

Métodos de Acesso Métrico

Arthur Emanuel de O. Carosia
Profa. Dra. Cristina Dutra de Aguiar Ciferri

Índice

Conceitos Básicos
Métodos de Acesso
Onion-tree

Índice

Conceitos Básicos
Métodos de Acesso
Onion-tree

Dados Complexos

Não possuem relação de ordem total, ao contrário dos dados convencionais.

Ex: dados multimídia.

Dados Complexos

Não possuem relação de **ordem total**, ao contrário dos dados convencionais.

Ex: dados multimídia.



>
<
?



Dados Complexos

Não possuem relação de ordem total, ao contrário dos dados convencionais.

Comparados por **similaridade**.

Dados Complexos

Não possuem relação de **ordem total**, ao contrário dos dados convencionais.

Comparados por **similaridade**.

Naturalmente representados em um **espaço métrico**.

Espaço Métrico

Definido como $M = \langle S, d \rangle$

S: **domínio** de elementos

d: função de **distância** sobre os elementos de S

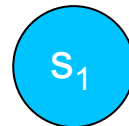
- Distância entre **dois elementos** do domínio
- Quanto **menor** a distância, **mais similares** os elementos são.

Espaço Métrico

A função de **distância** deve atender às condições:

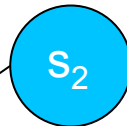
Identidade

$$d(s_1, s_1) = 0$$



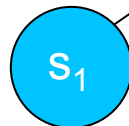
Simetria

$$d(s_1, s_2) = d(s_2, s_1)$$



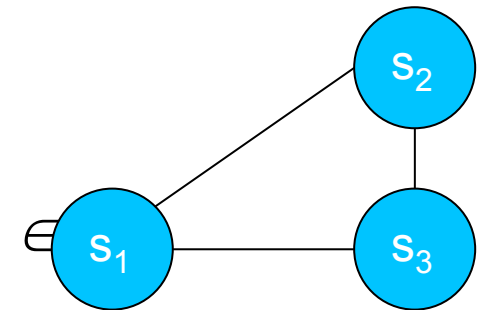
Não negatividade

$$0 \leq d(s_1, s_2) \leq \infty$$



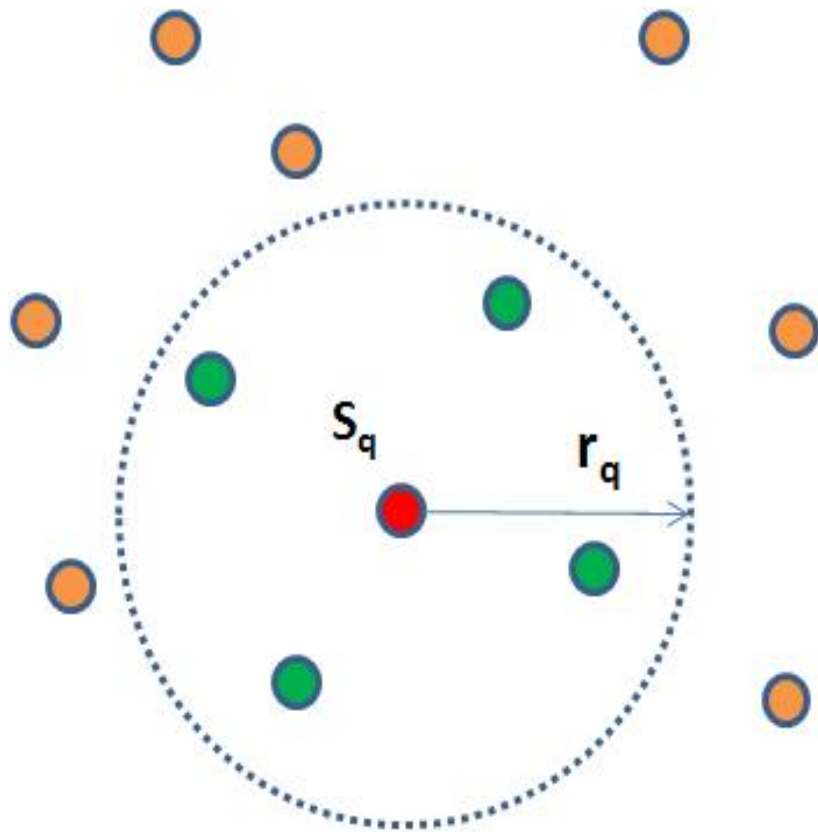
Desigualdade triangular

$$d(s_1, s_2) \leq d(s_1, s_3) + d(s_2, s_3) \text{ onde } s_1, s_2, s_3$$

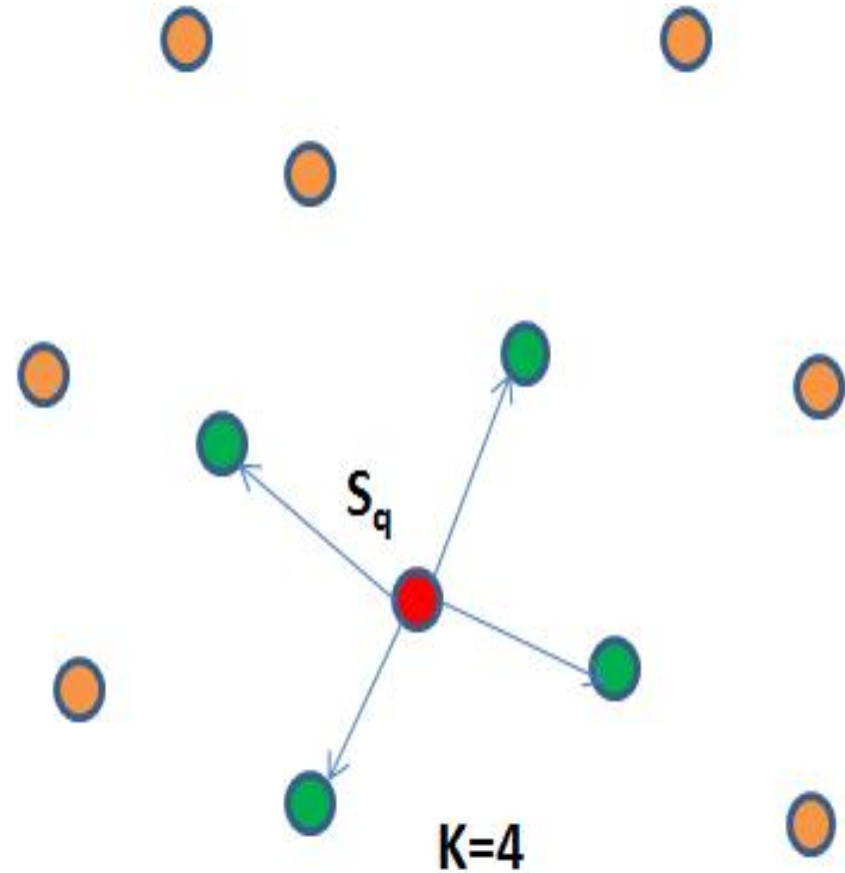


Consultas por Similaridade

Consulta por Abrangência



Consulta aos k vizinhos mais próximos



Índice

Conceitos Básicos
Métodos de Acesso
Onion-tree

Métodos de Acesso

Estrutura de dados e conjunto de **algoritmos** que manipulam esta estrutura.

Métodos de Acesso

Estrutura de dados e conjunto de **algoritmos** que manipulam esta estrutura.

- Inserção
- Remoção
- Busca
- ***Bulk-loading***

Métodos de Acesso Multidimensionais

Estrutura de dados com o objetivo de **indexar** coleções de números que representam **objetos espaciais**.

Rápida recuperação de objetos que satisfaçam **relacionamentos topológicos**.

Uso de **aproximações** para indexar os dados.

Métodos de Acesso Métricos (Bola)

Particionam o espaço de dados em várias regiões.

Escolhe **elementos representantes** e agrupam os demais elementos à sua volta.

Hierarquia de elementos representantes.

Métodos de Acesso

Nome	Tipo de Método de Acesso	Memória	Conceito de Divisão do Espaço	Algoritmos Descritos	Tratamento de Sobreposição	Construção
-------------	---------------------------------	----------------	--------------------------------------	-----------------------------	-----------------------------------	-------------------

Métodos de Acesso

Nome	Tipo de Método de Acesso	Memória	Conceito de Divisão do Espaço	Algoritmos Descritos	Tratamento de Sobreposição	Construção
R-tree	Multidimensional	Secundária	Aproximação (MBR)	Inserção e Busca	Sim	Bottom-up

Guttman, A. (1984). R-trees: a dynamic index structure for spatial searching. In Proceedings of the 1984 ACM SIGMOD International Conference on Management of Data, pages 47–57, Boston, USA.

Métodos de Acesso

Nome	Tipo de Método de Acesso	Memória	Conceito de Divisão do Espaço	Algoritmos Descritos	Tratamento de Sobreposição	Construção
R-tree	Multidimensional	Secundária	Aproximação (MBR)	Inserção e Busca	Sim	Bottom-up
M-tree	Métrico	Secundária	Bola	Inserção e Busca	Sim	Bottom-up

Ciaccia, P., Patella, M., and Zezula, P. (1997). M-tree: an efficient access method for similarity search in metric spaces. In Proceedings of the 23rd International Conference on Very Large Data Bases (VLDB), pages 426–435, Athens, Greece.

Métodos de Acesso

Nome	Tipo de Método de Acesso	Memória	Conceito de Divisão do Espaço	Algoritmos Descritos	Tratamento de Sobreposição	Construção
R-tree	Multidimensional	Secundária	Aproximação (MBR)	Inserção e Busca	Sim	Bottom-up
M-tree	Métrico	Secundária	Bola	Inserção e Busca	Sim	Bottom-up
Slim-tree	Métrico	Secundária	Bola	Inserção e Busca	Sim	Bottom-up

Traina-Jr, C., Traina, A. J. M., Seeger, B., and Faloutsos, C. (2000). Slim-trees: high performance metric trees minimizing overlap between nodes. In Proceedings of the 7th International Conference on Extending Database Technology (EDBT), pages 51–65, Konstanz, Germany.

Métodos de Acesso

Nome	Tipo de Método de Acesso	Memória	Conceito de Divisão do Espaço	Algoritmos Descritos	Tratamento de Sobreposição	Construção
R-tree	Multidimensional	Secundária	Aproximação (MBR)	Inserção e Busca	Sim	Bottom-up
M-tree	Métrico	Secundária	Bola	Inserção e Busca	Sim	Bottom-up
Slim-tree	Métrico	Secundária	Bola	Inserção e Busca	Sim	Bottom-up
MM-tree	Métrico	Primária	Bola	Inserção, Semi-Balanceamento e Busca	Desnecessário, sem sobreposição	Top-down

Pola, I. R. V., Traina-Jr, C., and Traina, A. J. M. (2007). The MM-tree: A memory-based metric tree without overlap between nodes. In Proceedings of the 12th East European Conference on Advances in Databases and Information Systems (ADBIS), pages 157–171.

Métodos de Acesso

Nome	Tipo de Método de Acesso	Memória	Conceito de Divisão do Espaço	Algoritmos Descritos	Tratamento de Sobreposição	Construção
R-tree	Multidimensional	Secundária	Aproximação (MBR)	Inserção e Busca	Sim	Bottom-up
M-tree	Métrico	Secundária	Bola	Inserção e Busca	Sim	Bottom-up
Slim-tree	Métrico	Secundária	Bola	Inserção e Busca	Sim	Bottom-up
MM-tree	Métrico	Primária	Bola	Inserção, Semi-Balanceamento e Busca	Desnecessário, sem sobreposição	Top-down
Onion-tree	Métrico	Primária	Bola	Inserção, Substituição de Pivôs e Busca	Desnecessário, sem sobreposição	Top-down

Carelo, C., Pola, I., Ciferri, R., Traina, A., Traina Jr, C., and Ciferri, C. (2011). Slicing the metric space to provide quick indexing of complex data in the main memory. *Information Systems*, pages 79-98.

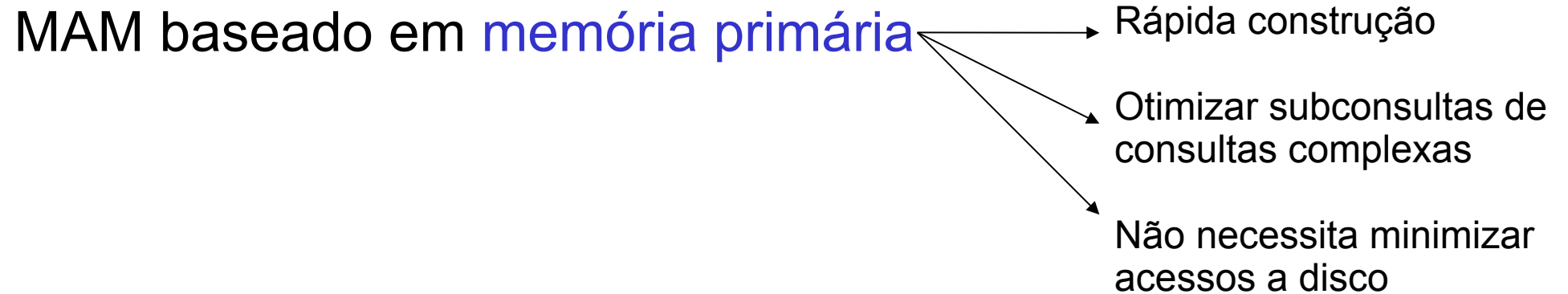
Índice

Conceitos Básicos
Métodos de Acesso
Onion-tree

Onion-tree

MAM baseado em memória primária

Onion-tree



Onion-tree

MAM baseado em **memória primária**

Pode dividir o espaço do nó em **mais que 4 regiões disjuntas**, conforme a necessidade

Onion-tree

MAM baseado em **memória primária**

Pode dividir o espaço do nó em **mais que 4 regiões disjuntas**, conforme a necessidade

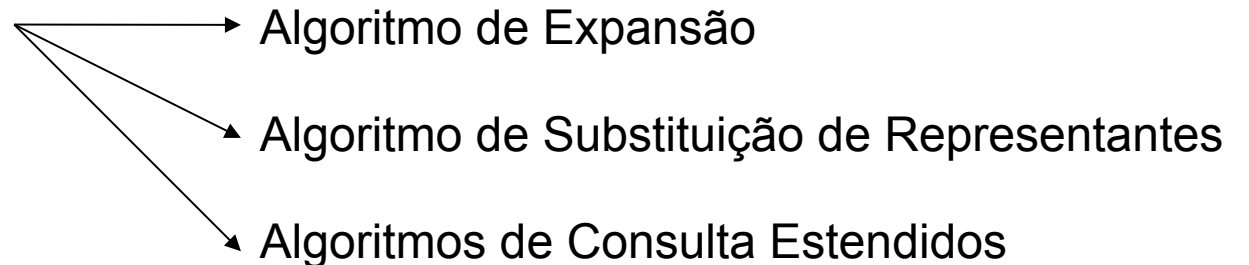
Extensão da MM-tree

Onion-tree

MAM baseado em **memória primária**

Pode dividir o espaço do nó em **mais que 4 regiões disjuntas**, conforme a necessidade

Extensão da MM-tree



Onion-tree

MAM baseado em **memória primária**

Pode dividir o espaço do nó em **mais que 4 regiões disjuntas**, conforme a necessidade

Extensão da MM-tree

Não possui operações de:

- *Bulk-loading*
- Remoção

Onion-tree

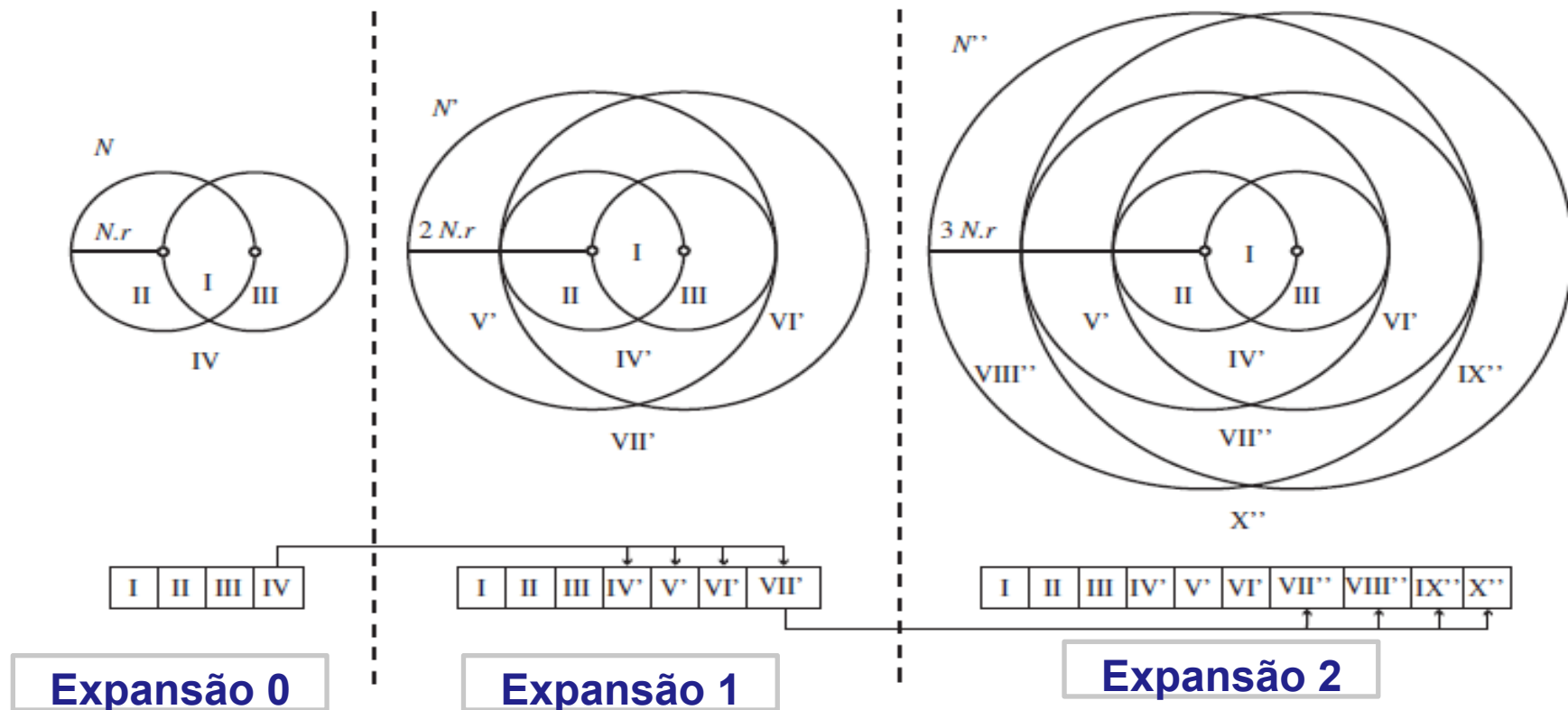
Algoritmo de expansão

Aumenta de forma flexível a quantidade de regiões disjuntas em cada nó

Onion-tree

Algoritmo de expansão

Aumenta de forma flexível a quantidade de regiões disjuntas em cada nó



Onion-tree

Políticas de expansão:

Fixa (F-Onion-tree)

Variável (V-Onion-tree)

Onion-tree

Políticas de expansão:

Fixa (F-Onion-tree)

Variável (V-Onion-tree) → Keep
Small
Strategy

Onion-tree

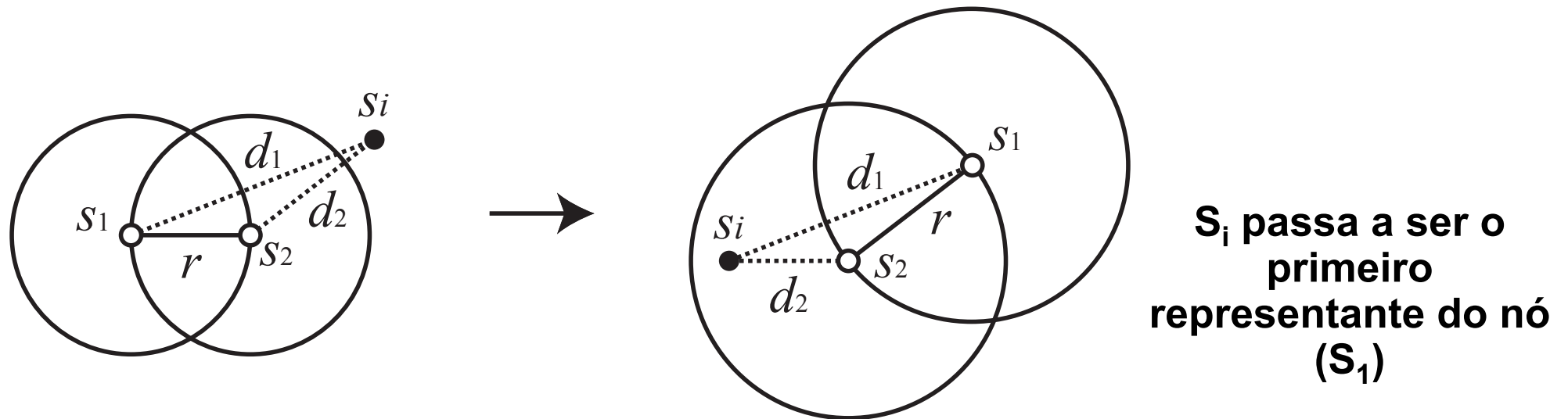
Algoritmo de substituição de representantes

Na inserção, identifica se o subespaço do nó folha pode ser melhor dividido a partir de outros representantes

Onion-tree

Algoritmo de substituição de representantes

Na inserção, identifica se o subespaço do nó folha pode ser melhor dividido a partir de outros representantes



Onion-tree

Algoritmo de substituição de representantes

Políticas de substituição

1. *Keep Small Strategy*
2. *Minimize Expansions*
3. *Maximize Expansions*

Onion-tree

Algoritmos de consulta estendidos

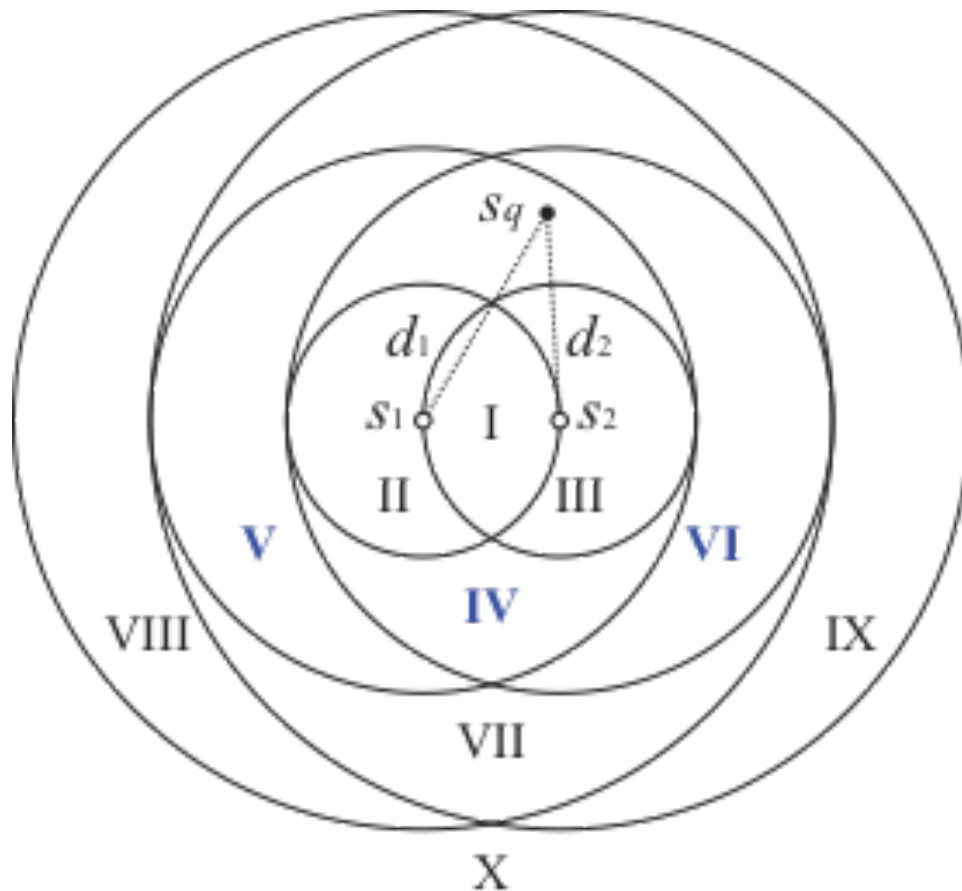
Estende os algoritmos de consulta da MM-tree para considerar as regiões adicionais criadas

Algoritmo de busca pelos k -vizinhos mais próximos: nova seqüência para a visitação das regiões do nó.

Onion-tree

Algoritmos de consulta estendidos – KNN

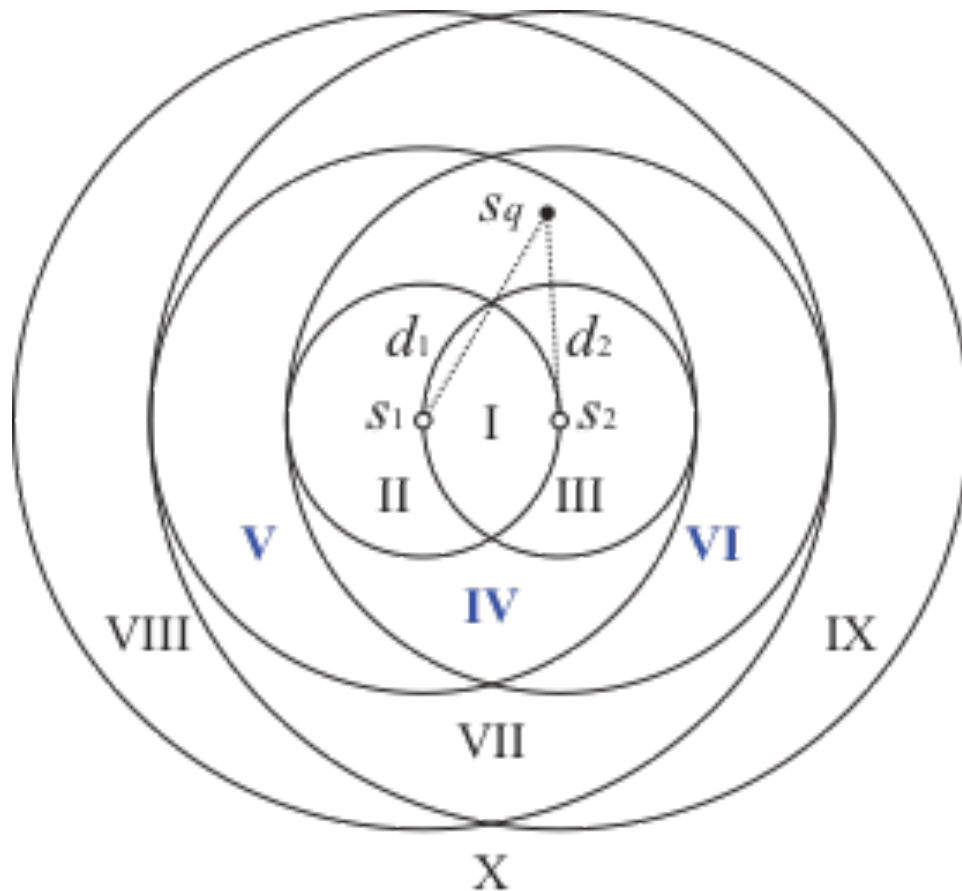
1. Visita a expansão E na qual se encontra o elemento S_q
2. Visita as outras regiões, de acordo com a proximidade ao elemento da query
3. Expansões $E - 1$ e $E + 1$, Expansões $E - 2$ e $E + 2$



Onion-tree

Algoritmos de consulta estendidos – KNN

1. Visita a expansão E na qual se encontra o elemento S_q
2. Visita as outras regiões, de acordo com a proximidade ao elemento da query
3. Expansões $E - 1$ e $E + 1$, Expansões $E - 2$ e $E + 2$

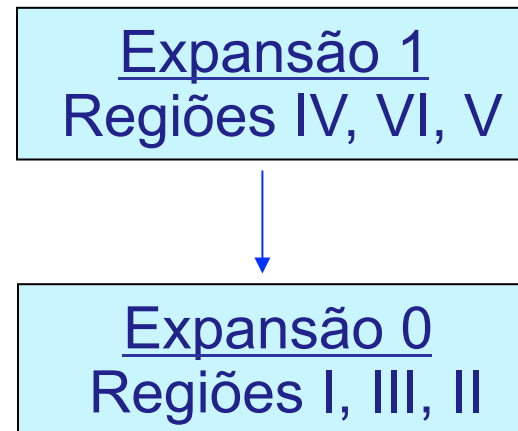
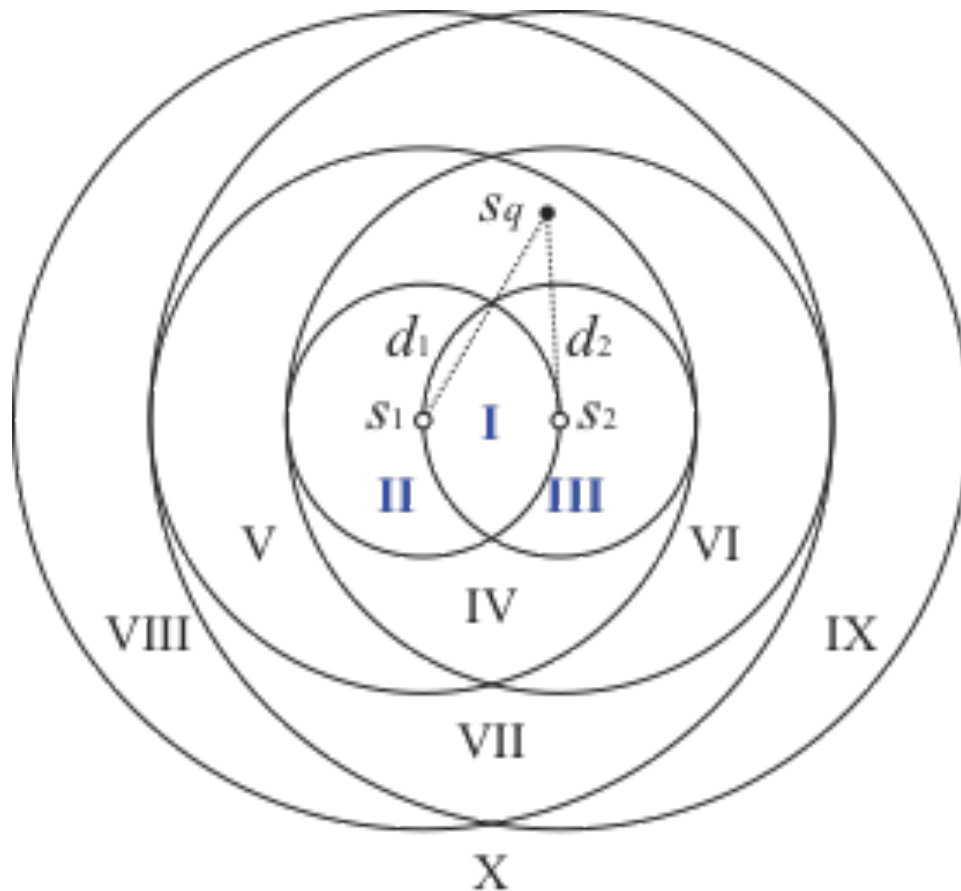


Expansão 1
Regiões IV, VI, V

Onion-tree

Algoritmos de consulta estendidos – KNN

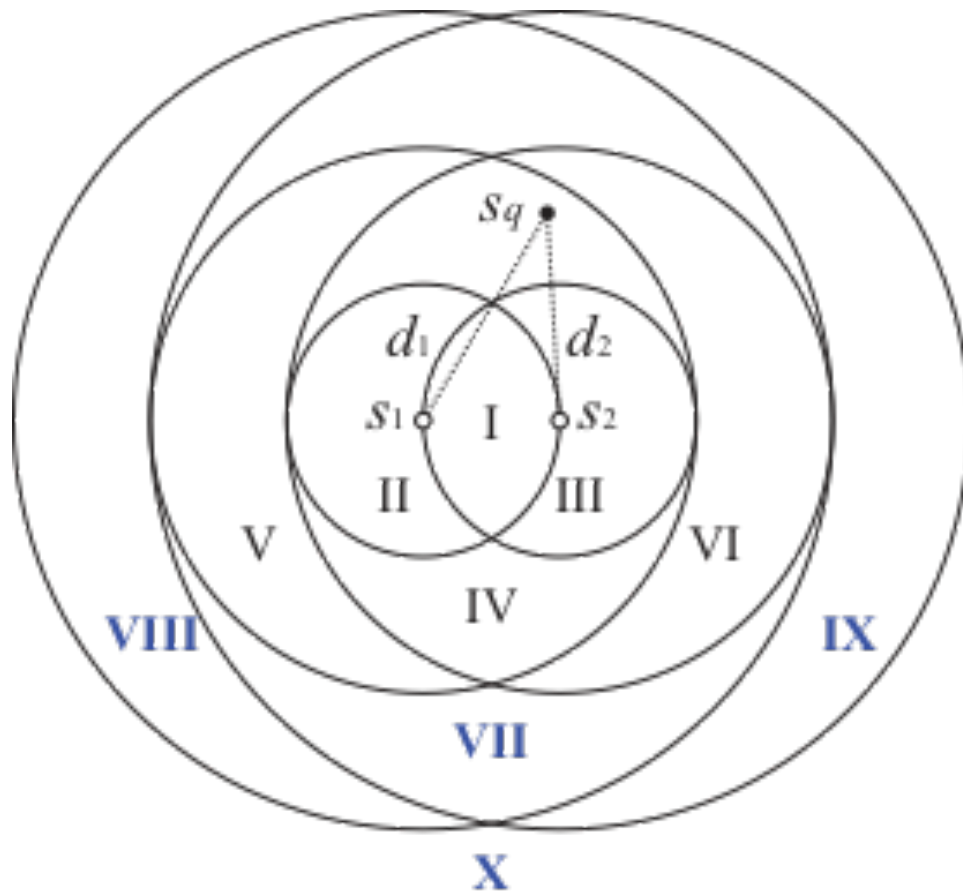
1. Visita a expansão E na qual se encontra o elemento S_q
2. Visita as outras regiões, de acordo com a proximidade ao elemento da query
3. Expansões $E - 1$ e $E + 1$, Expansões $E - 2$ e $E + 2$



Onion-tree

Algoritmos de consulta estendidos – KNN

1. Visita a expansão E na qual se encontra o elemento S_q
2. Visita as outras regiões, de acordo com a proximidade ao elemento da query
3. Expansões $E - 1$ e $E + 1$, Expansões $E - 2$ e $E + 2$



Expansão 1
Regiões IV, VI, V

Expansão 0
Regiões I, III, II

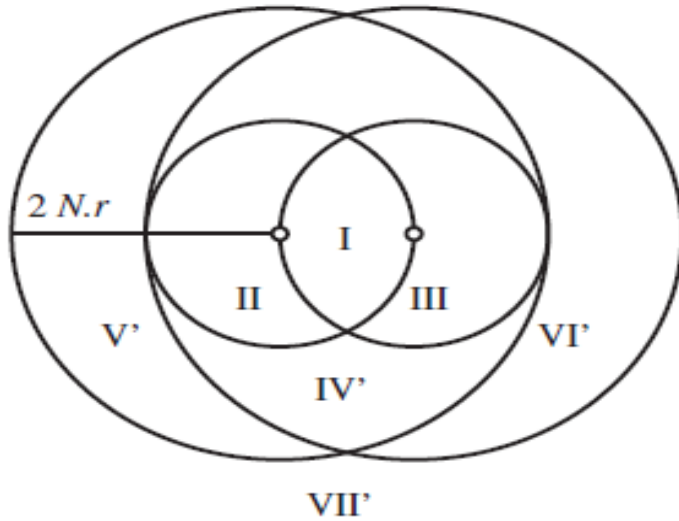
Expansão 2
Regiões VII, IX,
VIII, X

Onion-tree

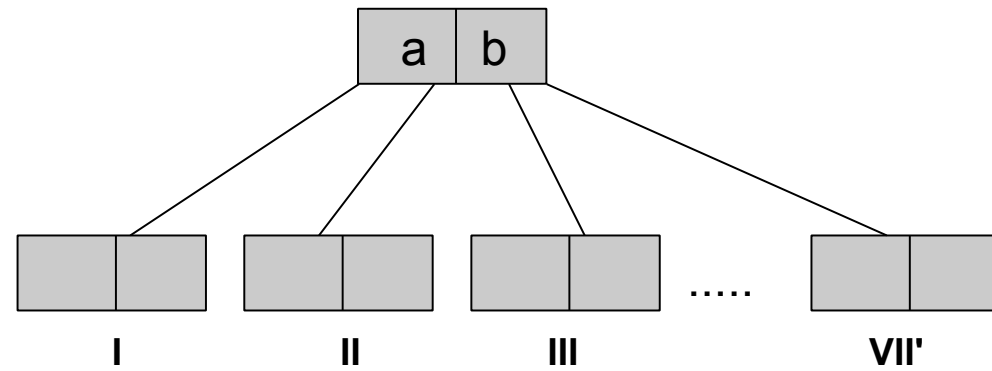
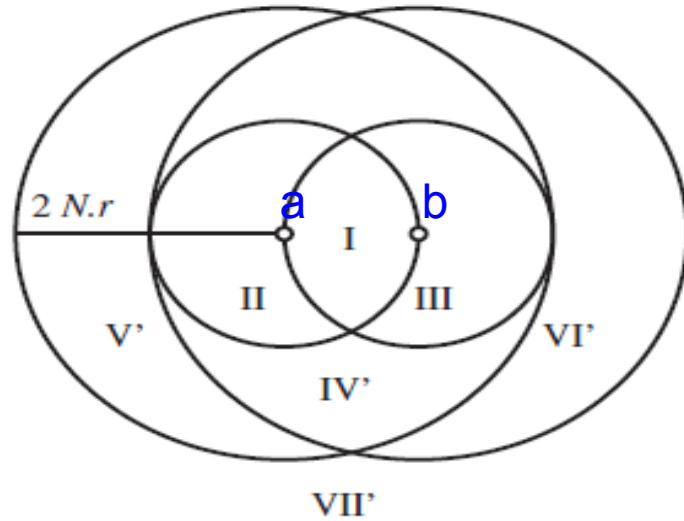
Algoritmo de Inserção

Leva em consideração a distância entre o novo elemento a ser inserido e os elementos representantes do nó

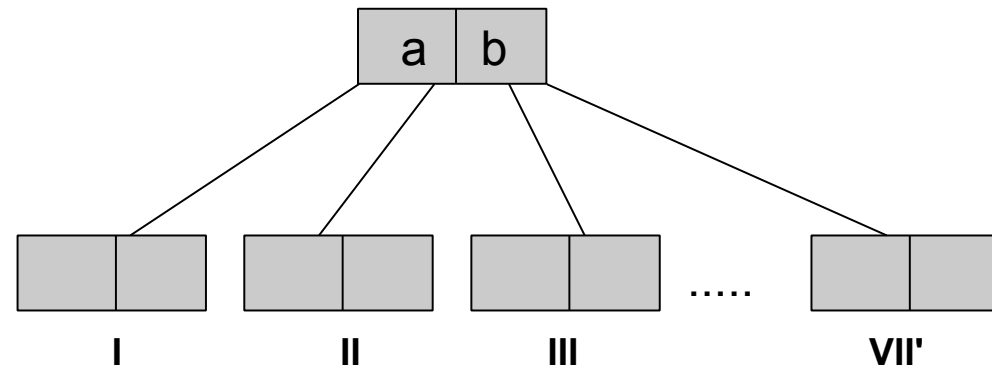
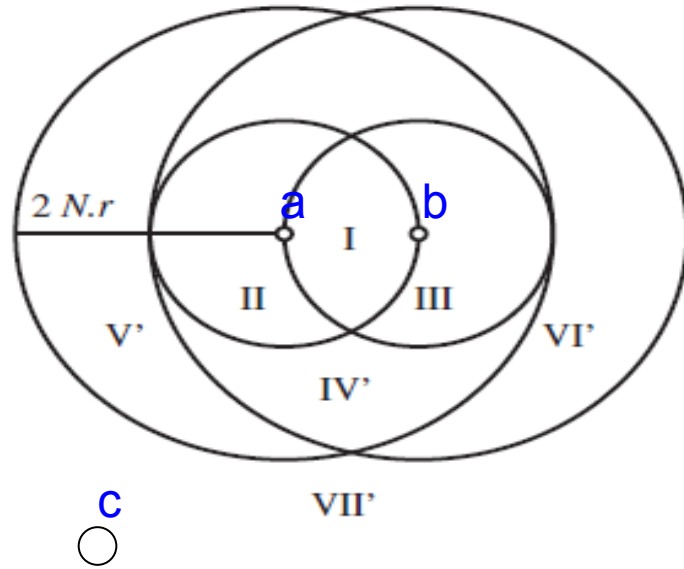
Onion-tree



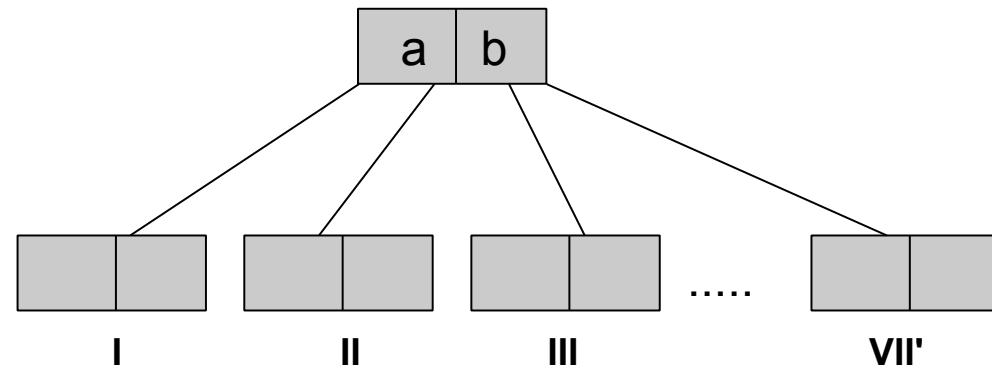
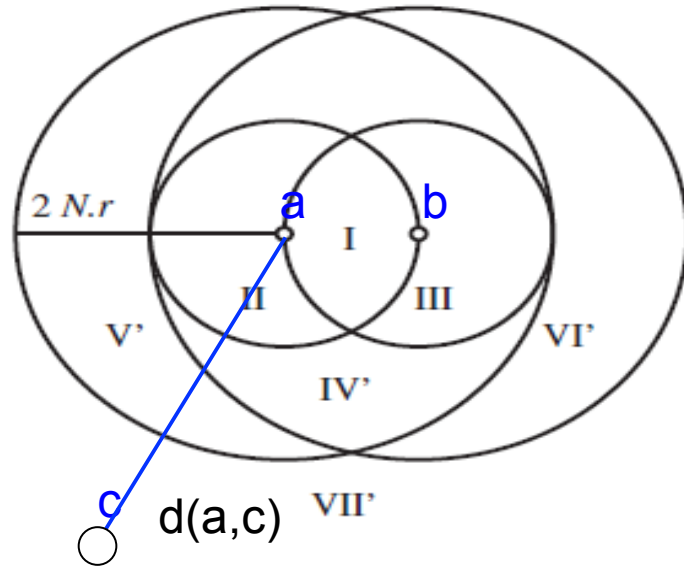
Onion-tree



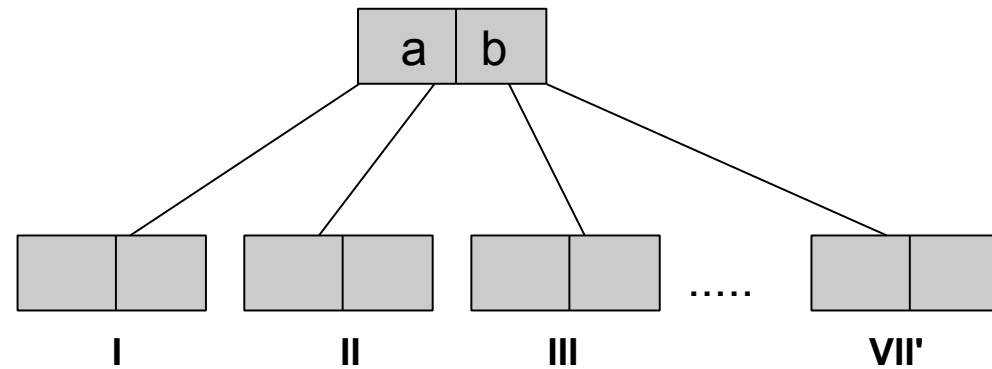
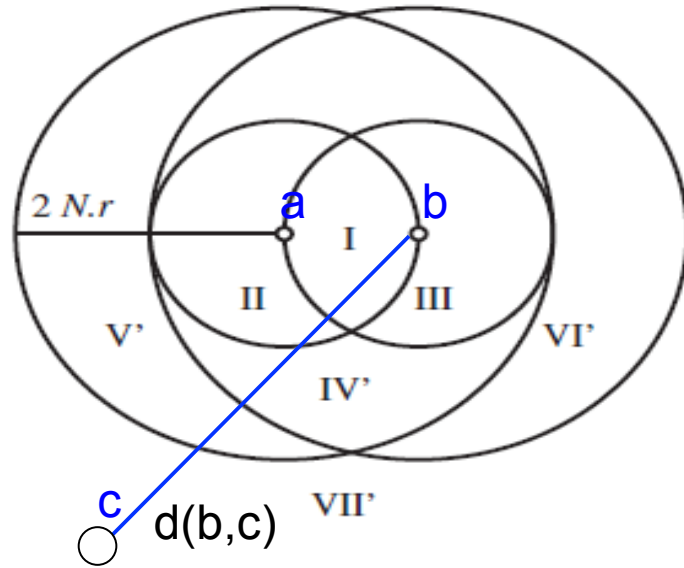
Onion-tree



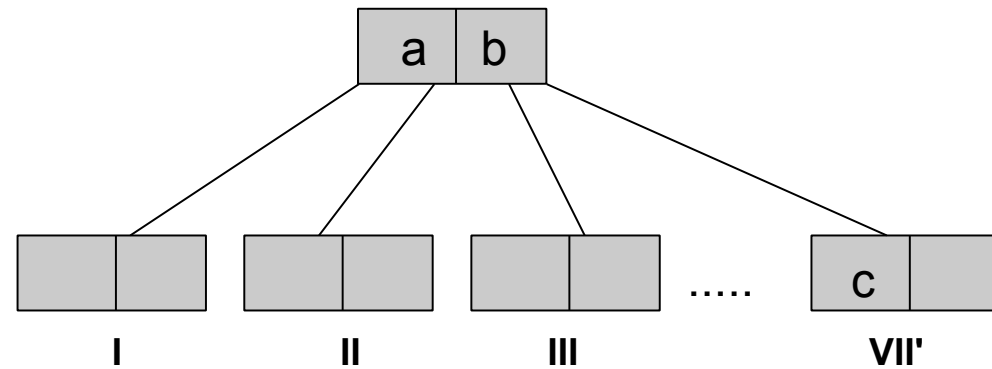
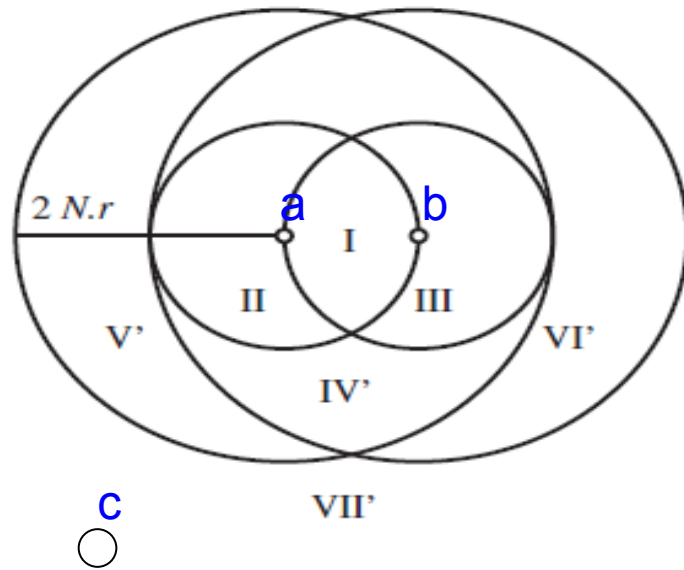
Onion-tree



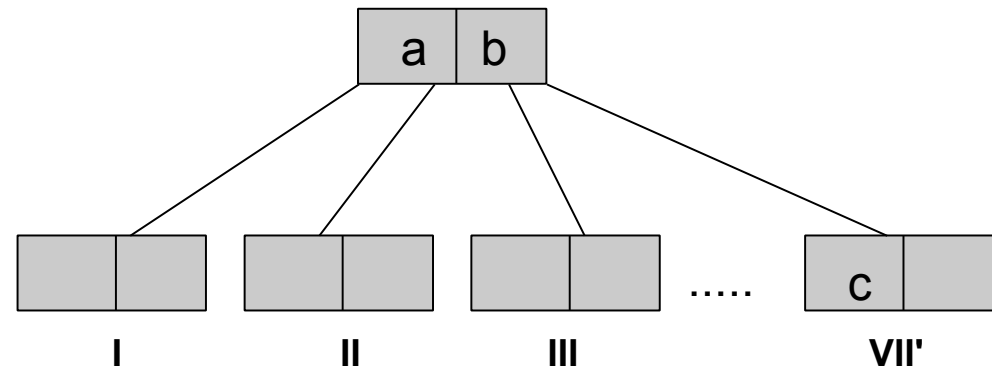
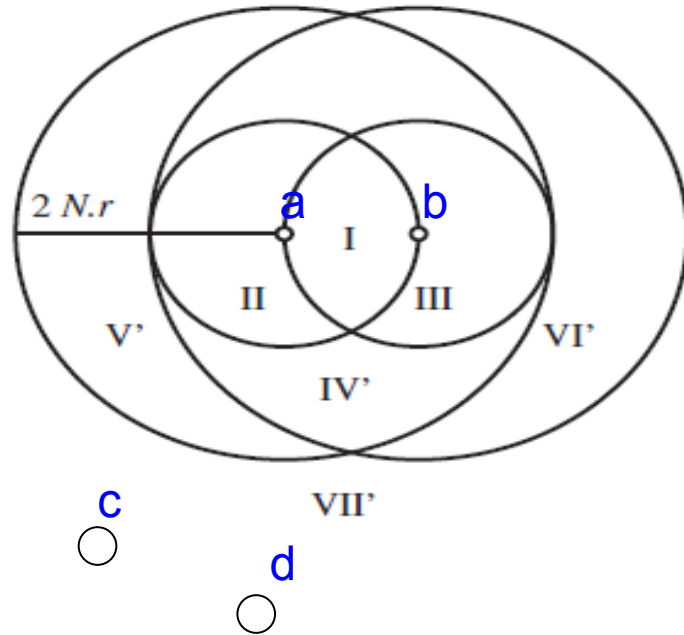
Onion-tree



Onion-tree



Onion-tree

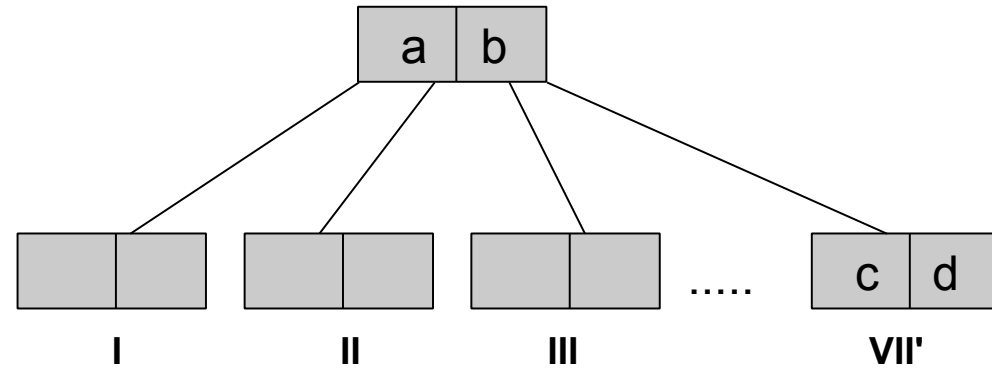
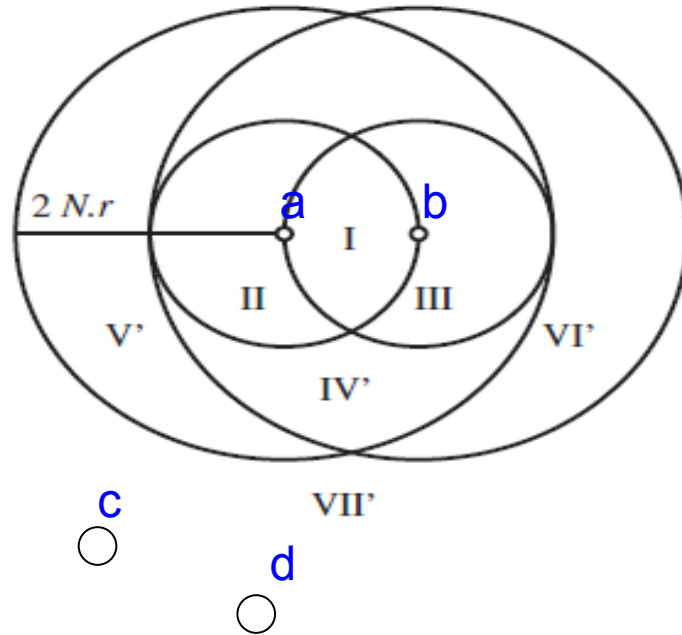


Algoritmo de Inserção

F-Onion-tree

1 Expansão

Onion-tree



Onion-tree

