El trabajo consistirá en una tarea simple pero tediosa, como por ejemplo ojear el texto deteniéndose sólo en las definiciones y afirmaciones de relación (estas frases serán señaladas) y analizar las frases señaladas para extraer los conocimientos buscados: conceptos, relaciones y definiciones de conceptos.

3.3. Educción del conocimiento

El IC deberá marcar la perspectiva y profundidad deseada en cada sesión. Dada la complejidad intrínseca del proceso, cada sesión debe ser controlada, además, en muchos otros aspectos.

Ciclo de educación:

1. **PREPARACIÓN DE LA SESIÓN. Información a tratar:** se debe elegir una perspectiva o área del dominio y concentrar una serie de sesiones en ella hasta que el IC se encuentre satisfecho con los conocimientos obtenidos. **Amplitud, profundidad:** decidir, para cada sesión, el grado de profundidad y detalle que se necesita. **Técnica adecuada:** Las distintas técnicas de educación poseen un método específico y alcanzan un grado de detalle concreto. El IC deberá decidir qué técnica se adapta mejor al objetivo de cada sesión. **Preparación de preguntas:** El IC debe reflexionar, en base a las sesiones anteriores, qué nueva información necesita, qué temas debe tratar y qué preguntas va a realizar para acometerlos.

2. **SESIÓN. Repaso del análisis de la última sesión:** El IC muestra al experto un breve resumen (notas, diagramas, ...) del análisis de la última sesión. El experto debe corregir o clarificar detalles del resumen. **Explicación al experto de los objetivos de la nueva sesión:** El IC comunica al experto cuáles son los objetivos de esta sesión. **Educción:** El IC procede a aplicar la técnica de educación elegida. **Resumen y comentarios del experto:** El experto añade comentarios o críticas al proceso de educación o aporta detalles que le parecen importantes con respecto al contenido de la sesión.
3. **TRANSCRIPCIÓN.**
4. **ANÁLISIS DE LA SESIÓN. Lectura para obtención de una visión general. Extracción de conocimientos concretos:** se realiza un análisis más formal. El IC busca conocimientos y los estructura en sus componentes importantes. **Lectura para recuperar detalles olvidados. Crítica para mejoras por parte del IC.**
5. **EVALUACIÓN.** Las preguntas a responder en este momento son, entre otras: ¿Se han conseguido los objetivos? ¿Es necesario volver sobre el mismo objetivo? Establecer el número y tipo de sesiones necesarias para cubrir el área. Esta fase del ciclo de educación se funde con la primera de preparación de la siguiente. Es decir, evaluando si se han alcanzado los objetivos, si se necesitan más sesiones sobre el mismo área, etc., se responde a las preguntas que el IC debe plantearse en la primera fase del ciclo.

3.4. Técnicas de educación

1. Métodos directo: le preguntan directamente al experto lo que sabe. El experto es la única fuente de información; el IC confía totalmente en lo que el experto le dice.
2. Métodos indirectos: se usan porque no siempre los expertos pueden acceder a los detalles de sus conocimientos o procesos mentales, y para confirmar lo adquirido mediante técnicas directas.

Técnicas: entrevistas (abiertas y estructuradas), cuestionarios, observación de tareas habituales, incidentes críticos, clasificación de conceptos, análisis de protocolos, emparillado e inducción.

3.4.1. Entrevista abierta

El IC plantea, más o menos espontáneamente, preguntas al experto. Espontáneamente no significa que esta técnica no necesite planificación y control. Como en cualquier sesión de educación, el IC habrá fijado un tema o perspectiva a tratar con el experto, así como una profundidad de los conocimientos a educir. La falta de conocimientos sobre la perspectiva fijada, y el requisito de un grano grueso en el tema, lleva al IC a seleccionar la entrevista abierta como la técnica más adecuada a usar en una determinada sesión de educación.

Preguntas típicas:

1. Iniciales: ¿Cómo resuelves este problema? ¿Cuáles son los elementos que influyen cuando resuelves el problema? ¿Cuáles son las informaciones que necesitas antes de empezar el tratamiento del problema?

2. Alentar: ¿Qué haces a continuación? ¿Puedes describir lo que quieres decir con eso? ¿Por qué haces eso?
3. Alerta: ¿En qué se parece y diferencia este problema con los típicos del dominio? ¿Qué tipos de datos necesita el problema? ¿Qué clase de soluciones son adecuadas para el problema? ¿Qué constituye una explicación o justificación adecuada de la solución del problema?

3.4.2. Entrevista estructurada

El IC planifica todas las preguntas que debe plantear al experto durante la sesión. Debe formular y agrupar las cuestiones lógicamente. Normalmente, los grupos conciernen a acciones o procesos que se han identificado en sesiones previas. El IC plantea un grupo de cuestiones sobre determinado objeto o atributo y, una vez resueltas, pasa al grupo siguiente. Las preguntas a plantear deberían centrarse sobre los conocimientos de los conceptos, relaciones e inferencias del experto.

Preguntas típicas: ¿Qué tipo de cosas te gustaría saber acerca del problema cuando empiezas a sopesarlo? ¿Qué hechos o hipótesis intentas establecer cuando piensas sobre el problema? ¿Cuáles son los factores que influyen en la forma en que razonas en el problema? ¿Qué tipo de valores puede tener ese objeto?, ¿qué rango de valores está permitido? ¿Este valor depende de otros factores? En caso afirmativo, ¿cuáles son? ¿Es este factor necesario para resolver todos los problemas en el dominio, o sólo para algunos? En este caso, ¿cuáles?

¿Podrías explicarme este concepto en mayor detalle? ¿Qué ocurre en este punto? ¿Por qué te planteaste este problema? ¿Es correcta esta secuencia? ¿Es esto lo que haces en esta situación? ¿Te parece que este diagrama muestra correctamente el orden de tus decisiones? ¿Están incluidos aquí todos los conceptos relacionados con tal tema?

3.4.3. Limitaciones de las entrevistas

1. Puntos fuertes: Sirven para compilar conocimientos básicos de la tarea, para obtener la información conceptual implicada en el problema y para extraer conocimientos de relaciones, valores y acciones.
2. Puntos débiles: Las entrevistas son consumidoras de tiempo, sobre del IC. Confían en la memoria del experto y existen problemas con el lenguaje, tales como eliminación de componentes clave en un proceso de razonamiento, palabras con referencias no especificadas, referencias hechas con comparativos, condensación con palabras de procesos complejos o implicación de conexión causal entre eventos.

3.4.4. Observación de tareas habituales

Con frecuencia, la mejor forma de descubrir cómo hace un juicio un experto, efectúa un diagnóstico, o diseña una solución, es observar a un experto trabajar en un problema real habitual. La primera decisión que se debe tomar al respecto es cómo registrar las prestaciones del experto. Una posibilidad es sencillamente observar, tomar notas e intentar seguir al vuelo el proceso de pensamiento del experto. Otra, es grabar todo el proceso para una posterior revisión

con el experto. El IC no interfiere en la actuación del experto en la solución de sus tareas reales cotidianas.

1. Ventajas: Proporciona al IC una primera idea de los tipos de conocimientos y habilidades implicados en el dominio, proporciona conocimiento básicos del dominio y ayuda a que el IC comprenda la tarea del experto y es útil para captar conocimientos procedimentales o para entender las características peculiares de los usuarios del SBC.
2. Inconvenientes: Si las tareas habituales no son muy informativas acerca del razonamiento del experto, el método suministra poco conocimiento, consume mucho tiempo y a veces es inoportuna y fastidiosa.

3.4.5. Otras técnicas

TÉCNICA DE LAS VEINTE PREGUNTAS

El IC presenta un caso inventado al experto. Para la resolución del problema, el experto sólo podrá plantear al IC, veinte preguntas sobre el mismo. De nuevo la técnica sirve para conocer la información relevante que usa el experto en la resolución.

INCIDENTES CRÍTICOS

Se pide al experto que describa casos especialmente interesantes o difíciles que se le hayan presentado. El experto, además, deberá contar cómo los resolvió. Esta técnica es muy interesante, ya que los casos especialmente complejos hacen comentar detalles que, en otro momento, hubiera pasado por alto. Una variante es plantear casos críticos imaginarios. Se le pide al experto que considere un caso cualquiera y se le convierte en un incidente crítico mediante el planteamiento de situaciones alternativas de la forma: ¿qué pasaría si ...?

IMPOSICIÓN DE RESTRICCIONES

Se pretende que el experto verbalice ciertas deducciones que, en un caso familiar, pasaría por alto, imponiéndole limitaciones en el uso de recursos. Los recursos a retringir pueden ser el tiempo de resolución o la información disponible (que falten datos de entrada). Es útil para educir cierto tipo de conocimientos como pueden ser las prioridades de atención por parte del experto.

CLASIFICACIÓN DE CONCEPTOS

En primer lugar se obtiene, a partir de un simple glosario o texto, un conjunto de conceptos que cubran ampliamente el dominio. A continuación, se transfiere cada concepto a una ficha y se le pide al experto que las clasifique en una serie de grupos, describiendo lo que cada grupo tiene en común. Por último, las fichas de cada grupo se comparan para formar jerarquías. Esta técnica es especialmente aconsejable cuando hay un gran número de conceptos en un dominio, de modo que requieren una estructuración para que sean manejables.

CUESTIONARIOS

Consiste en realizar una entrevista estructurada al experto de forma indirecta a través de cuestionarios. Esta técnica tiene la ventaja, frente a las entrevistas, de ser una forma eficiente de acumular información. Los cuestionarios pueden ser particularmente apropiados para describir los conceptos, revelar las relaciones en el dominio y determinar incertidumbre.