

# Mapas Conceptuales como primera fase del Diseño de Bases de Datos

Héctor Gómez Gauchía  
Facultad de Informática  
Universidad Complutense de Madrid  
28040 Madrid  
[hector@sip.ucm.es](mailto:hector@sip.ucm.es)

## Resumen

En nuestra asignatura de Bases de Datos los alumnos no entienden el proceso de diseño de bases de datos fácilmente. Hemos descubierto que es un problema previo a donde empieza el diseño de base de datos habitual, el problema está en la fase de formación de conceptos. Nuestra propuesta consiste en incluir, en el proceso, el uso de mapas conceptuales. Y, así, representar conocimiento declarativo a un nivel muy bajo, usando conceptos relacionados con frases de enlace. Hemos definido un proceso de construcción de mapas específico, una estrategia de enseñanza y algunas técnicas adhoc para evaluación del impacto en los alumnos.

## 1. Motivación

Los enfoques en la enseñanza de las bases de datos están basados, en general, en la creencia de que los alumnos y los diseñadores tienen una idea clara del dominio, sus conceptos y la funcionalidad que necesitan. En nuestra experiencia esta creencia es falsa en muchas ocasiones. El desarrollo profesional de bases de datos se realiza usando las especificaciones del análisis de requisitos. Sin embargo, frecuentemente, las especificaciones no son muy claras y, a veces, simplemente no existen. Entonces, las fuentes alternativas de información son los expertos del dominio, que suelen ser imprecisos al describir el dominio y, como es lógico, no hablan en términos de modelos de bases de datos, tales como *entidades* o *atributos*. En estas situaciones, hemos encontrado muy útil el uso de herramientas que ayuden en el proceso de conceptualización. Nos referimos al proceso de describir explícitamente los elementos que forman

el dominio, i.e.: conceptos y relaciones entre ellos.

Por otra parte, hemos encontrado que, para nuestros alumnos como diseñadores, es difícil entender el proceso de conceptualización necesario para obtener un modelo de datos de la base de datos, p. e.: el diagrama del modelo de Entidad / Relación Extendido (EER) o el modelo de objetos de UML. Los alumnos tienen dudas principalmente con estos temas:

- Qué es un concepto y qué es una relación entre conceptos.
- Entender el significado de cada relación dentro de un dominio desconocido para ellos.
- Diferenciar conceptos, atributos y valores.
- Cómo hacer corresponder los términos del dominio con los elementos de los diagramas EER: entidades, relaciones, claves, cardinalidad, participación o cobertura.
- Encontrar qué relaciones son necesarias para la funcionalidad requerida.
- Cómo navegar a través de los conceptos y las relaciones en el diagrama EER
- Los diagramas EER que obtienen son: ambiguos, incompletos e incluso incorrectos, no reflejando la funcionalidad requerida.

Podemos resumir el problema desde los puntos de vista de los dos participantes: para un experto del dominio, explicar el dominio es una tarea complicada en sí y para los diseñadores de bases de datos, entender un dominio nuevo también es complicado.

Esta situación tiene varias razones concretas. Nos hemos centrado en dos importantes: la primera es que el lenguaje natural es ambiguo y hay una tendencia a usar verbos y conceptos genéricos en la descripción oral del dominio. La consecuencia directa de esta tendencia es que es

difícil pensar y razonar con precisión sobre el dominio usando esos elementos genéricos. La segunda es que los alumnos, durante la conceptualización, están desbordados con los detalles del dominio y tienden a olvidar el propósito de la base de datos. Entendemos por propósito la funcionalidad y los datos que se necesitan para el proyecto a realizar en ese dominio. Es importante el propósito porque es muy diferente el modelo de datos resultante para propósitos diferentes.

La solución propuesta es modificar el proceso de diseño de la base de datos añadiendo fases nuevas para incluir una herramienta que ayude a pensar sobre el dominio, i.e. el proceso de formación de conceptos.

En la siguiente sección presentamos una lista de las teorías que hemos analizado y pueden servir para este propósito. Entre ellas, hemos escogido los mapas conceptuales (cmaps) como la más adecuada. A continuación se expone una breve descripción de los elementos básicos de los cmaps. En la siguiente sección se describe el proceso de construcción de cmaps que hemos creado como paso previo a hacer los diagramas EER. A continuación se presenta un caso de estudio, la estrategia de enseñanza que hemos aplicado a nuestros alumnos, la evaluación informal y los resultados.

## 2. Mejorar la etapa de formación de conceptos

Para la ayuda a pensar sobre un dominio se han desarrollado varias herramientas. Algunas son para la elaboración de ideas abstractas y la agrupación de éstas, como los *Diagramas de Afinidad* o método KJ [10] muy usado en Japón como herramienta para estructurar las tormentas de ideas. Para nuestro objetivo necesitamos representar conceptos a varios niveles de abstracción, incluyendo los de muy bajo nivel, algo que no se puede con esta herramienta.

Otras herramientas, tales como los *mapas cognitivos* se centran en el proceso mental en sí y de cómo el dominio es afectado en su evolución a lo largo del tiempo. Se basan en la teoría de la construcción personal [7], que ha evolucionado hacia los mapas difusos [8]. En nuestro problema no representamos la evolución en el tiempo.

Otras herramientas aunque son conceptuales, están orientadas a resolver problemas, como los

*diagramas en uve* [6]. En ellos, hay una descripción conceptual de un problema sobre un dominio y unos pasos metodológicos para resolver el problema. Aquí la descripción no es detallada ni exhaustiva, solo es general y orientada a explicar el problema. Nosotros necesitamos una descripción del dominio en un nivel de detalle grande.

Otras herramientas se centran más en la representación visual y las razones cognitivas del uso de dibujos e imágenes. Entre ellas están los *mapas mentales* [2], muy usadas para organización del conocimiento de un modo jerárquico. Partiendo de una idea central, salen ramas en forma de estrella con ideas en las que se divide la idea central. Cada rama tiene subramas con ideas que se relacionan con la idea de esa rama. Es todavía un nivel muy abstracto para usarlo en el diseño de bases de datos, y no se definen claramente las relaciones entre las ideas que participan en mapa.

Otras herramientas se centran en los conceptos y en las relaciones entre ellos. Por ejemplo los *mapas conceptuales* [9], que se usan mucho en la conceptualización para la enseñanza de muy diferentes disciplinas, como se puede observar en el último congreso sobre ellos [3]. Los mapas conceptuales son los más adecuados para nuestros objetivos, porque en la herramienta necesitamos:

*Flexibilidad*, para poder hacer conceptualización a varios niveles de abstracción, desde el concepto más elemental a frases complejas que describan un tema muy abstracto. Los cmaps permiten esto porque sus diferentes componentes son pocos y con muy pocas restricciones comparados con otras herramientas, p. e. los diagramas de afinidad, que tienen unos pasos muy concretos a seguir con unos elementos complejos como son las tarjetas.

*Simplicidad*, para poder enseñarla y usarla rápidamente. Los cmaps tienen unas directrices fácilmente asimilables [1] y de uso inmediato, como hemos comprobado en otro de nuestros trabajos en un dominio distinto [5].

### 2.1. Elementos básicos de los Cmaps

Los cmaps representan conocimiento declarativo. Contienen proposiciones significativas llamadas *tripletras*, que se representan gráficamente con un concepto origen, una frase de enlace y un concepto destino. En la Figura 1 hay un ejemplo.

Los conceptos se escriben dentro de un rectángulo y pueden ser nombres comunes o frases descriptivas de un objeto. Las frases de enlace se escriben sobre una línea terminada en flecha hacia



Figura 1: Ejemplo de un cmap

el concepto destino y son verbos o frases que describen una acción hecha por el concepto origen que afecta al concepto destino. Los cmap pueden ser jerarquías de conceptos o redes. Como implementación de los cmap destaca un programa muy completo llamado CmapTools [3], desarrollado por el IHMC (Institute for Human & Machine Cognition). Incluye la gestión de servidores de mapas en internet de libre acceso. Los cmap de este artículo están hechos con ese programa.

### 3. Extendiendo el proceso de diseño de bases de datos

Hemos encontrado mejoras en el aprendizaje al apoyar el proceso de diseño con un cmap, cuando las especificaciones no están claras. Creamos el cmap interactuando con los expertos del dominio para explicitar los conceptos y sus relaciones de una forma visual. Definimos cuatro fases, que se muestran, junto a los resultados que generan, en la Figura 2:

- Fase 1: Construir el cmap.
- Fase 2: Restringir los elementos de cmap para aproximarlos a los elementos del diagrama EER.
- Fase 3: Crear el diagrama EER usando el cmap anterior.
- Fase 4: Transformar el diagrama EER en el esquema del Modelo Relacional.

No describimos en este artículo la Fase 4 porque no hay ninguna innovación, puesto que usamos el proceso habitual, ver [4] para detalles.

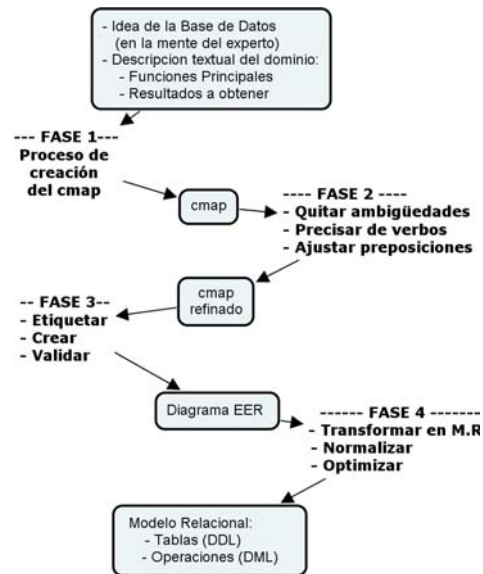


Figura 2. Proceso extendido del Diseño de Bases de Datos

#### 3.1. Fase 1: Construir el CMAP

No hay una metodología clara definida para construir cmap, solo hay algunas directrices. Hemos diseñado un proceso para construir los cmap de una forma sistemática basado en las directrices aplicables de las herramientas mencionadas. Lo hemos pulido basándonos en los comentarios que han hecho nuestros alumnos después de usarlo. Es un ciclo de refinamiento con los siguientes pasos:

1. Construir una lista de conceptos conocidos en el dominio a modelar. Un concepto se refiere a cualquier elemento que sea nombrable. No se crean frases de enlace en este paso. El dominio se ve sólo desde el punto de vista del propósito de la base de datos.
2. Marcar, como puntos de comienzo para el siguiente paso, los conceptos más generales de la lista.

3. Hacer grupos de los conceptos seleccionados en el paso 2. Para ello se usa, como criterio, algún aspecto común a varios conceptos, que sea relevante para el propósito de la Base de Datos. Esos conceptos formarán un grupo. Se crea un concepto nuevo, que describa dicho criterio, y se conecta a cada uno de los conceptos del grupo mediante líneas con flecha que apunten a dichos conceptos. Por ejemplo, un experto está describiendo un tratamiento, separado del cliente, de los conceptos sueldo y título sin mencionar ningún concepto que los agrupe. En esta situación, hay que crear un nuevo concepto, *puesto*, que agrupe sueldo y título, como se muestra en la Figura 1 de la sección anterior.
4. Crear frases de enlace para conectar el resto de los conceptos seleccionados en el paso 2. Cada frase debe contestar a la pregunta: ¿qué criterio tienen en común esos conceptos a conectar?. La respuesta es la frase de enlace que estará en la línea que los conecta. Conforme se agrupan los conceptos, se van quitando de la lista inicial.
5. Aplicar el paso 4 a los conceptos que quedan en la lista inicial. Se hace uno a uno para colocarlo en el cmap, allí donde haya una frase de enlace con otro concepto que está ya en el cmap. Si el concepto no tiene relación con ninguno de los que está en el cmap se marca y se continúa con los demás. Si después de enlazar todos quedan algunos marcados sin conectar, es muy probable que no se necesiten en el modelo de datos, así que se pueden eliminar en el paso 6.
6. Verificar y validar el cmap obtenido con expertos en el dominio o fuentes documentales. En esta fase pueden aparecer elementos nuevos o bien eliminar elementos antiguos volviendo al punto 1 si es necesario.

En la Figura 1, está un ejemplo del cmap obtenido. Normalmente, en esta fase, el cmap tiene varios problemas que serán tratados en la fase siguiente.

### 3.2. Fase 2: Refinando el cmap

El objetivo de esta fase es acercar el cmap, obtenido en la fase anterior, a los elementos que participan en el diagrama EER mediante la transformación de los términos en lenguaje natural usados en el cmap. Esta transformación es muy dependiente del lenguaje natural aplicado y el modo de usarlo específico de cada diseñador de

bases de datos. Por ejemplo, en inglés, el sujeto de una frase es obligatorio, esto simplifica la composición de las frases. Mientras que en español no lo es, necesitando una atención especial para determinar cuál es el sujeto en cada frase a crear. Otro ejemplo de transformación es el uso de verbos genéricos, como *haber*, *tener*, *ser*, en vez de verbos específicos. Para hacer la transformación hemos definido una serie de directrices. Nos basamos en el análisis del lenguaje que usan nuestros alumnos cuando hablan sobre el dominio que han escogido ellos para hacer una base de datos. A continuación mostramos las directrices y después un ejemplo de su aplicación:

1. *Quitar ambigüedades*: En todo el cmap los mismos conceptos deben tener el mismo significado y diferentes conceptos, diferentes significados.
2. *Dar precisión a las frases de enlace y a los verbos*: Se quitan verbos genéricos del tipo de tener, ser, haber, etc. Se transforman en verbos precisos de acción, o, si se puede, en lo que hemos llamado *verbos especiales*, que corresponden a elementos del diagrama EER. En la lista siguiente, para cada verbo especial se describe cuál debe ser el segundo concepto de la frase que se forma con ese verbo:
  - *Tiene-parte*: para un componente físico.
  - *Tiene-subclase* para subtipo o subclase.
  - *Tiene-valor* para un posible valor.
  - *Tiene-propiedad* para una característica.
3. *Transformar preposiciones* en verbos de acción. Esto se hace cuando aparece una preposición sola como frase de enlace, en vez de un verbo.
4. *Transformar dos frases de enlace* en una, cuando las dos frases forman parte de una sola acción. Por ejemplo, en la situación del paso 3, la preposición puede ser, en realidad, parte de la frase de enlace que la precede. En tal caso también se unen los conceptos que participan para formar uno más amplio.
5. *Validar el mapa refinado*, que obtenemos después de estos pasos, con los expertos en el dominio.

Estas directrices las hemos aplicado al cmap de la Figura 1. El resultado es el mapa refinado de la Figura 3. Aunque es un ejemplo simple se incluyen las transformaciones más relevantes. En

la Figura 1, el verbo genérico ‘tiene’ se refiere a tres acciones diferentes: *tiene-propiedad* en la frase que se forma ‘cliente tiene nombre’, *invierte* en ‘cliente tiene inversiones’ y *trabaja-en-un* en ‘cliente tiene un puesto’. En la misma figura, la preposición que está sola y que forma la frase ‘cliente tiene inversiones *con* compañía’ se refiere a la frase de enlace ‘tiene’ que esta antes de ‘inversiones’. Podemos transformarla, siguiendo la última directriz, en ‘cliente *invierte-en* compañía’.

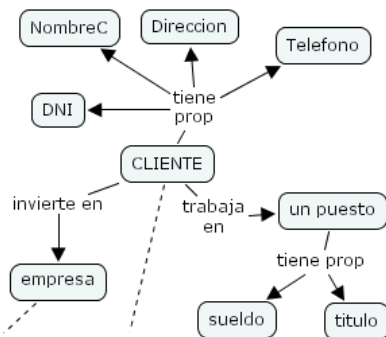


Figura 3: Ejemplo de Cmap refinado

3.3. Fase 3: Obtener el diagrama EER

En esta fase se usan, como fuentes de información, el mapa refinado, los expertos del dominio y las descripciones escritas que existan sobre el dominio. A continuación se describen las tres tareas de esta fase y después un ejemplo de su aplicación:

1. *Clasificar elementos* del cmap refinado. Para ello se usa la correspondencia entre los elementos del cmap y del diagrama EER de la Tabla 1. La correspondencia se hace poniendo etiquetas a los diversos elementos del cmap, con nombres de los elementos del diagrama EER, incluyendo la cardinalidad y la participación o cobertura.
2. *Construir el diagrama EER preliminar* haciendo las correspondencias indicadas por las etiquetas creadas en la tarea anterior.
3. *Verificar el diagrama EER preliminar* comprobando si faltan elementos, si hay elementos incorrectos o si hay elementos inútiles que se pueden eliminar. En esta tarea también se completan los elementos del EER

que no existen en los mapas conceptuales, tales como las claves primarias y las distinciones entre entidades fuertes y débiles.

Elementos EER	Usa el cmap refinado
Relación.....	Frase de enlace
depende del verbo especial:	Conceptos precedidos de:
Atributo .....	..Tiene-parte
Relación <i>es-un</i> .....	..Tiene-subclase
Valor (no hay en EER)..	..Tiene-valor
Atributo o Entidad .....	..Tiene-propiedad
(según casos particulares)	

Tabla 1: Correspondencia elementos Cmap y de EER

En la Figura 4 está el diagrama EER que se obtiene usando el cmap de la Figura 3 después de la presente fase. Esta fase puede llegar a ser complicada, porque la transformación de algunos elementos depende de los otros elementos que les rodean. En estos casos solo queda aplicar la teoría del modelo EER directamente. En la Figura 5 hay un ejemplo de estos casos donde tenemos dos conceptos apuntados por una frase de enlace *has-part*. Aunque parece la misma situación, cada uno es transformado de forma diferente: se crea un atributo *tipo-motor* del coche porque no tenemos mas información sobre ‘motor’. Sin embargo creamos una entidad *ruedas* porque sabemos, sobre ellas, la *marca* y el *tamaño*. Y creamos una sola relación *tiene-parte* para conectar *ruedas* y coche.

En la Tarea 1 se dan también casos particulares. Por ejemplo, si del concepto apuntado por un verbo *tiene-parte* cuelga, a su vez, otro verbo *tiene-parte*, entonces se crea una entidad con dicho concepto, en vez de un atributo como sería de esperar según las directrices indicadas.

4. Caso de estudio

Para hacer la evaluación de esta propuesta, primero hemos desarrollado una estrategia para enseñarla a nuestros alumnos de 3º de Ingeniería Informática. Posteriormente hemos aplicado cuatro técnicas informales para evaluar el impacto en los alumnos en comparación con el método habitual que veníamos usando hasta ahora.

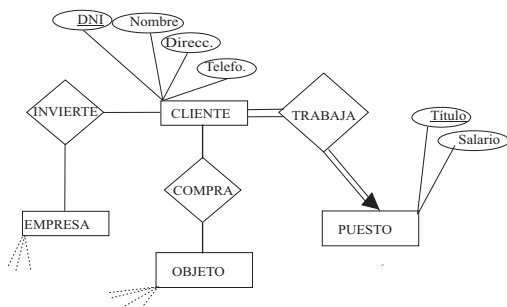
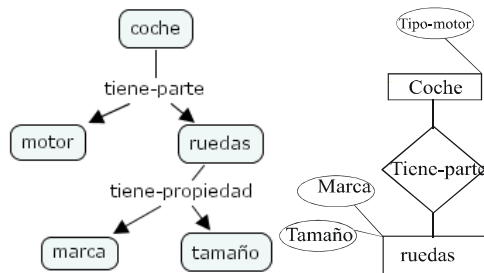


Figura 4: Diagrama EER

Figura 5: Transformación de *tiene-parte*

#### 4.1. Estrategia de enseñanza

La estrategia está basada en el paradigma de *aprender haciendo* y es guiada por ejercicios que tienen que hacer los alumnos. Empleamos seis horas para enseñarles los fundamentos de cmaps. Los alumnos trabajan en grupos de dos. En primer lugar, tienen que crear pequeños cmaps, asistidos por el profesor. Para enseñarles el enfoque propuesto de diseño de bases de datos empleamos tres horas en cada fase. Los alumnos aplican cada fase a los cmaps que han creado previamente, también asistidos por el profesor. En la siguiente etapa se les pide que hagan todo el proceso de diseño sobre su dominio favorito, creando una base de datos de tamaño medio. En esta etapa está incluida su implementación, empleando doce horas para ello. En este tiempo no se incluye las dieciocho horas que usamos para enseñar la teoría sobre el modelo EER, el modelo relacional, la fase 4 y el lenguaje SQL.

Después de cada paso, se les pide a los alumnos que escriban una *lista de problemas del alumno* con aquellas dificultades que han encontrado en ese paso. Al escribir la lista asimilan cuáles han sido sus dificultades principales. Sin estas listas tienden a olvidar muchas de las soluciones que han encontrado. Con las listas evitan, en sucesivas ocasiones, volver a explorar alternativas que han descrito ya en su lista de problemas, ahorrando tiempo en la clásica estrategia de prueba y error que usan los alumnos en general. Estas listas también sirven para la evaluación del proceso, así como para que el profesor resuelva los temas más relevantes que no han sido solucionados.

#### 4.2. Evaluación y resultados

Para comparar resultados tomamos, como grupo de referencia, los alumnos del año anterior que no habían visto cmaps. Así como las encuestas sobre sus dificultades en el aprendizaje del diseño de bases de datos. Establecimos un proceso de evaluación basado en cuatro técnicas diseñadas *ad hoc* para nuestro caso de estudio. Cada técnica mide uno de nuestros objetivos de evaluación, en comparación con las del grupo de referencia:

1. Para medir la comprensión global del proceso de diseño: comparamos las dificultades actuales de las *listas de problemas del alumno*.
2. Para medir la comprensión de los conceptos y del propósito de la base de datos: comparamos las preguntas sobre el dominio que hacen los alumnos a los profesores durante el proceso de desarrollo.
3. Para medir la desviación final del modelo pedido: comparamos la importancia de los cambios hechos del diseño que creían definitivo en relación con lo que se implementó finalmente. Estos cambios son motivados por los comentarios hechos por expertos en el dominio al ver los modelos preliminares implementados.
4. Para medir la satisfacción de nuestros alumnos, con todo el proceso: comparamos la encuesta sobre la facilidad de comprensión, de desarrollo y de corrección de errores. La valoración de los alumnos no puede considerarse como un criterio definitivo, ya que los estudiantes no son capaces de

evaluar, con perspectiva, sus propias mejoras en un dominio que están aprendiendo y solo conocen una parte de él.

El resultado cualitativo de este proceso indica una mejora apreciable en la comprensión de la teoría de base de datos y en los elementos del dominio sobre el que se crea la base de datos. Los alumnos muestran facilidad en el aprendizaje de la nueva herramienta, cmaps. Además de las encuestas y listas de problemas, como evaluación cualitativa, también hemos comparado las notas obtenidas para una evaluación cuantitativa. Ha habido una mejora general, que destaca en el porcentaje de las notas más bajas que han mejorado alrededor de un 20%, mientras que las medias y altas se mantienen similares.

## 5. Conclusiones

Después de enseñar varios años diseño de bases de datos, descubrimos un problema serio: los alumnos no entienden fácilmente el proceso de conceptualización necesario para obtener el diseño de una base de datos. Y este problema sucede antes de donde empieza la enseñanza actual del diseño. Es decir, en el modo de pensar sobre el dominio y comprender sus elementos.

Nuestra propuesta es modificar el proceso de diseño de base de datos para incluir fases previas al proceso habitual que ayuden en el proceso de formación de conceptos, usando la herramienta de mapas conceptuales o cmaps. Estos representan conocimiento declarativo a diferentes niveles de abstracción, usando conceptos y frases de enlace entre ellos para componer proposiciones sobre el dominio.

En la propuesta definimos una técnica de construcción de cmaps específica como fase previa para el diseño de bases de datos. Definimos el proceso de diseño de la base de datos que incluye los cmaps, con tareas de aproximación sucesiva al diagrama EER. Definimos una estrategia de enseñanza de este proceso, que está basada en el paradigma de aprendizaje haciendo. Aplicamos la propuesta a un caso de estudio: nuestros alumnos de ingeniería de informática. Y desarrollamos algunas técnicas *ad hoc* para la evaluación del impacto en los alumnos.

## 6. Trabajo futuro

Como trabajo futuro, en el curso próximo haremos una evaluación más formal con encuesta detallada para extraer estadísticas. Una vez validada la aproximación propuesta se planea hacer una herramienta que asista al alumno en el proceso de diseño. También planeamos introducir la propuesta del proceso de diseño en una empresa de desarrollo de software para evaluar los resultados en desarrollo profesional.

## 7. Agradecimientos

Financiado por la Comisión Interministerial de Ciencia y Tecnología (TIC2002-01961)

## Referencias

- [1] Ahlberg, M. *Varieties of concept mapping in Concept Maps: Theory, Methodology, Technology*, Proc. of the First Int. Conference on Concept Mapping A. J. Cañas, J. D. Novak, F. M. González (Eds.) Pamplona, Spain. 2004.
- [2] Buzan, T., Buzan, B. *The mind map book* (Millenium ed.). London: BBC Books. 2000.
- [3] Cañas, A. J., Hill, G., Carff, R., Suri, N., Lott, J., Gómez, G., Eskridge, T. C., Arroyo, M., Carvajal, R. *CmapTools: a knowledge modeling and sharing environment* In Concept Maps: Theory, Methodology, Technology Proc. of the First Int. Conference on Concept Mapping A. J. Cañas, J. D. Novak, F. M. González (Eds.) Pamplona, Spain. 2004.
- [4] Elmasri, R., Navathe, S.R. *Fundamentals of Database Systems*. 3ª Ed. 1997. Addison-Wesley. 2000.
- [5] Gómez-Gauchía, H., Díaz-Agudo, B., González-Calero, P.A., *A Pragmatic Methodology for Conceptualization with two layered Knowledge Representation: a case study*. In Conceptual Structures at Work, Contributions to 12th International Conference on Conceptual Structures. ICCS2004. Pfeiffer, H.D., Wolff, K.E., Delugach, H.S. (Eds.) Shaker Verlag. 2004.
- [6] Gowin, B. *Educating*, Ithaca, NY: Cornell University Press. 1981/1987.

- [7] Kelly, G., S. *The Psychology of Personal Constructs*. Ed. Norton, New York, 1955.
- [8] Kosko, B. *Fuzzy cognitive maps* International Journal of Man-Machine studies 1986,24, 65-75
- [9] Novak, J. D., Gowin, D. B.. *Learning how to learn*. New York: Cambridge University Press. 1984.
- [10] Ohiwa, H. *KJ Editor for Creative Work Support and Collaboration* in Proceedings of the First Conference on Creating, Connecting and Collaborating through Computing (C5'03) 2003.