

# Un sistema de razonamiento basado en casos para apoyar la toma de decisiones en la industria de la construcción

Cruz Ac J. M.<sup>1</sup>, González Fajardo J. A.<sup>2</sup>, Zaragoza Grifé J. N.<sup>3</sup>

*Fecha de recepción: 3 julio 2013 – Fecha de aprobación: 30 agosto 2013*

## RESUMEN

En la industria de la construcción es común que no se lleve un registro documental de todos los problemas que resuelven. Específicamente durante la fase de ejecución, el gerente del proyecto es el encargado de resolver todos los problemas que se presentan con la ayuda de su equipo de trabajo. Desafortunadamente el aprendizaje obtenido por el gerente del proyecto y su equipo se queda únicamente en su experiencia personal y no en la empresa ni en la industria lo que resulta costoso para muchas organizaciones.

El Razonamiento Basado en Casos (RBC) es una herramienta usada en Inteligencia Artificial que permite resolver problemas nuevos a través de la búsqueda de problemas viejos documentados dentro de una base de datos o base de casos.

En este trabajo de investigación se presentan las bases conceptuales para realizar el desarrollo de un sistema de RBC aplicado a la gerencia de proyectos de construcción masiva de vivienda (CMV). En los resultados obtenidos se muestra la conceptualización de cada uno de los componentes del sistema RBC propuesto (base de casos, recuperación, adaptación y retención), mostrando claramente cuáles son los pasos a seguir para lograr la automatización de cada una de las fases del proceso RBC.

**Palabras clave:** CBR, vivienda, gerente, construcción, proyectos.

## ABSTRACT

In the construction industry it is not common to keep a record of the problems solved. Specifically during execution phase, the construction project manager is the person responsible for solving all problems that arise during the work, with the help of his team. Unfortunately all lessons learned by the project manager and his team remains only as individual or team experience at most. Construction industry and firms do not gain the experience, but the people who work for them, this results costly for organizations.

Case based Reasoning (CBR) is a tool used in Artificial Intelligence which helps to solve new problems by searching into old problems in a case base. This paper presents the foundations to develop a CBR System applied to the construction project management in mass construction of housing. Results of this work show the conceptualization of every part of the CBR System proposed (case base, retrieve, adaptation and retain). It also clearly shows the steps necessary for automation of every phase of the system.

**Keywords:** CBR, housing, manager, construction, projects.

---

<sup>1</sup>Estudiante de Posgrado de la Facultad de Ingeniería de la UADY.

<sup>2</sup>Universidad Autónoma de Yucatán, Cuerpo Académico Ingeniería de la Construcción, correo electrónico: jagonz@uady.mx

<sup>3</sup>Universidad Autónoma de Yucatán, Cuerpo Académico Ingeniería de la Construcción, correo electrónico: zgrife@uady.mx

**Nota:** El período de discusión está abierto hasta el 1° de marzo de 2014. Este artículo de investigación es parte de **Ingeniería–Revista Académica de la Facultad de Ingeniería, Universidad Autónoma de Yucatán**, Vol. 17, No. 2, 2013, ISSN 1665-529-X.

## **INTRODUCCIÓN**

Es común que durante la ejecución de un proyecto de construcción surjan problemas que afecten los planes formulados al inicio, algunos ejemplos son: demoras y retrasos en los tiempos programados de construcción, retrasos en los suministros de materiales, problemas con proveedores, falta de liquidez, falta de productividad de la mano de obra, entre otros. Ante esta problemática el gerente del proyecto, en su rol de líder, una de sus principales responsabilidades es la de tomar decisiones pertinentes para corregir los problemas para mantener el proyecto dentro de lo planeado.

Para decidir con respecto a algún caso en particular, el gerente del proyecto primero recaba toda la información disponible para estructurar un panorama completo del problema; esto lo realiza apoyándose en su equipo de trabajo (residente de obra, superintendente de obra, jefe de compras, contador, etc.) quienes le proporcionan la información necesaria para poder realizar un diagnóstico del problema y posteriormente proponer una solución. Es aquí donde la experiencia del gerente del proyecto juega un papel importante, ya que apoyándose en los problemas que resolvió en el pasado podrá formular nuevas soluciones a los problemas de futuros proyectos.

El problema empieza a generarse cuando la persona que debe tomar las decisiones no tiene la experiencia suficiente o no conoce a profundidad sobre el caso en cuestión, lo que origina que tome decisiones poco acertadas. Otro problema común que sucede en la industria de la construcción es que cuando se obtiene un aprendizaje nuevo resolviendo un problema mediante la experiencia, dicho aprendizaje se queda únicamente en la mente de la persona o del grupo que lo resolvió, lo que origina que ante la ausencia de dicha persona o grupo con esa experiencia, se tenga que replantear el problema como si fuera nuevo, es decir, existe un retroceso en el aprendizaje de los problemas lo cual resulta costoso para la organización.

Un sistema de Razonamiento Basado en Casos (RBC) es una herramienta perteneciente al campo de la inteligencia artificial que emula el proceso de razonamiento que realiza un ser humano para

resolver un problema. El RBC utiliza una base de casos de problemas resueltos en el pasado. Ante un nuevo problema, recupera los casos más similares almacenados en la base de casos, adapta los casos recuperados, propone una nueva solución y finalmente almacena el nuevo caso resuelto con el fin de enriquecer la base de casos.

El RBC permite solucionar un problema nuevo sin partir de cero, es decir, se aprovecha las experiencias previas. A quienes tienen experiencia, el RBC les permite proponer soluciones más satisfactorias y los que carecen les permite proponer soluciones conociendo de antemano que es lo que ya se ha hecho previamente con respecto al problema en cuestión (Watson 1997).

Muchas de las decisiones que se necesitan tomar durante el proceso de ejecución de un proyecto de construcción están bajo presión del tiempo (Richard et al. 2000); el RBC permite disminuir el tiempo de respuesta a un problema (Luger 2005). En el campo de la administración de proyectos de construcción, en trabajos recientes se ha aplicado a problemas tales como identificar peligros que pueden ocurrir durante la ejecución de proyectos de construcción (Goh y Chua 2009) o a estimar la duración del tiempo de construcción de un edificio (Han et al., 2007)

En este trabajo es presentado el desarrollo de las bases conceptuales del funcionamiento de un sistema que utiliza el RBC aplicado al contexto de la gerencia de proyectos en la construcción masiva de vivienda (CMV), para apoyar el proceso de toma de decisiones administrativas. El sistema fue conceptualizado en primera instancia conociendo cuales son los casos o problemas más comunes que se hayan presentado durante el proceso de construcción de vivienda masiva y a partir de ellos se generó una base de casos como ejemplo, un mecanismo de recuperación y un procedimiento de adaptación para poder proponer soluciones a problemas futuros.

## **METODOLOGÍA**

La metodología utilizada se dividió en dos fases: primero se realizó un trabajo de investigación de campo con el objetivo de recabar el mayor número de casos acontecidos en el contexto real

de la CMV. Por último se modeló conceptualmente el funcionamiento del sistema a partir de los casos recabados.

### **Primera fase: Investigación en Campo**

En general, existen tres tipos de representación de casos que pueden adoptarse en el desarrollo de un sistema RBC (Bergmann et. al. 2005): vectores de valores-atributos, casos relacionales y la representación textual. No fueron encontrados trabajos de investigación que proporcionen un registro documental de problemas que hayan sido resueltos por los gerentes de proyectos de CMV, de tal manera que no fue posible tener una idea de cuál enfoque de representación de datos se debiera utilizar de los tres mencionados.

La estrategia utilizada para obtener la información requerida para generar y representar la base de casos del sistema RBC, fue la de aplicar una entrevista exploratoria a los gerentes de proyectos de la región de estudio. Dicha estrategia incluyó la definición de la terminología que a continuación se describe:

**Diseño de la Investigación.** El tipo de investigación que se llevó a cabo en este trabajo consistió en realizar un desarrollo tecnológico porque se aplicó un procedimiento sistemático basado en el conocimiento previo, el cual fue obtenido mediante la investigación y/o la práctica (Busch 1999). Como producto final de esta investigación fueron obtenidas las bases conceptuales para el desarrollo de una herramienta computacional dirigida a la gerencia de proyectos de construcción, que permita apoyar el proceso de toma de decisiones.

**Unidad de Análisis.** En este trabajo se tomó como unidad de análisis al gerente de proyectos de una empresa de CMV. Para establecerla se necesitó definir las características del sujeto “tipo”, quienes para este caso, cumplieron con: *a)* se desempeñó como gerente de proyectos, gerente de construcción, superintendente de obras, gerente de suministros o dueño, *b)* sus funciones estuvieron dentro de una empresa de CMV en la región de estudio y *c)* al menos 5 años de experiencia en su puesto. El universo estuvo conformado por la totalidad de los gerentes de

proyectos de construcción masiva de vivienda que se desempeñan en la región de estudio.

**Muestra.** Para calcular el tamaño de la muestra se siguió un procedimiento no aleatorio porque no se buscaba que los resultados del estudio fueran representativos de toda la población y no probabilístico porque los elementos que integran la muestra fueron seleccionados con base en las características que debían cumplir los sujetos “tipo”.

El objetivo de la recolección de datos en campo fue documentar la experiencia de los gerentes de proyectos de construcción, de tal manera que fueron seleccionados primeramente los gerentes de las principales empresas que desarrollan vivienda a gran escala (de 2000 a 5000 unidades por año) en la localidad del estudio y en segunda instancia fueron consideradas el resto de las empresas que desarrollan vivienda en menor escala (de 500 a 2000 unidades por año).

Fueron seleccionadas un total de 10 empresas constructoras de toda la población y un gerente de proyecto por cada empresa. El número total de empresas que conforman la población fue tomado del padrón de contratistas afiliados a la Cámara Nacional de la Industria de Desarrollo y Promoción de Vivienda (CANADEVI, Delegación Yucatán).

**Instrumento para la Recolección de Datos en Campo.** Para conocer cuáles son los principales problemas que tuvieron que resolver los gerentes de proyectos de vivienda durante la fase de ejecución de algún proyecto, se diseñó una entrevista exploratoria la cual fue aplicada a todos los gerentes que conformaban la muestra. Los principales aspectos contenidos en la entrevista fueron: cuál fue el caso o problema acontecido; cuáles fueron los indicadores de la existencia de un problema, cuál fue el proceso para obtener su solución y finalmente, cuál fue la solución que se implementó y los resultados obtenidos.

### **Segunda Fase: Desarrollo Conceptual del Sistema RBC**

Para explicar el funcionamiento conceptual del sistema RBC propuesto se utilizó el esquema básico utilizado por varios autores (Kolodner

1993; Watson 1997). La Figura 1 muestra los componentes de un sistema RBC. Se describió de forma detallada la manera en que se conceptualizó cada uno de los componentes del sistema tomando en consideración todos los casos recabados en el trabajo de campo.

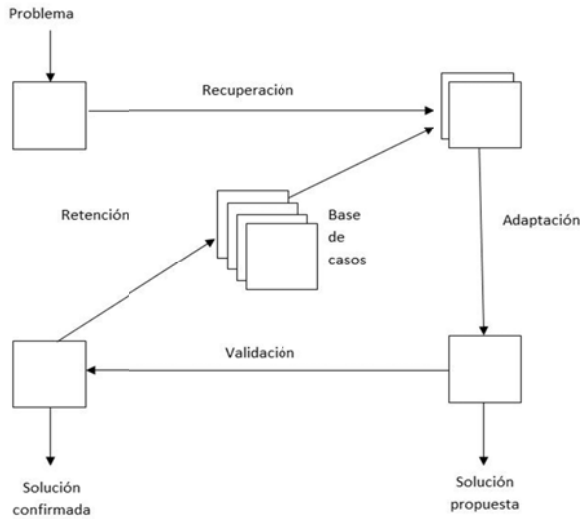


Figura 1. Ciclo de un sistema de Razonamiento Basado en Casos. Fuente: Applying Case Based Reasoning (Watson 1997).

**La Base de Casos.** Del análisis realizado a las grabaciones de las entrevistas fueron extraídos todos los casos que tuvieron que resolverse durante la fase de ejecución de algún proyecto de CMV. Los casos obtenidos se concentraron en una Tabla identificando los siguientes aspectos: 1) una descripción textual; 2) una descripción textual de las causas; y 3) una descripción textual de las soluciones implementadas para resolverlos.

Para su representación en la base, se decidió que fuese de forma textual; lo anterior debido a que los casos que fueron recabados del campo no proporcionaban mucha información que permitiera realizar una representación vectorial de valores-atributos del problema, el cual contiene atributos que están presentes en todos los casos de la base.

**Fase de Recuperación.** Consistió en dos pasos: 1) realizar una búsqueda textual de la descripción de un nuevo problema dentro de la base de casos obtenida en el trabajo de campo y 2) una vez

obtenidos todos los casos similares con el nuevo problema, se realizó un dimensionamiento de la similitud de dichos casos recuperados con el nuevo problema.

Para explicar el funcionamiento del proceso de recuperación fueron utilizados 2 enfoques: el Modelo del Espacio de Vectores (Salton y Buckley 1988) y la Red para la Recuperación de Casos Mejorada (Chakraborti et. al. 2006). La razón de haber utilizado dos enfoques diferentes fue la falta de experiencia en la aplicación de un sistema RBC en el contexto de la CMV, de tal manera que no se conocía cuál de los dos enfoques podría proporcionar los mejores resultados.

En el enfoque del Modelo del Espacio de Vectores (VSM) toda la base de casos, representada de forma textual, fue transformada en un espacio vectorial en donde cada caso en particular es representado mediante un vector. Los componentes numéricos de cada vector se definieron en base a la descripción del problema. Para poder transformar la descripción textual de un caso en un vector con componentes numéricos y para evaluar la similitud del nuevo caso con los existentes fue utilizada la metodología propuesta por Salton (Salton y Buckley 1988).

En el enfoque de la Red para la Recuperación de Casos Mejorada (FCRN) se construyó una red que relaciona las entidades de información (IEs) del nuevo caso con los casos existentes en la base. Una Entidad de Información (IE) es la representación de una pieza atómica de conocimiento, es la unidad más pequeña en la cual puede dividirse la información para poder utilizarla sin perder su estructura original. Para evaluar la importancia y similitud entre las IEs del nuevo caso con los casos existentes en la base, se utilizaron las fórmulas desarrolladas por Chakraborti (Chakraborti et. al. 2006).

**Fase de Adaptación.** Para explicar el funcionamiento de la fase de adaptación del sistema RBC se utilizó el enfoque de la adaptación composicional, que consiste en combinar soluciones de múltiples casos para generar una solución al problema en cuestión (Arshadi y Badie 2001).

Para que el sistema pudiera proponer una solución por sí solo (de manera “inteligente”) se hizo una combinación de todas las soluciones implementadas en los casos similares que fueron recuperados en el paso anterior.

**Fase de Validación.** Durante esta fase con respecto a la solución propuesta al usuario se podrían presentarse 3 diferentes situaciones:

- 1.- Es completamente satisfactoria. En caso de ocurrir esto se pasaría automáticamente a la siguiente fase.
- 2.- Es parcialmente satisfactoria. En este caso el usuario tiene la oportunidad de poder “editar” la solución propuesta. Durante el proceso de edición se pueden eliminar, agregar o modificar las soluciones propuestas.
- 3.- Es insatisfactoria. En este caso el usuario tendría la oportunidad de ingresar una nueva causa del problema y una nueva solución.

**Fase de Retención.** Independientemente de la situación que se haya tenido de la fase anterior, se propuso que el sistema almacenara automáticamente el caso resuelto y validado sin importar que la descripción textual del nuevo caso resuelto coincidiera con alguno de los que estuvieran en la base, es decir, el sistema no eliminaría la redundancia de los casos. Esta consideración se hizo tomando en cuenta que, en el trabajo de campo se observó que la descripción de un mismo problema puede tener diferentes causas y diferentes soluciones dependiendo de la empresa en donde se genere. Al nuevo caso resuelto se le asignó un nuevo identificador único (ID) y pasó a formar parte de la base de casos.

## **RESULTADOS**

**La Base de Casos.** Los resultados obtenidos del trabajo de campo sirvieron para poder generar una base de casos de problemas que son representativos del sector de la CMV en la región de estudio. Los casos fueron problemas reales que tuvieron que resolver los gerentes de proyectos que participaron en la investigación. Los casos obtenidos sirvieron como ejemplo para modelar conceptualmente todas las fases del sistema RBC propuesto.

Para que cada empresa constructora pueda generar su propia base de casos es preciso que se lleve a

cabo la documentación de todos los problemas que se resuelven durante la ejecución de algún proyecto de CMV. Todas las demás fases del sistema RBC propuesto en este trabajo son aplicables a cualquier base de casos que sea generada tanto a nivel sector como a nivel empresa, es decir, su aplicabilidad no se limita a ninguna de las dos antes mencionadas.

En el análisis de las entrevistas aplicadas durante el trabajo de campo, a nivel sector, se obtuvieron inicialmente 42 diferentes casos provenientes de los 10 diferentes gerentes de proyecto entrevistados. Durante una segunda revisión se desecharon los casos redundantes y los casos no representativos de algún problema, por lo que al final sólo quedaron 29 casos en la base. En este artículo no se muestran la totalidad de los casos recabados en el trabajo de campo, por razón de espacio. Posteriormente se elaboró un resumen de los casos más representativos con el objetivo de poder ejemplificar de qué manera sería el funcionamiento de cada una de las fases del sistema RBC propuesto.

El resumen de la base de casos mostrada en la Tabla 1, se elaboró tomando en cuenta que para ejemplificar la operación del sistema RBC propuesto en este trabajo se requiere de la realización de muchas operaciones matemáticas que deben ser programadas utilizando algún lenguaje de programación (principalmente durante la fase de recuperación), lo cual no fue el objetivo de este trabajo. En el ejemplo descrito, todas las operaciones matemáticas utilizadas para realizar la recuperación y evaluación de la similitud de los casos fueron realizadas utilizando únicamente una calculadora de bolsillo.

**Fase de recuperación.** Dadas las condiciones en las que fue estructurada y representada la base de casos (representación textual) obtenida del trabajo de campo, se decidió que la manera más práctica de poder realizar la recuperación de los casos similares en la base fue a través de la descripción textual del problema.

Supóngase que se tiene la descripción del problema nuevo: “*Los materiales requisitados no llegan satisfactoriamente a la obra*”.

Antes de aplicar las funciones de recuperación y evaluación de los casos fue necesario definir cuáles serían las Entidades de Información (IEs) de cada uno de los documentos de la base. Para el ejemplo, todas las IEs de los casos contenidos en la base resumida se definieron utilizando palabras “clave” de la descripción del problema (Tabla 1).

De igual forma el nuevo problema expresado de forma de *IEs* quedaría de la siguiente manera: “*materiales, requisitados, no llegan, satisfactoriamente, obra*”.

Una vez definidas las IEs de todos los casos de la base y del problema a resolver se procedió a aplicar los procesos de recuperación mencionados. Bajo el enfoque del Modelo del Espacio de Vectores el proceso de recuperación de los casos más similares fue conceptualizado de la siguiente manera:

**Paso 1.** Realizar una búsqueda simple dentro de la base de casos para seleccionar todos los documentos que tengan coincidencia con las entidades de información de la nueva consulta.

Para realizar la búsqueda de las coincidencias se puede utilizar el algoritmo Boyer-Moore (Cole, 1991) o una consulta *Select Query Language*, (SQL). Se consideró todos los casos de la Tabla 1, excepto el 4, tienen coincidencia.

**Paso 2.** Una vez que se ha realizado un primer filtro de todos los caso de la base que tengan cuando menos una IE coincidente con la descripción del nuevo caso a resolver, se procedió a crear una nueva Tabla en donde las filas representan a los documentos con alguna coincidencia y las columnas representan a las IEs de cada documento, como lo muestra la Tabla 2. Esta Tabla se denomina “pseudo-base” ya que en realidad es una base de casos temporal creada a partir de la base de casos original y solamente se genera con el objetivo de eliminar el número de operaciones matemáticas que tendría que realizarse para evaluar la similitud de los documentos de la base. En esta Tabla se agregó el nuevo problema para poder evaluar su similitud con los casos de la “pseudo-base”

Tabla 1. Base de casos resumida del trabajo de campo y Entidades de Información correspondientes.

Id caso	Descripción del Problema	Entidades de información	Causas	Solución implementada
1	Los materiales requisitados no llegan oportunamente a la obra	materiales, requisitados, no llegan, oportunamente	Falta de coordinación entre el departamento de compras y el departamento de construcción	Establecer un canal de comunicación directa entre el jefe de compras y el jefe de construcción para resolver la falta de coordinación en las requisiciones
2	Existen retrasos en el suministro de materiales en obra	retrasos, suministro, materiales, obra	El proveedor no suministra oportunamente la cantidad de materiales solicitados	Establecer un canal de comunicación directa entre el jefe de compras y el representante del proveedor para coordinar el suministro de los materiales
3	Las requisiciones de materiales no se cumplen satisfactoriamente	requisiciones, materiales, no se cumplen	La empresa no tiene dinero para solventar los pagos a proveedores	Llegar a una promesa de pago con los diferentes proveedores a través de un cheque post-fechaado
4	Hay escasez de mano de obra	escasez, mano de obra	Existe mucha competencia en el mercado	Modificar los procedimientos constructivos de las viviendas para agilizar el tiempo de ejecución
5	Los materiales especiales solicitados no llegan a la obra	materiales especiales, no llegan, obra	No se piden con suficiente tiempo de anticipación	Todos los materiales especiales se piden con 3 meses de anticipación

Tabla 2. “Pseudo-base” creada con las coincidencias y con la nueva consulta.

Documento	IE1	IE2	IE3	IE4	IE5
D1	materiales	requisitados	no llegan	oportunamente	
D2	retrasos	suministro	materiales	obra	
D3	requisiciones	materiales	no se cumplen		
D5	materiales especiales	no llegan	obra		
D6	materiales	requisitados	no llegan	satisfactoriamente	obra

**Paso 3.** Convertir los casos recuperados en vectores.

Para realizar la conversión de los documentos D1, D2, D3, Dn de la “pseudo-base” en vectores numéricos se utilizaron las fórmulas propuestas por Salton (Salton y Buckley 1988). En el Modelo del Espacio de Vectores los documentos textuales son representados vectorialmente de la siguiente manera:

$$V_d = [W_{1,d}, W_{2,d}, W_{3,d}, \dots, W_{n,d}]^T$$

En donde:

W= es el peso o importancia de cada una de las Entidades de Información (IEs) tanto en el documento D respectivo (parámetro local) como en toda la base de casos D1, D2, D3, Dn (parámetro global)

El superíndice “T” no expresa ninguna potencia, únicamente denota que cada uno de los valores “W” esta referido a cada uno de los términos o Entidades de Información de cada uno de los documentos de la base de casos.

Para calcular los valores de “W” se procedió de la siguiente manera:

$$W_{t,d} = tf_{t,d} \times \log \frac{|C|}{|\{d' \in C | t \in d'\}|}$$

En donde:

$tf_{t,d}$  = es la frecuencia de la Entidad de Información  $IE_n$  en el documento  $D_n$  (parámetro local)

$|C|$  = es el número total de documentos en la base de casos

$|\{d' \in C | t \in d'\}|$  = es el número total de documentos de la base que contienen a la Entidad de Información  $IE_n$  (parámetro global).

Con el fin de comprender la fórmula expuesta con anterioridad se expone el siguiente ejemplo. Para calcular el valor “W” de la Entidad de Información  $IE_1$  en el documento  $D_1$  se procedió de la siguiente manera:

Ya que  $tf_{t,d}$  es la frecuencia de la Entidad de Información  $IE_1$  en el documento  $D_1$ . En la “pseudo-base” contenida en la Tabla 2 se puede observar que la Entidad de Información  $IE_1$  que corresponde al documento  $D_1$  es la palabra

“materiales”. El número de veces que aparece la palabra “materiales” en el documento  $D_1$  es igual a la unidad, por lo tanto la frecuencia o valor  $tf_{t,d} = 1$ . Por convención del autor del cual fue tomada la fórmula (Salton y Buckley 1988), el valor de la frecuencia se expresó como una fracción del total de documentos de la base quedando el valor  $tf_{t,d} = 1/6$ .

El valor  $|C|$  es igual a 6 ya que éste es el número total de documentos contenidos en la “pseudo-base”. Dicho valor permanece constante durante todo el proceso de recuperación.

El valor  $|\{d' \in C | t \in d'\}|$  corresponde al número total de documentos de la “pseudo-base” que contienen a la Entidad de Información  $IE_1$  del documento  $D_1$ , es decir, a la palabra “materiales” que para este ejemplo dicho valor es 4.

Una vez obtenidos todos los valores necesarios de la formula, se procedió a calcular el valor “W” de la Entidad de Información  $IE_1$  del documento  $D_1$  quedando de la siguiente manera:

$$W_{IE_1,D_1} = \frac{1}{6} \times \log \frac{6}{4} = 0.016$$

Para calcular los valores de “W” en todos los demás documentos se procedió de manera similar Tabla 3, 4 y 5.

Tabla 3. Valores de  $tf_{t,d}$  de las Entidades de Información de la “pseudo-base”.

Documento	tf1	tf2	tf3	tf4	tf5
D1	1/6	1/6	1/6	1/6	0
D2	1/6	1/6	1/6	1/6	0
D3	1/6	1/6	1/6	0	0
D5	1/6	1/6	1/6	0	0
D6	1/6	1/6	1/6	1/6	1/6

Tabla 4. Valores de  $\log \frac{|C|}{|\{d' \in C | t \in d'\}|}$  de las Entidades de Información de la “pseudo-base”.

Documento	IE1	IE2	IE3	IE4	IE5
D1	0.097	0.398	0.222	0.699	0
D2	0.699	0.699	0.097	0.222	0
D3	0.699	0.097	0.699	0	0
D5	0.699	0.222	0.222	0	0
D6	0.097	0.398	0.222	0.699	0.222

Tabla 5. Valores de “W” de todas las Entidades de Información de la “pseudo-base”: representación vectorial de los documentos recuperados:

Vector	W1	W2	W3	W4	W5
V1	0.016	0.066	0.037	0.116	0.000
V2	0.116	0.116	0.016	0.037	0.000
V3	0.116	0.016	0.116	0.000	0.000
V5	0.116	0.037	0.037	0.000	0.000
V6	0.016	0.066	0.037	0.116	0.037

En la Tabla anterior se puede observar que cada documento textual de la “pseudo-base” ya fue transformado en valores numéricos, los cuales permiten realizar una evaluación numérica de la similitud que guardan los documentos entre sí.

**Paso 3.** Evaluar la similitud de la consulta con todos los casos de la base. Para realizar este paso se calculó el coseno del ángulo  $\alpha$  entre el vector de la consulta ( $V_6$ ) y todos los demás vectores de la sub-base. Para realizar el cálculo se utilizó la siguiente formula:

$$\begin{aligned} \text{sim}(D_j, D_k) &= \frac{(D_j)(D_k)}{|D_j||D_k|} \\ &= \frac{\sum_{i=1}^n W_{i,j}W_{i,k}}{\sqrt{\sum_{i=1}^n W_{i,j}^2} \sqrt{\sum_{i=1}^n W_{i,k}^2}} \end{aligned}$$

Aplicando esta función los resultados fueron los mostrados en la Tabla 6.

Tabla 6. Evaluación de la similitud de la consulta con todos los casos de la base.

sim (V6,V1)	0.967
sim (V6,V2)	0.591
sim(V6,V3)	0.303
sim(V6,V5)	0.308

Como se puede ver de la Tabla anterior, la mayor similitud está entre los documentos (V6, V1) y (V6, V2). El rango de valores mínimos de similitud aceptados puede ser definido por el usuario del sistema.

Para describir el proceso de recuperación utilizado bajo el enfoque de la Red para la Recuperación de

Casos Mejorada (FCRN) se utilizó la misma base de datos resumida y la misma consulta utilizada con anterioridad en el Modelo del Espacio de Vectores.

**Paso 1.** Realizar una búsqueda simple dentro de la base de casos para seleccionar todos los documentos que tengan coincidencia con las entidades de información de la nueva consulta. Para realizar la búsqueda de las coincidencias se puede utilizar el algoritmo Boyer-Moore (Cole 1991) o una consulta SQL (Tabla 2).

**Paso 2.** Una vez obtenidos todos los casos con coincidencias se construyó una red que relaciona las IEs de la consulta con los casos en donde existen coincidencias.

Consulta: “materiales, requisitados, no llegan, satisfactoriamente, obra”.

En la red de la Figura 2 se puede observar que todas las IEs que tuvieron coincidencia tienen un valor de “1” en el caso respectivo. La IE “satisfactoriamente” no tuvo ninguna coincidencia por lo que su valor es cero.

**Paso 3.** Aplicar la función de similitud entre las IEs de la consulta y las IEs de la base de casos. Para realizar este paso es necesario que desde la construcción de la base y la definición de las IEs, se establezca un vocabulario alternativo para cada una de las IEs de la base (Figura 3).

Para el ejemplo se propuso que la única IE de la base que tenía un vocabulario alternativo fue la palabra “oportunamente”; dicha IE pertenece a la descripción del problema en el caso 1. Se estableció que la palabra “oportunamente” se parece en 0.80 con la palabra “satisfactoriamente”. Aplicando esta función de similitud la nueva red se muestra en la Tabla 7.



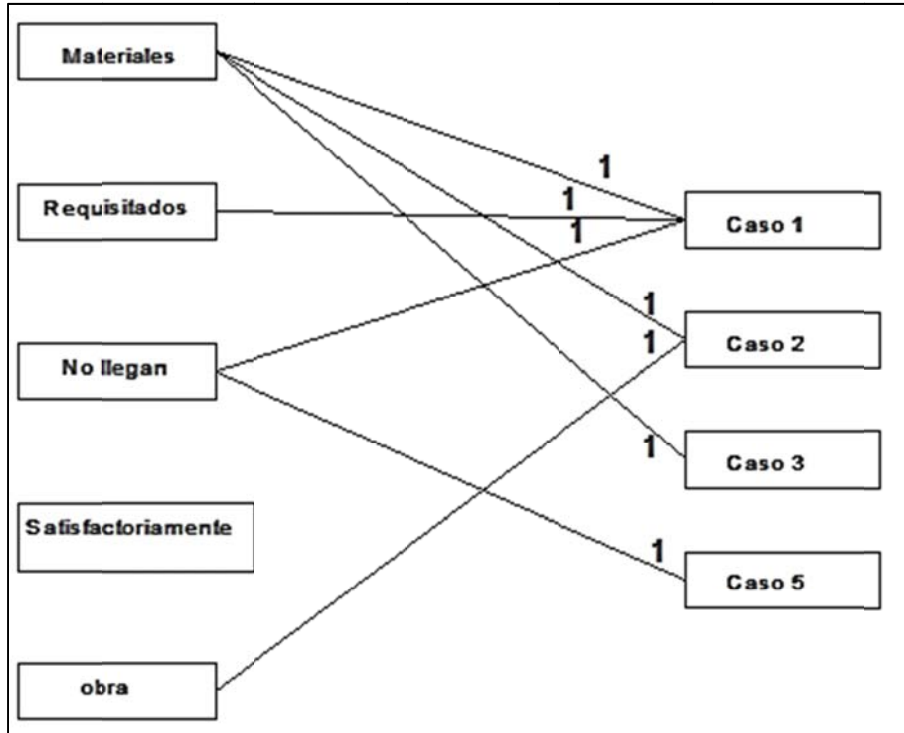


Figura 2. Red para la Recuperación de Casos (CRN) del problema

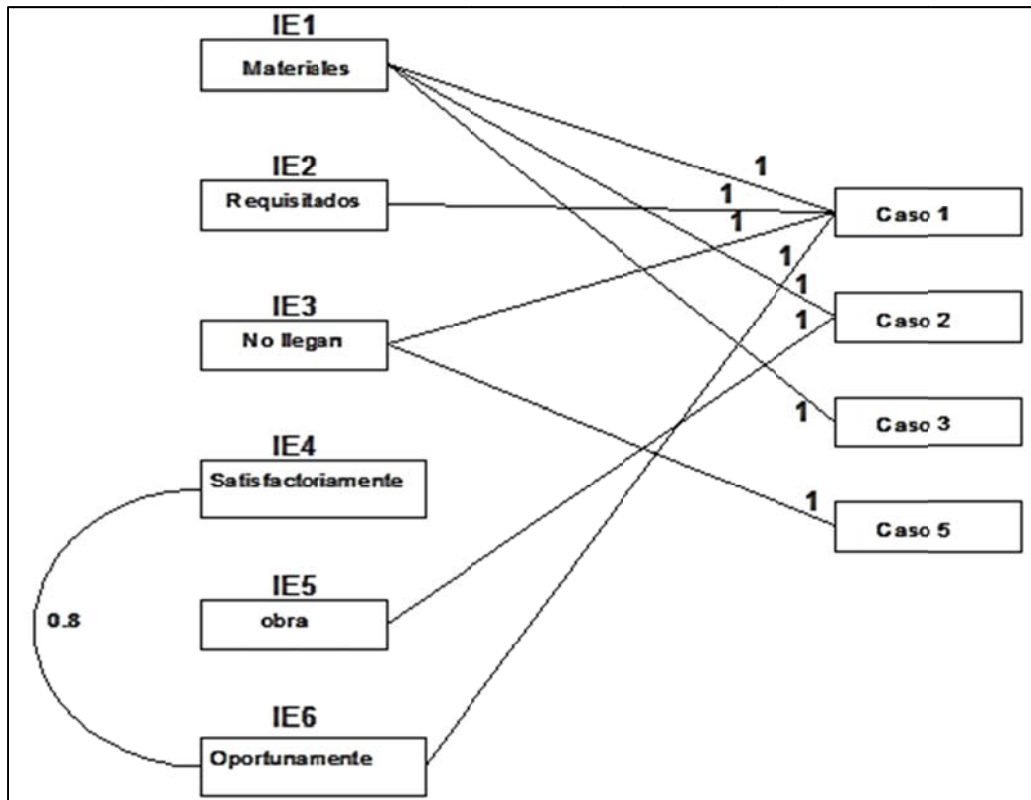


Figura 3. Red para la Recuperación de casos Mejorada (FCRN) del problema.

Tabla 7. La función de relevancia para la red construida.

ID Caso	Entidades de Información				
	Materiales	Requisitados	No llegan	Satisfactoriamente	Obra
1	1	1	1	0	0
2	1	0	0	0	1
3	1	0	0	0	0
5	0	0	1	0	0

De acuerdo con la función desarrollada por Chakraborti (Chakraborti et. al., 2006) los valores de relevancia efectiva de las IEs se calcularon con la siguiente fórmula:

$$\alpha_2(c) = \sum_{i=1}^s \Delta(e_i, c) \alpha_0(e_i)$$

En donde:

$\alpha_2(c)$  Representa el valor de relevancia efectiva de cada Entidad de Información (IE1, IE2, IE3, IE4, IE5, IE6) en cada uno de los casos respectivos.

$\Delta(e_i, c)$  Es una función de relevancia que relaciona todas IEs de la red con todos los casos respectivos. Cuando una IE esta presente en algún caso de la red el valor de  $\Delta(e_i, c)$  es igual a la unidad, en caso contrario el valor es cero.

$\alpha_0(e_i)$  Es función que representa la similitud semántica entre todas las Entidades de Información (IEs) de la red, es decir, esta función define que tanto se parecen entre sí todos los términos de la red. Este valor puede estar entre cero y 1.

Con el fin de ilustrar la aplicación de la fórmula antes mencionada, se procedió a calcular el valor de relevancia efectiva de todas las Entidades de Información (IE1, IE2, IE3, IE4, IE5, IE6) con respecto al caso C1 que se ilustra en la figura 3, resultando los siguientes valores:

$$\begin{aligned} \Delta(IE_1, C_1) &= \rho(IE_1, C_1)\sigma(IE_1, IE_1) \\ &+ \rho(IE_2, C_1)\sigma(IE_1, IE_2) \\ &+ \rho(IE_3, C_1)\sigma(IE_1, IE_3) \\ &+ \rho(IE_4, C_1)\sigma(IE_1, IE_4) \\ &+ \rho(IE_5, C_1)\sigma(IE_1, IE_5) \\ &+ \rho(IE_6, C_1)\sigma(IE_1, IE_6) \\ &= (1x1) + (1x0) + (1x0) \\ &+ (0x0) + (0x0) + (1x0) = 1 \end{aligned}$$

$$\begin{aligned} \Delta(IE_2, C_1) &= \rho(IE_2, C_1)\sigma(IE_2, IE_2) \\ &+ \rho(IE_3, C_1)\sigma(IE_2, IE_3) \\ &+ \rho(IE_4, C_1)\sigma(IE_2, IE_4) \\ &+ \rho(IE_5, C_1)\sigma(IE_2, IE_5) \\ &+ \rho(IE_6, C_1)\sigma(IE_2, IE_6) \\ &+ \rho(IE_1, C_1)\sigma(IE_2, IE_1) \\ &= (1x1) + (1x0) + (0x0) \\ &+ (0x0) + (1x0) + (1x0) = 1 \end{aligned}$$

$$\begin{aligned} \Delta(IE_3, C_1) &= \rho(IE_3, C_1)\sigma(IE_3, IE_3) \\ &+ \rho(IE_4, C_1)\sigma(IE_3, IE_4) \\ &+ \rho(IE_5, C_1)\sigma(IE_3, IE_5) \\ &+ \rho(IE_6, C_1)\sigma(IE_3, IE_6) \\ &+ \rho(IE_1, C_1)\sigma(IE_3, IE_1) \\ &+ \rho(IE_2, C_1)\sigma(IE_3, IE_2) \\ &= (1x1) + (0x0) + (0x0) \\ &+ (1x0) + (1x0) + (1x0) = 1 \end{aligned}$$

$$\begin{aligned} \Delta(IE_4, C_1) &= \rho(IE_4, C_1)\sigma(IE_4, IE_4) \\ &+ \rho(IE_5, C_1)\sigma(IE_4, IE_5) \\ &+ \rho(IE_6, C_1)\sigma(IE_4, IE_6) \\ &+ \rho(IE_1, C_1)\sigma(IE_4, IE_1) \\ &+ \rho(IE_2, C_1)\sigma(IE_4, IE_2) \\ &+ \rho(IE_3, C_1)\sigma(IE_4, IE_3) \\ &= (0x1) + (0x0) + (1x0.8) \\ &+ (1x0) + (1x0) + (1x0) = 0.8 \end{aligned}$$

$$\begin{aligned} \Delta(IE_5, C_1) &= \rho(IE_5, C_1)\sigma(IE_5, IE_5) \\ &+ \rho(IE_6, C_1)\sigma(IE_5, IE_6) \\ &+ \rho(IE_1, C_1)\sigma(IE_5, IE_1) \\ &+ \rho(IE_2, C_1)\sigma(IE_5, IE_2) \\ &+ \rho(IE_3, C_1)\sigma(IE_5, IE_3) \\ &+ \rho(IE_4, C_1)\sigma(IE_5, IE_4) \\ &= (0x1) + (1x0) + (1x0) \\ &+ (1x0) + (1x0) + (0x0) = 0.8 \end{aligned}$$

$$\begin{aligned} \Delta(IE_6, C_1) &= \rho(IE_6, C_1)\sigma(IE_6, IE_6) \\ &+ \rho(IE_1, C_1)\sigma(IE_6, IE_1) \\ &+ \rho(IE_2, C_1)\sigma(IE_6, IE_2) \\ &+ \rho(IE_3, C_1)\sigma(IE_6, IE_3) \\ &+ \rho(IE_4, C_1)\sigma(IE_6, IE_4) \\ &+ \rho(IE_5, C_1)\sigma(IE_6, IE_5) \\ &= (1x1) + (1x0) + (1x0) \\ &+ (1x0) + (0x0.8) + (0x0) = 1 \end{aligned}$$

Procediendo de la misma manera con los demás casos se obtuvieron los valores de relevancia efectiva mostrados en la Tabla 8.

Tabla 8. Valores de relevancia efectiva.

	<b>IE1</b>	<b>IE2</b>	<b>IE3</b>	<b>IE4</b>	<b>IE5</b>	<b>IE6</b>
<b>Caso 1</b>	1	1	1	0.8	0	1
<b>Caso2</b>	1	0	0	0	1	0
<b>Caso 3</b>	1	0	0	0	0	0
<b>Caso 5</b>	0	0	1	0	0	0

La relevancia efectiva nos muestra cuál es la importancia de cada una de las IEs con respecto a los casos recuperados. Para evaluar el nivel de similitud de los casos recuperados con respecto a la consulta inicial se procedió a calcular el módulo de los vectores fila de la matriz de relevancia efectiva.

Los resultados obtenidos fueron los siguientes:

Módulo del vector caso 1= 2.15

Módulo del vector caso 2= 1.41

Módulo del vector caso 3= 1.00

Módulo del vector caso 5 = 1.00

De los resultados obtenidos se pudo observar que el caso que más similitud tuvo con respecto a la consulta inicial fue el caso 1. El segundo caso en importancia fue el caso 2 y así sucesivamente. Estos resultados fueron similares a los que se obtuvieron utilizando el Modelo del Espacio de Vectores.

**Fase de Adaptación.** Para ejemplificar el funcionamiento de la fase de adaptación se utilizaron los dos casos recuperados de la base que tuvieron una mayor similitud con el problema inicial.

Recordando, el problema inicial fue el siguiente: *“Los materiales requisitados no llegan satisfactoriamente a la obra”*

Y los casos de la base recuperados de la base que tuvieron una mayor similitud con la descripción del problema anterior se muestran en la Tabla 9:

Para poder generar una solución automatizada el sistema debió realizar una serie de preguntas en función de las causas identificadas de los problemas recuperados. Para que el sistema

realice las preguntas se propuso utilizar la estructura computacional IF-THEN. Esta estructura es frecuentemente utilizada en el desarrollo de sistemas expertos que permiten pre-programar reglas que deben cumplirse. Para este ejemplo, permiten establecer cuál es la solución que se debe implementar de acuerdo con la causa del problema que ya ha sido identificada previamente en la investigación de campo.

Aplicando la estructura descrita anteriormente, la secuencia de preguntas quedó como:

**IF:** *¿Falta de coordinación entre el departamento de compras y el departamento de construcción?*

**THEN:** *“Establecer un canal de comunicación directa entre el jefe de compras y el jefe de construcción para resolver la falta de coordinación en las requisiciones”*

**IF:** *¿El proveedor no suministra oportunamente la cantidad de materiales solicitados?*

**THEN:** *“Establecer un canal de comunicación directa entre el jefe de compras y el representante del proveedor para coordinar el suministro de los materiales”*

Suponiendo que el usuario respondió afirmativamente a las dos preguntas con anticipación, el sistema propondría una solución compuesta al problema inicial se muestra en la Tabla 10.

El orden de importancia que tienen las soluciones propuestas por el sistema depende del grado de similitud de los casos recuperados con el problema original. En caso de que el usuario sólo haya respondido afirmativamente a una de las respuestas, el sistema únicamente propondrá una solución al problema. En caso de que el usuario no haya respondido afirmativamente a alguna respuesta, entonces el sistema no propondrá solución alguna al problema y automáticamente pasará a la siguiente fase del proceso.

**Fase de Validación.** Durante esta fase podrían presentarse tres diferentes situaciones con respecto a la solución propuesta por el sistema al usuario:

1.- Es completamente satisfactoria. En caso de ocurrir esta situación se pasaría automáticamente a la siguiente fase del proceso (Tabla 11).

Tabla 9. Casos de la base recuperados de la base que tuvieron una mayor similitud con “*Los materiales requisitados no llegan satisfactoriamente a la obra*”

Id caso	Descripción del Problema	Causas	Solución implementada
1	Los materiales requisitados no llegan oportunamente a la obra	Falta de coordinación entre el departamento de compras y el departamento de construcción	Establecer un canal de comunicación directa entre el jefe de compras y el jefe de construcción para resolver la falta de coordinación en las requisiciones
2	Existen retrasos en el suministro de materiales en obra	El proveedor no suministra oportunamente la cantidad de materiales solicitados	Establecer un canal de comunicación directa entre el jefe de compras y el representante del proveedor para coordinar el suministro de los materiales

Tabla 10. Suponiendo que el usuario respondió afirmativamente a las dos preguntas con anticipación.

Descripción del problema	Solución 1	Solución 2
Los materiales requisitados no llegan satisfactoriamente a la obra	Establecer un canal de comunicación directa entre el jefe de compras y el jefe de construcción para resolver la falta de coordinación en las requisiciones	Establecer un canal de comunicación directa entre el jefe de compras y el representante del proveedor para coordinar el suministro de los materiales

Tabla 11. Si es completamente satisfactoria.

Descripción del problema	Status	Solución 1	Solución 2
Los materiales requisitados no llegan satisfactoriamente a la obra	Se acepta	Establecer un canal de comunicación directa entre el jefe de compras y el jefe de construcción para resolver la falta de coordinación en las requisiciones	Establecer un canal de comunicación directa entre el jefe de compras y el representante del proveedor para coordinar el suministro de los materiales

Tabla 12. Si es parcialmente satisfactoria

Descripción del problema	Status	Solución 1
Los materiales requisitados no llegan satisfactoriamente a la obra	Parcialmente se acepta	Establecer un canal de comunicación directa entre el jefe de compras y el jefe de construcción para resolver la falta de coordinación en las requisiciones

Tabla 13. Si es insatisfactoria.

Descripción del problema	Status	Agregar Causa	Agregar Solución
Los materiales requisitados no llegan satisfactoriamente a la obra	No se acepta		

2.- Es parcialmente satisfactoria. En este caso el usuario tendría la oportunidad de poder “editar” la solución propuesta por el sistema. Durante el proceso de edición se pueden borrar, agregar o modificar las soluciones propuestas por el sistema Tabla 12.

3.- Es insatisfactoria. En este caso el usuario tendría la oportunidad de ingresar una nueva causa del problema y una nueva solución Tabla 13.

**Fase de Retención.** Durante esta fase el sistema almacena automáticamente el caso resuelto y validado de la fase anterior. Al nuevo caso resuelto se le asignó un nuevo identificador único (ID) y pasó a formar parte de la base de casos original. Suponiendo que durante la fase de validación la solución propuesta por el sistema no fue aceptada, habría que agregar una nueva solución para el problema. La base de casos final se muestra en la Tabla 14.

Tabla 14. Base de datos final.

Id caso	Descripción del Problema	Causas	Solución implementada
1	Los materiales requisitados no llegan oportunamente a la obra	Falta de coordinación entre el departamento de compras y el departamento de construcción	Establecer un canal de comunicación directa entre el jefe de compras y el jefe de construcción para resolver la falta de coordinación en las requisiciones
2	Existen retrasos en el suministro de materiales en obra	El proveedor no suministra oportunamente la cantidad de materiales solicitados	Establecer un canal de comunicación directa entre el jefe de compras y el representante del proveedor para coordinar el suministro de los materiales
3	Las requisiciones de materiales no se cumplen satisfactoriamente	La empresa no tiene dinero para solventar los pagos a proveedores	Llegar a una promesa de pago con los diferentes proveedores a través de un cheque post-fecha
4	Hay escasez de mano de obra	Existe mucha competencia en el mercado	Modificar los procedimientos constructivos de las viviendas para agilizar el tiempo de ejecución
5	Los materiales especiales solicitados no llegan a la obra	No se piden con suficiente tiempo de anticipación	Todos los materiales especiales se piden con 3 meses de anticipación
6	Los materiales requisitados no llegan satisfactoriamente a la obra	El residente de obra no elabora oportunamente sus requisiciones de materiales	Las requisiciones de materiales se elaboran con una semana de anticipación a la fecha de entrega en obra.

El caso número 6 representa el enriquecimiento de la base con la descripción de un nuevo problema, una nueva causa y una nueva solución implementada.

## DISCUSIÓN

Durante la aplicación de las entrevistas a los gerentes de proyectos se pudo notar que el 100% de los entrevistados respondió que utiliza la experiencia como la principal herramienta para resolver los problemas que se originan en la fase de ejecución de los proyectos de vivienda, lo que hace concluir que la experiencia sigue siendo la mejor maestra para los gerentes de proyecto.

También se observó que en el contexto de la CMV, cuando se presenta un problema, ya sea de suministro de insumos o de procedimientos constructivos, el problema en sí puede ser el mismo para dos empresas diferentes y lo único que varía es la causa o causas del problema, dependiendo de la empresa. Por ejemplo, una empresa de CMV de menor escala podría tener un problema de retrasos en el suministro de materiales a la obra siendo la causa del problema que la empresa tiene problemas financieros lo que hace difícil que pueda cumplir con sus compromisos económicos con el proveedor; por otro lado una empresa de CMV de mayor escala

podría tener el mismo problema de retrasos en el suministro de materiales, siendo la causa principal del problema que el proveedor no tiene la capacidad suficiente para suministrar el volumen de material solicitado o que no se hicieron las negociaciones necesarias. En ambos casos el problema es el mismo, pero se tienen causas diferentes según la empresa.

En lo que se refiere a la conceptualización del sistema RBC, es preciso señalar que una de las fases más importantes de su desarrollo es la fase de recuperación de los casos dentro de la base.

Para la recuperación de los casos con mayor similitud es necesario que el procedimiento de recuperación tenga la capacidad de extraer de la base todos los casos que tengan un parecido con la descripción del nuevo problema a resolver. En este trabajo se planteó utilizar dos enfoques diferentes. El modelo FCRN demostró ser un sistema de recuperación de casos más poderoso que el utilizado por el modelo VSM, debido a que el modelo FCRN tiene la capacidad de aceptar semántica de palabras, lo que ayuda a que si un caso se redacta con palabras similares a las palabras con las que están redactados los casos dentro de la base, va a ser posible recuperarlos satisfactoriamente.

La parte fundamental de un sistema CBR es la base de casos. Si se tiene una base de casos lo suficientemente amplia con los problemas más comunes que se presentan en el contexto de un mismo tipo de proyectos, se podrían obtener excelentes resultados durante la fase de recuperación, más aún, se podría tener un conocimiento más amplio de los problemas que ocurren durante la ejecución de los proyectos.

### **CONCLUSIONES**

De acuerdo con los resultados obtenidos se observó que un sistema RBC sí tiene la capacidad de ser “inteligente” para poder proponer una solución a un problema por sí solo, pero al final del proceso todavía se requiere de la intervención de un experto que pueda validar la solución propuesta.

Respecto a la automatización del sistema, es posible lograrlo utilizando las herramientas que están disponibles en el mercado. Para poder modelar todas las fases del sistema de razonamiento basado en casos (base, recuperación, adaptación y retención) se pueden utilizar lenguajes de programación orientada a objetos como el “C”, *JAVA* o aplicaciones

específicas como el *Jcolibri*, el cual fue diseñado específicamente para construir sistemas RBC. Sin importar cuál sea el lenguaje de programación utilizado, es importante que el sistema RBC sea construido de tal forma que garantice los mejores resultados posibles para solucionar los problemas que se presenten. Es importante que la automatización del sistema integre reglas y operaciones que permitan evaluar la semántica de las palabras ya que es común que en la industria de la construcción se utilicen palabras coloquiales del sector o de la región.

Para que un sistema como el que se propone sea útil dentro de la industria de la construcción es necesario que las empresas lleven a cabo un registro de los problemas que se resuelven durante la ejecución de los proyectos. Preservar la experiencia obtenida de la resolución de problemas es muy importante para lograr el aprendizaje organizacional. Un sistema RBC como el que se propone en este trabajo es perfectamente aplicable en todos los niveles de la empresa constructora (operativo, administrativo, etc.).

### **BIBLIOGRAFIA**

- Arshadi, N., Badie K. (2001) “A Compositional Approach to Solution Adaptation in Case-Based Reasoning and its Application to Tutoring Library”, Software Engineering Group and Iran Telecom Research Center, Tehran, Iran.
- Bergmann R., Kolodner, J.L. Plaza, E., (2005) Representation in case-based reasoning. “The Knowledge Engineering Review”, 1-4.
- Busch T. (1999) “Sistemas Tecnológicos”, Editorial Aique, Buenos Aires, Argentina
- Chakraborti, S., Lothian R., Wiratunga N., Orecchioni A., Watt S. (2006) Fast Case Retrieval Nets. Memorias del Congreso Internacional de “Case Based Reasoning”, 400-414.
- Cole R. (1991) Tight Bounds on the Complexity of the Boyer-Moore Algorithm. Memorias del Segundo “Simposium Anual sobre Algoritmos Discretos”, 224-233.
- Goh Y.M., Chua D.K.H. (2009) Case-Based Reasoning for Construction Hazard Identification: Case Representation and Retrieval, “Journal of Construction Engineering and Management”, A.S.C.E., 135(11).
- Han G. R., Lee H.-S., Park M. (2007) Construction Planning Method Using Cased-Based Reasoning (CONPLA-CBR), “Journal of Computing in Civil Engineering”, A.S.C.E., 21(6), 410-422.
- Kolodner J.L. (1993) “Case-Based Reasoning”, Editorial Morgan Kaufmann, San Mateo.

Luger G. F. (2005) "Artificial Intelligence". Structures and Strategies For Complex Problem Solving, 5<sup>ta</sup> Edición, Addison Wesley.

Richard C., Sears G., Keoki S. (2000) "Construction Project Management", 4<sup>ta</sup> edición, John Wiley & Sons, U.S.A.

Salton G., Buckley C. (1988) Term-Weighting Approaches in Automatic Text Retrieval, "Information Processing & Management", 24(5)513-523.

Watson I. (1997) "Applying Case-Based Reasoning: Techniques for Enterprises Systems", 1<sup>ra</sup> Edición, Morgan Kaufmann Publishers, U.S.A.

---

Este documento debe citarse como: Cruz Ac J. M., González Fajardo J. A., Zaragoza Grifé J. N. (2013). **Un sistema de razonamiento basado en casos para apoyar la toma de decisiones en la industria de la construcción**. Ingeniería, Revista Académica de la FI-UADY, 17-2, pp 111-125, ISSN 1665-529-X.