



estudios de postgrado
en computación



Bases de datos avanzadas

Universidad de Los Andes

Postgrado en Computación

Prof. Isabel M. Besembel Carrera

***Unidad IV. Sesiones 25 y 26. Bases de datos
espaciales.***



Estructuras de datos espaciales

- Datos espaciales es el término utilizado para describir los datos que pertenecen al espacio ocupado por los objetos en la BD.
- Ellos pueden ser geométricos y variados. D. Peuquet lo define como el término que se aplica a cualquier dato de un fenómeno distribuido en 2-, 3- o n-dimensiones que son ortogonales y homogéneas.
- Los datos espaciales consisten del conjunto de objetos espaciales construidos con: puntos, líneas, regiones, rectángulos, superficies, volúmenes, etc. Ejm: Ciudades, ríos, estados, etc.
- El espacio puede ser discreto o continuo, limitado o no
- Si se considera limitado normalmente se aplica un rango de 70x106 mts.
- También puede ser absoluto o relativo.



Estructuras de datos espaciales

- Los objetos espaciales contiene además datos no espaciales como nombre de la ciudad, número de habitantes, etc.
- Las consultas que pueden ser hechas sobre este tipo de datos normalmente involucran relaciones entre los objetos, dadas por su ubicación u ocupación en el espacio, lo cual indica que la consulta es espacial.
- Para responder consultas espaciales se utilizan los métodos espaciales de indexación.
 - ✓ Ejm: encontrar todas las ciudades dentro de la región andina.
- Los métodos se clasifican como: de puntos multidimensionales (MAPM) y espaciales (MAE).



Métodos de acceso

➤ Por puntos multidimensionales

- ✓ Divide el espacio de datos en regiones de tal forma que todos los registros de una región se almacenan en la misma página de datos.
- ✓ Dicha división se hace dependiendo del número de puntos en cada subregión que por lo general son disjuntas.
- ✓ Representantes de estos son:
 - el árbol_Bk-d,
 - el archivo malla,
 - el archivo BANG (Balanced And Nested Grid File),
 - el árbol compañero (Buddy tree) y
 - los árboles LSD

Métodos de acceso

- Ellos se clasifican de acuerdo a tres factores:
 - ✓ La posición del hiperplano de corte puede ser fija o no.
 - ✓ La dimensión de corte que puede ser solo una o sobre todas (quadrees)
 - ✓ La localización del hiperplano que puede afectar a todas las regiones en una dirección dada o dentro de la región de corte.

- Árboles_Bk-d (Robinson, 1981)
- Archivo malla (Nievergelt, Hinterberger y Sevcik, 1984)
- Árbol compañero (Seeger y Kriegel, 1990)



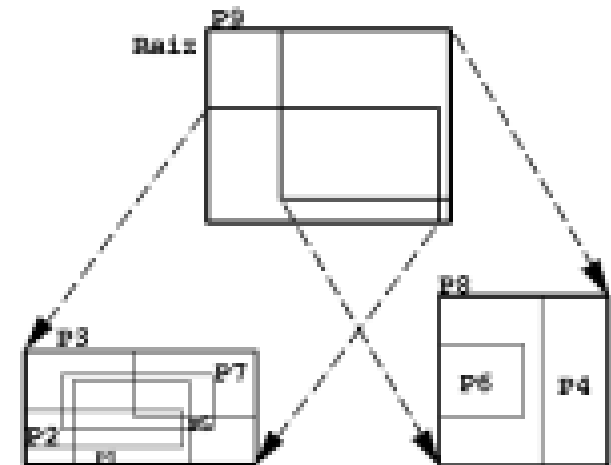
Métodos de acceso

➤ Espaciales:

- ✓ Organizan objetos espaciales multidimensionales.
 - ✓ Se categorizan como:
 - Métodos que transforman rectángulos en puntos de mayores dimensiones.
 - Métodos que dividen el espacio original en subregiones solapadas o disjuntas.
 - Ejemplos:
 - el árbol_R,
 - el árbol_R+ y
 - el árbol celda.
-
- Árbol binario k-d espacial (Ooi et al. 1991)
 - Árbol_R (Guttman, 1984)
 - Árbol_R+ (Sellis, Roussopoulos y Faloutsos, 1987)
 - Árbol celda (Günther, 1989)

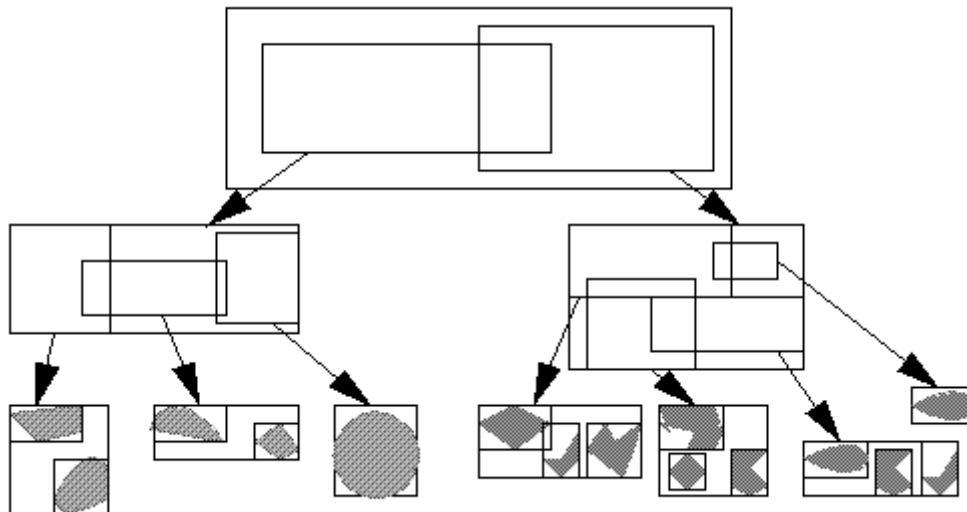
Árbol_R

- Los nodos hojas contienen entre m y M entradas a menos que ella sea la raíz.
- Cada entrada en un nodo hoja es de la forma $(I, \text{identificador})$, I es el rectángulo más pequeño que espacialmente contiene al objeto n -dimensional referenciado por el identificador.
- Cada nodo rama tiene entre m y M hijos a menos que sea la raíz.
- Cada entrada en un nodo rama es de la forma $(I, \text{punteroAlHijo})$, I es el rectángulo más pequeño que espacialmente contiene los rectángulos en el nodo hijo.
- La raíz tiene al menos dos hijos a menos que sea una hoja.
- Todos los nodos aparecen en el mismo nivel.



Árbol_R bidimensional

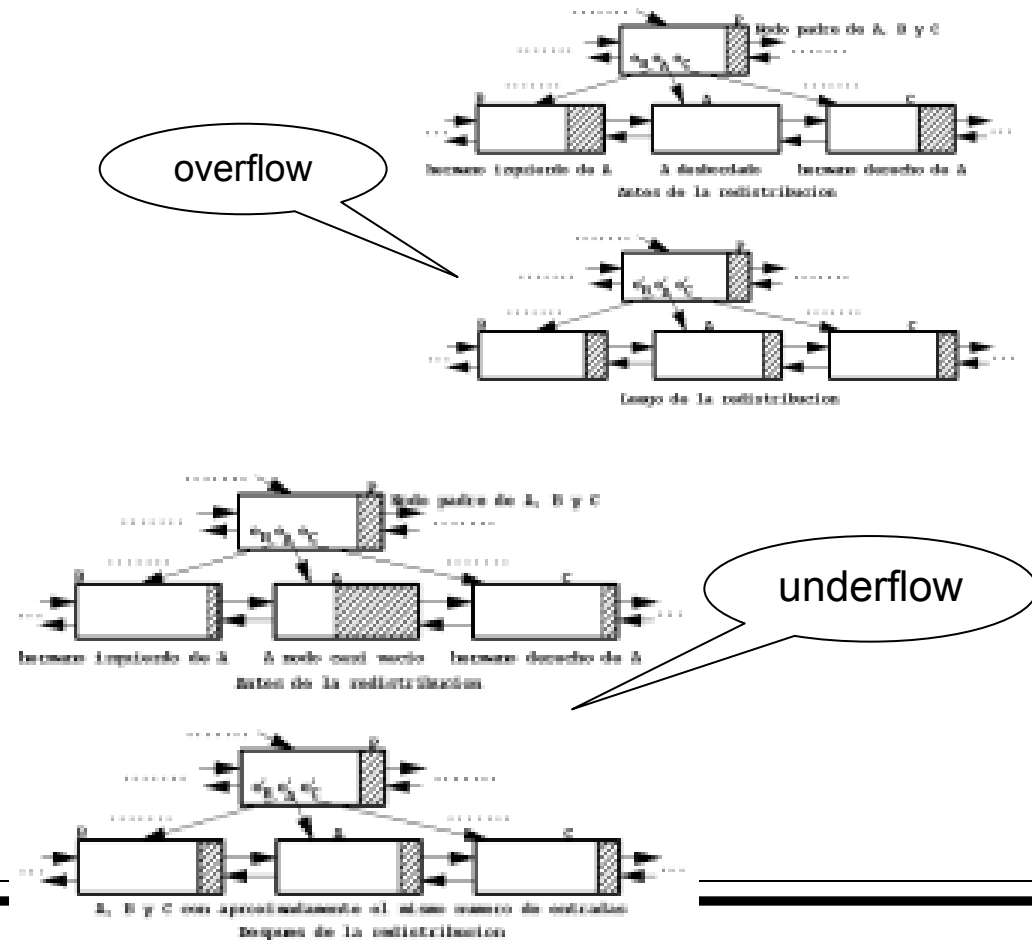
Árbol_R



- Generalización del árbol_B a n dimensiones
- Indexación de los objetos por su mbr (minimum bounding rectangle)
- Balanceado en altura
- Completamente dinámico en inserciones y eliminaciones

Árbol_R

- **nA** el número de entradas de nodo actual,
- **nl** el número de entradas del hermano a la izquierda y
- **nD** el de su hermano a la derecha,
- Si $\lceil (nA+nI+nD)/3 \rceil \geq m$ se realiza una redistribución, sino se realiza la fisión del nodo.
- Si $\lceil (nA+nI+nD)/2 \rceil < M$ se realiza la eliminación del nodo actual.



Una consulta expresada en CPOO posee la siguiente forma:

{Cláusula destino, Cláusula rango, Cláusula cualificación}

donde:

Cláusula destino: que recuperar al realizar la interrogación.

Cláusula rango: enlace de una variable objeto con la clase

Cláusula cualificación: fórmula (combinación de predicados lógicos)

- El conjunto de operadores antes definido debe ser expandido para contener a los operadores propios de los objetos espaciales.
- Debe incluir operaciones sobre regiones, líneas, puntos y sus combinaciones.
- Para ello se deben definir las relaciones entre los objetos espaciales
- *Relaciones topológicas:*
 - ✓ Se basan en las nociones de borde, interior y acotamiento de conjuntos abiertos.



➤ *Relaciones topológicas:*

- ✓ “Sea M un conjunto y N un subconjunto de potencia igual a la del conjunto M , entonces el par (M, N) se denomina un espacio topológico si cumple lo siguiente:

$$\emptyset \in N, M \in N$$

$$O_1, O_2 \in N, \text{ entonces } O_1 \cup O_2 \in N$$

$$R \in N, \text{ entonces } \bigcup_{O \in R} O \in N$$

Sea N la topología sobre M , los elementos de N se denominan conjuntos abiertos y sus complementos se denominan conjuntos cerrados, los elementos de M son puntos” [Kain90]

- ✓ “Dado $Y \subset X$, el interior de Y es la unión de todos los conjuntos abiertos que se encuentran contenidos en Y y se denota por Y° . El acotamiento se denota por Y' y representa la intersección de todos los conjuntos cerrados que contiene Y . El borde de Y se denota por ΨY ” [Kain90]

Relaciones topológicas

Relación topológica	$\Psi A \cap \Psi B$	$A^\circ \cap B^\circ$	$\Psi A \cap B^\circ$	$A^\circ \cap \Psi B$
A y B son regiones disjuntas (disjoint)	\emptyset	\emptyset	\emptyset	\emptyset
A y B se tocan (meet)	$\neg \emptyset$	\emptyset	\emptyset	\emptyset
A y B son iguales (match)	$\neg \emptyset$	$\neg \emptyset$	\emptyset	\emptyset
A dentro de B (inside)	\emptyset	$\neg \emptyset$	$\neg \emptyset$	\emptyset
A cubierto por B (coveredBy)	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	\emptyset
A contiene B (contain)	\emptyset	$\neg \emptyset$	\emptyset	$\neg \emptyset$
A cubre B (cover)	$\neg \emptyset$	$\neg \emptyset$	\emptyset	$\neg \emptyset$
A y B se solapan con bordes disjuntos	\emptyset	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$
A y B se solapan (overlap)	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$

- Ellos se incluyen como operadores de comparación de objetos espaciales en los predicados de búsqueda en una consulta espacial, o como funciones de post-procesamiento una vez que el resultado de una consulta espacial se ha obtenido.
- Ellos devuelven un valor lógico que establece la relación topológica entre dos objetos espaciales.
- Si los operandos son objetos no espaciales se abortará el proceso en la fase de chequeo de tipos
- De las 16 posibilidades definidas por Egenhofer, las anteriores se aplican a líneas y a regiones.



Relaciones de dirección: Indican si un objeto espacial está arriba, abajo, a la izquierda o a la derecha de otro [KGK93].

Operador	Relación
k.Arriba(m)	$k.y_i \geq m.y_s$
k.Izquierdo(m)	$k.x_i \geq m.x_s$
k.Abajo(m)	$k.y_s \geq m.y_i$
k.Derecha(m)	$k.x_s \geq m.x_i$
k.dArriba(m)	$k.Arriba(m) \wedge m.x.cubre(k.x)$
k.dIzquierda(m)	$k.Izquierdo(m) \wedge m.y.cubre(k.y)$
k.dAbajo(m)	$k.Abajo(m) \wedge m.x.cubre(k.x)$
k.dDerecha(m)	$k.Derecha(m) \wedge m.y.cubre(k.y)$
k.arribalIzquierda(m)	$k.Arriba(m) \wedge k.Izquierdo(m)$
k.abajolIzquierda(m)	$k.Abajo(m) \wedge k.Izquierdo(m)$
k.arribaDerecha(m)	$k.Arriba(m) \wedge k.Derecha(m)$
k.abajoDerecha(m)	$k.Abajo(m) \wedge k.Derecha(m)$



CPEOO

- El procesamiento de productos espaciales se realiza basándose en el algoritmo hash híbrido
- Ejemplo: Prototipo
- Recupere los nombres, latitud y longitud de todas las ciudades localizadas al este de Lagunillas.

MARINA Lenguaje de consulta gráfico para manejo de datos espacio temporales

Consulta en CPOO:
 { 9.Nom_Ciu, 9.Lat_Min , 9.Lon_Min ; 9/Ciudad ; ESTE "Lagunillas" }

Clase: Tipos

ADD

Base de Datos	Atributos	Sub Clases
Persona	Ced_Per	Desempleado
Vehículo	Nom_Per	Profesional
Estado	Ape_Per	Estudiante
Accidente	Dir_Per	
	Nom_Ciu	

Símbolos { } () " ; .

Símbolos	59	44
Timotes	59	44
Chachopo	56	46
Pueblo Llano	55	40
Santo Domingo	51	42
Las Piedras	53	37
Apartaderos	47	48
San Rafael	45	52
Mucuchies	44	55
Mucurubá	42	0
Tabay	38	4
Mérida	36	7
Ejido	33	13
San Juan	30	21
Pico El Aguila	50	50
La Venta	55	46
Escambay	42	59
Cacute	40	0
Chorros de Milla	37	7
San Jacinto	35	8

Combine los operadores espaciales y temporales para realizar la consulta

Sintaxis del Cálculo de Predicado Orientado por Objeto (CPOO):
 {Cláusula Destino; Cláusula Rango; Cláusula Cualificación}

- CLAUSULA DESTINO: Especifica lo que se deberá recuperar al realizar la interrogación. (SELECT en SQL)
- CLAUSULA RANGO : Define el enlace de una variable objeto con la clase. (En el ejemplo: 1 es la Variable Objeto y Persona es la Clase, similar al FROM de SQL)
- CLAUSULA CUALIFICACION: Esta definida por una fórmula expresada por la combinación de predicados booleanos, espaciales, temporales, los cuales estan conectados por operadores lógicos (El WHERE en SQL).

Por Ejemplo: Recupere los nombres de las ciudades que estan al este de Mérida.

{ 9.Nom_Ciu ; 9/Ciudad ; ESTE "Mérida" }