

# Processamento e Otimização de Consultas

Banco de Dados

Profa. Dra. Cristina Dutra de Aguiar Ciferri

# Motivação

- Consulta
  - pode ter sua resposta computada por uma variedade de métodos (geralmente)
- Usuário (programador)
  - sugere uma estratégia para achar a resposta, independentemente de ser a estratégia mais eficiente
- SGBD
  - responsável por transformar a consulta realizada pelo usuário em uma consulta **equivalente mais eficiente**

# Processamento de Consultas

- Característica
  - melhoria da estratégia para processamento de uma consulta
  - não apresenta uma estratégia **ótima**
  - porém apresenta uma estratégia **eficiente**
- Modelo relacional
  - facilita a otimização da consulta
  - permite que uma consulta seja expressa inteiramente em uma linguagem de consulta relacional (i.e., SQL) sem o uso de uma linguagem hospedeira

# Processamento de Consultas

- Técnica utilizada para processar, otimizar e executar consultas de alto nível
- Objetivo:
  - produzir uma **estratégia de consulta** para recuperar o resultado da mesma

**plano para:** executar a consulta  
acessar os dados  
armazenar resultados intermediários

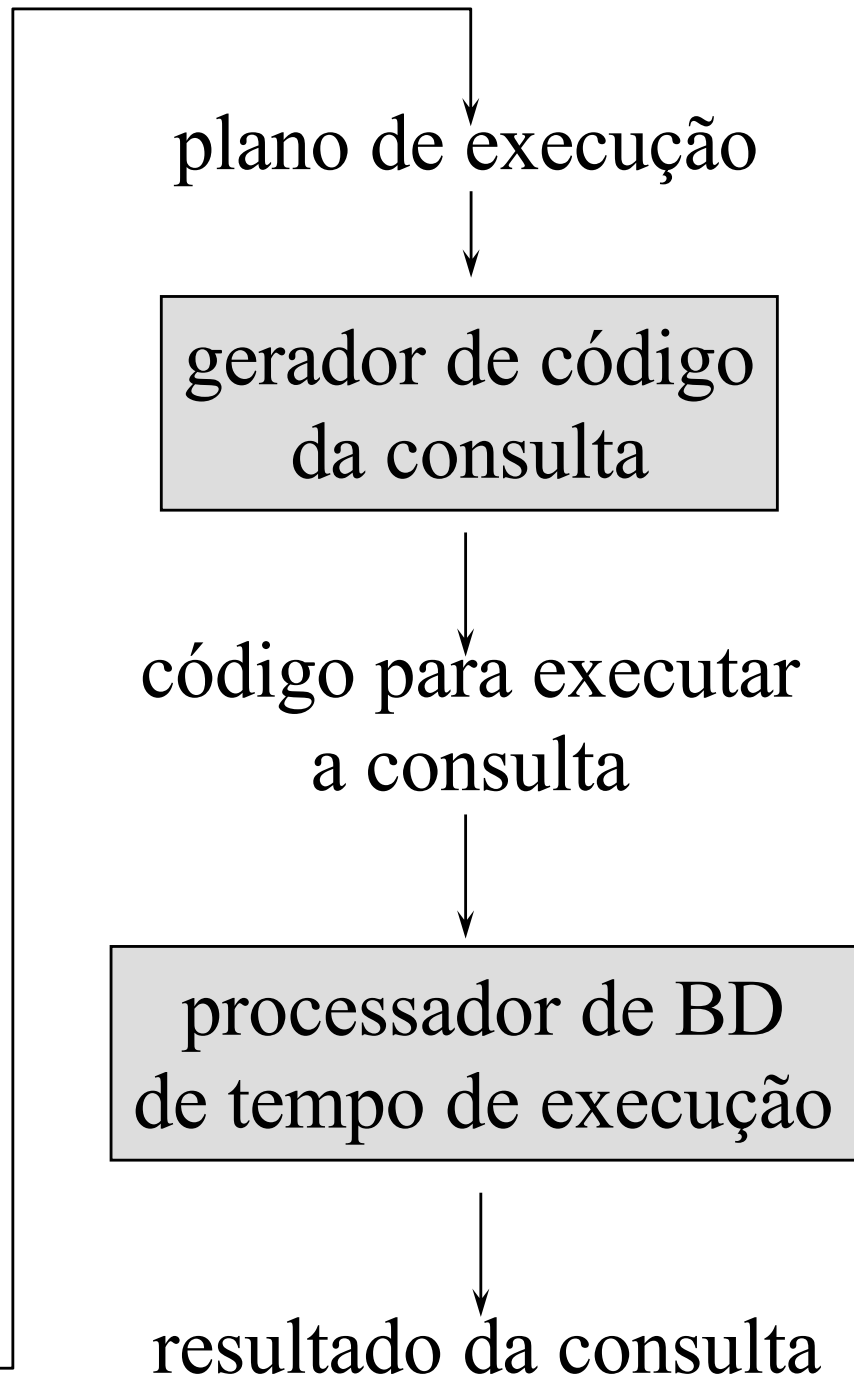
SQL consulta em uma linguagem de alto nível

análise léxica, sintática e semântica

álgebra forma intermediária de consulta

otimizador de consulta

baseado em regras	baseado em custo
-------------------	------------------



# Árvore de Consulta

- Estrutura de árvore que corresponde a uma expressão da álgebra relacional
- Representação
  - nós folhas
    - ◆ **relações** de entrada para a consulta
  - nós internos
    - ◆ **operações** da álgebra relacional
- Indica uma ordem específica das operações durante a execução de uma consulta

# Árvore de Consulta

- Execução
  - efetuar uma operação do nó interno sempre que os seus operandos estiverem disponíveis
  - substituir o nó interno pela relação que resulta da execução da operação
  - final da execução:
    - ◆ o nó raiz é executado
    - ◆ a relação resultado para a consulta é produzida

# Exemplo

- Dado o seguinte comando SQL

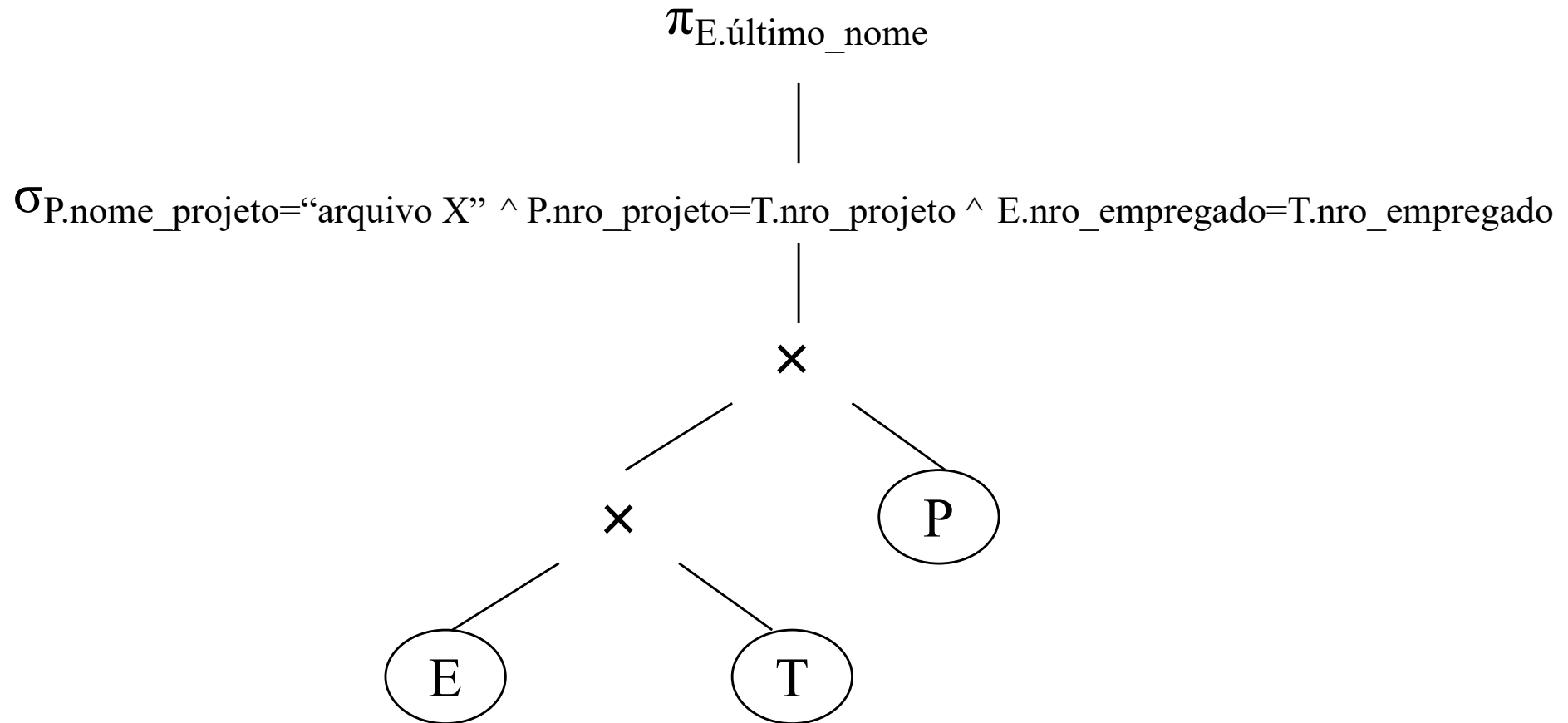
```
SELECT E.último_nome  
FROM empregado E, trabalha T, projeto P  
WHERE P.nome_projeto = "arquivo X"  
      AND P.nro_projeto = T.nro_projeto  
      AND E.nro_empregado = T.nro_empregado
```

- Construa a árvore de consulta canônica



# Árvore de Consulta Canônica

árvore de consulta 1

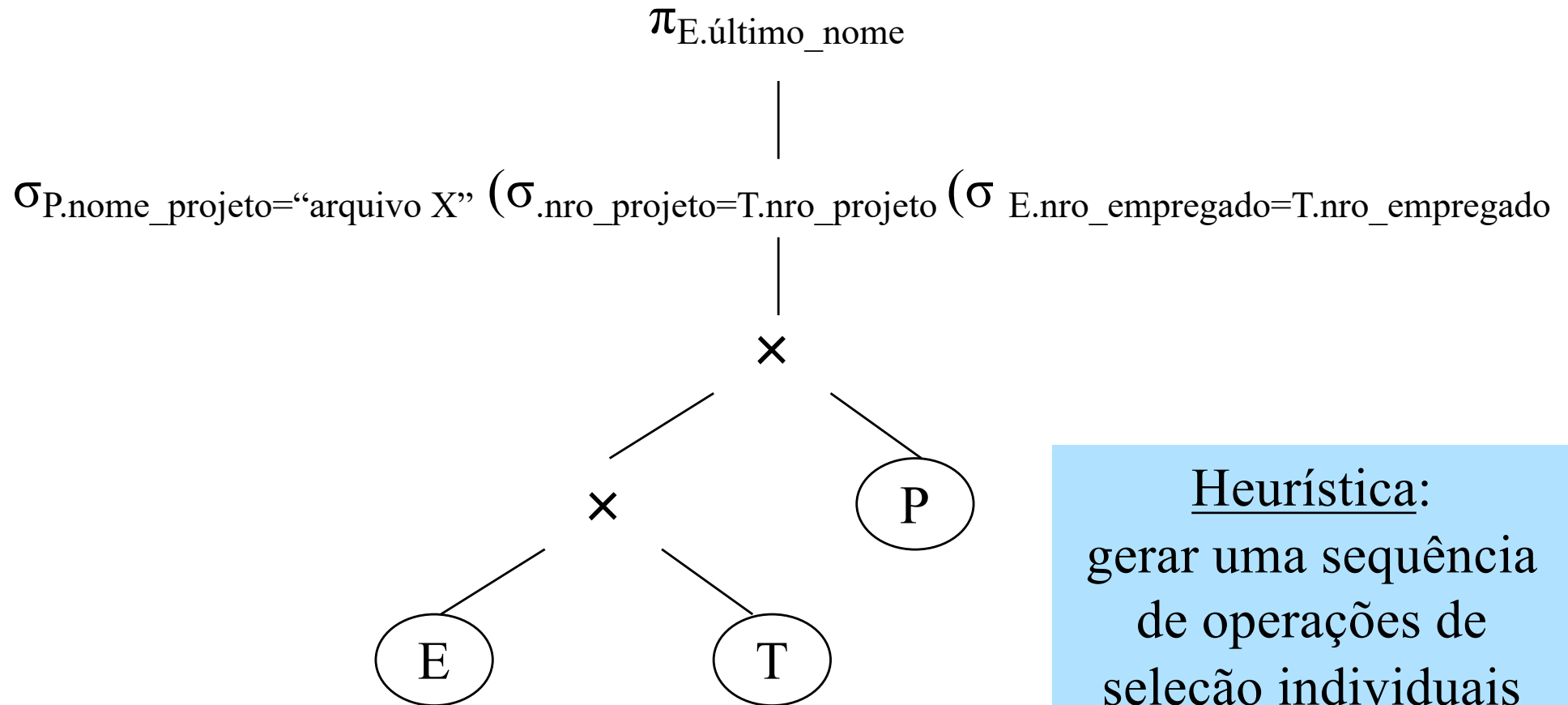


# Conversão da Árvore de Consulta

- Observações
  - o produto cartesiano das relações empregado, trabalha e projeto produz uma grande relação, a qual provavelmente precisará ser armazenada em disco
  - acesso a disco
    - ◆ para ler as relações empregado, trabalha e projeto
    - ◆ para ler e escrever resultados intermediários
  - objetivo
    - ◆ reduzir o tamanho dos resultados intermediários

# Conversão da Árvore de Consulta

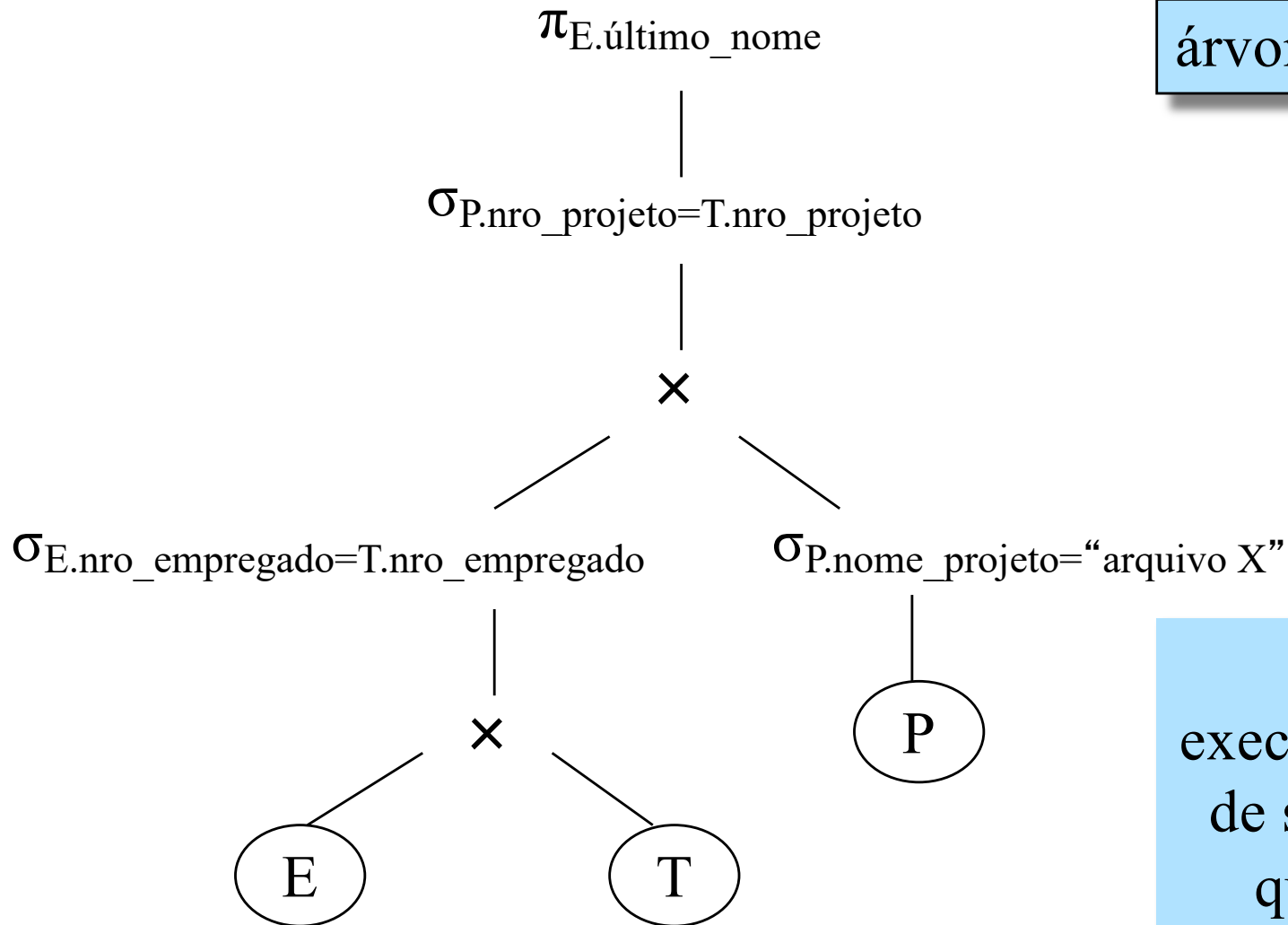
árvore de consulta 2



Heurística:  
gerar uma sequência  
de operações de  
seleção individuais

*algoritmo passo 1*

# Conversão da Árvore de Consulta



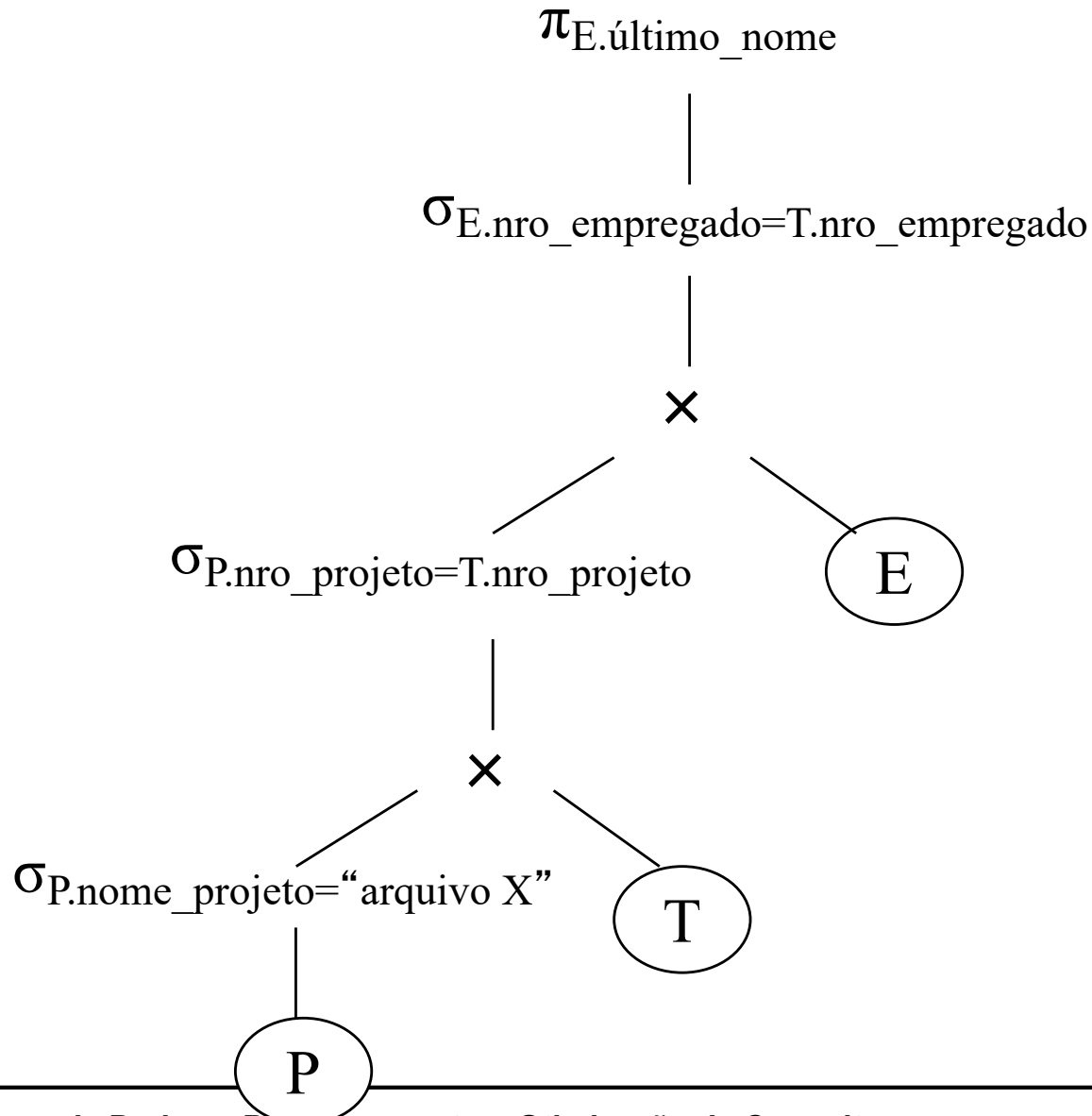
árvore de consulta 3

Heurística:  
executar as operações  
de seleção tão cedo  
quanto possível

*algoritmo passo 2*

# Conversão da Árvore de Consulta

árvore de consulta 4

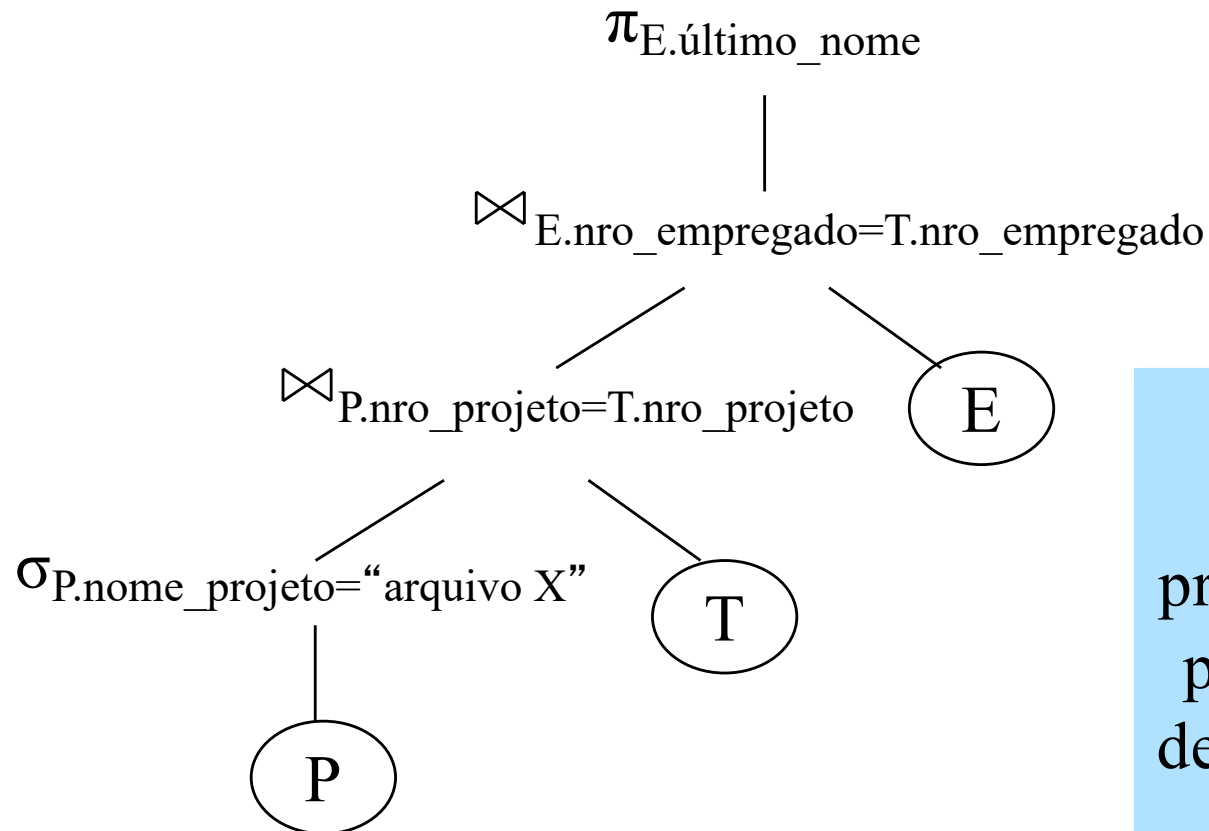


Heurística:  
diminuir os tamanhos  
das relações a serem  
utilizadas no produto  
cartesiano

*algoritmo passo 3*

# Conversão da Árvore de Consulta

árvore de consulta 5



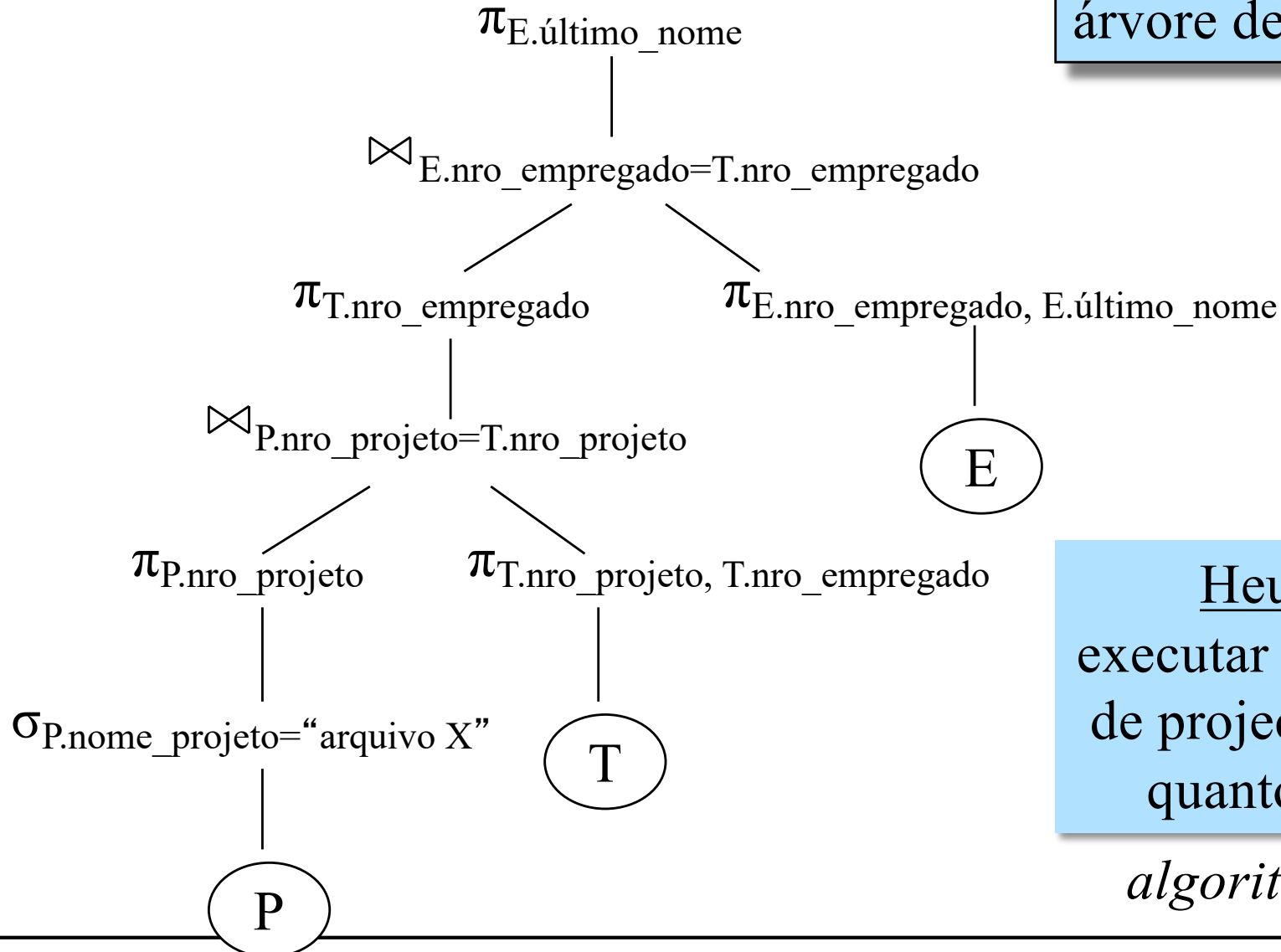
## Heurística:

substituir operações de produto cartesiano seguidas pelos respectivos critérios de seleção por operações de junção

*algoritmo passo 4*

# Conversão da Árvore de Consulta

árvore de consulta 6



Heurística:  
executar as operações  
de projeção tão cedo  
quanto possível

*algoritmo passo 5*

# Heurísticas Básicas

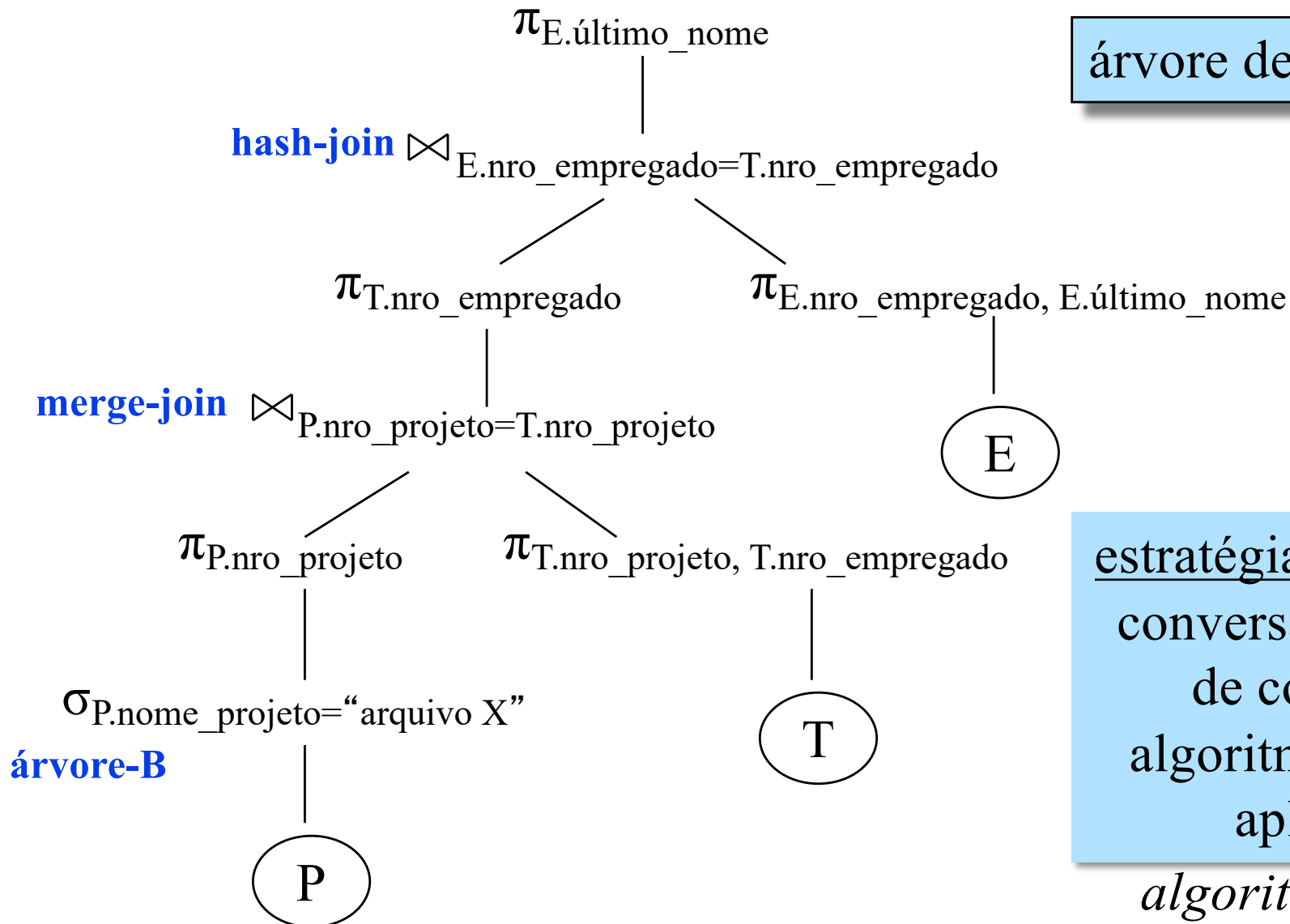
- Aplicar primeiro as operações que reduzem o tamanho dos resultados intermediários
  - operações de seleção
    - ◆ reduzem o número de tuplas
  - operações de projeção
    - ◆ reduzem o número de atributos
- Aplicar primeiro as operações de seleção e de junção mais restritivas
  - reordenar os nós folha da árvore de consulta
  - evitar a operação de produto cartesiano
  - ajustar o restante da árvore de forma apropriada



A geração de expressões é apenas parte do processo de otimização de consultas. Cada operação na expressão pode ser implementada com **diferentes algoritmos**. Uma estratégia de consulta define exatamente que algoritmo é utilizado para cada operação e como a execução das operações é coordenada.

# Exemplo

árvore de consulta 7



estratégia de consulta:

conversão da árvore  
de consulta +  
algoritmos a serem  
aplicados

*algoritmo passo 6*