

FLEXIBILIZACION DE BUSQUEDA EN BASES DE DATOS RELACIONALES

Plüss Jorge, Bender Cristina, Casali Ana, Corti Rosa, D'Agostino Estela & Deco Claudia (*)

RESUMEN

En base de datos relacionales existe un amplio espectro de posibilidades de búsqueda. Para realizar la recuperación de información, se han desarrollado diferentes métodos que permiten al usuario especificar exactamente qué datos desea recuperar de una base de datos. El más popular es el Lenguaje de Consulta Estructurado SQL. A pesar de ser un lenguaje relacionamente completo, tiene ciertas limitaciones. Una de ellas, es que el usuario está forzado a establecer en forma precisa el criterio de búsqueda, obteniendo exactamente y solamente lo que ha pedido. En todos estos casos las tuplas resultantes en una recuperación tienen un mismo nivel de satisfacción respecto a la búsqueda planteada. SQL plantea una búsqueda dicotómica en el sentido de que una tupla integra el conjunto resultante si satisface totalmente la condición de búsqueda.

Este trabajo presenta una alternativa de flexibilización de la recuperación de información basándose en la lógica difusa. A través de ella, se puede permitir la representación de datos de entrada imprecisos y recuperar las tuplas que satisfagan en alguna medida la condición de búsqueda.

El planteo de esta flexibilización surge de distintas aplicaciones donde es necesario realizar una búsqueda a partir de información imprecisa, pudiendo ser esta imprecisión difusa. Esto ocurre cuando el concepto buscado está representado por un conjunto de valores y un coeficiente de compatibilidad asociado a cada uno de ellos, tomando estos coeficientes valores en el intervalo $[0,1]$. Esta imprecisión puede representarse mediante un conjunto difuso y la lógica difusa resulta muy adecuada para expresar consultas en bases de datos a partir de este tipo de información.

Si bien la imprecisión puede darse en datos de naturaleza cualitativa o cuantitativa, su tratamiento se analiza en forma individual. En cada caso, se describen dos alternativas, dependiendo si la imprecisión asociada a un dato ingresado por el usuario está predeterminada, o es modelizada por el mismo usuario. La primera alternativa, consiste en que el sistema determina la extensión difusa del valor buscado, armando el conjunto de valores y compatibilidades, de acuerdo al valor ingresado por el usuario. La otra, es que el usuario defina el término impreciso, determinando el conjunto de valores junto a los coeficientes de compatibilidad. En ambos casos se recuperan los registros semánticamente aproximados a las condiciones de búsqueda

La búsqueda se formaliza a través de predicados difusos. Se utiliza la lógica difusa para resolver la combinación de los distintos coeficientes de compatibilidad, a través de distintos conectivos lógicos. De esta forma se obtienen como resultado tuplas que satisfacen la condición de búsqueda en distintos grados, estos coeficientes formarán también parte del resultado, ordenando las tuplas de acuerdo a su grado de aproximación.

PALABRAS CLAVES:

Base de datos - Lógica difusa - Sistemas de Información .

* Depto de Sistemas e Informática / Facultad de Cs. Exactas Ingeniería y Agrimensura / UNR

Av Pellegrini 250 (2000) Rosario, Argentina.

Tel / Fax (041) 217998 - email: jpluss@dsifei.edu.ar