Applied Faculty

Database 3

Query Processing II

معالجة الاستعلامات ٢

Dr. Ramez Alkhatib

معالجة الاستعلامات

- مراحل معالجة الاستعلامات
- ترجمة استعلامات SQL إلى الجبر العلائقي
- تمثيل الاستعلامات على شكل أشجار
- استخدام الإرشادات في تحسين الاستعلامات
- قواعد التحويل العامة لعمليات الجبر العلائقي

استخدام الإرشادات في تحسين الاستعلامات
Heuristics for query optimization

- تقنيات تحسين الاستعلامات التي تطبق قواعد إرشادية (heuristic rules) لتعديل التمثيل الداخلي للاستعلام (عادة شجرة استعلام) لتحسين الأداء.
- تطبيق قواعد التحسين الإرشادية يؤدي إلى تحويل شجرة استعلام إلى شجرة استعلام مكافئة لها ، لكن أكثر فعالية في التنفيذ، و تعطي نفس نتيجة الاستعلام الأصلي

- بشكل عام، العديد من تعابير الجبر العلائقي المختلفة – و بالتالي العديد من أشجار الاستعلام المختلفة – يمكن أن تكون متكافئة.

استخدام الإرشادات في تحسين الاستعلامات
Heuristics for query optimization

- إحدى أهم القواعد الإرشادية :
- تطبيق عمليات القسر σ و الإسقاط π قبل عملية الربط $\bowtie$ أو العمليات الثنائية الأخرى
  - لأن حجم الملف الناتج عن عملية ثنائية مثل الربط يتناسب مع حجوم ملفات الدخل
  - و عمليات القسر σ و الإسقاط π تخفض حجم الملف، لذا يجب تنفيذها قبل الربط أو أي عملية ثنائية

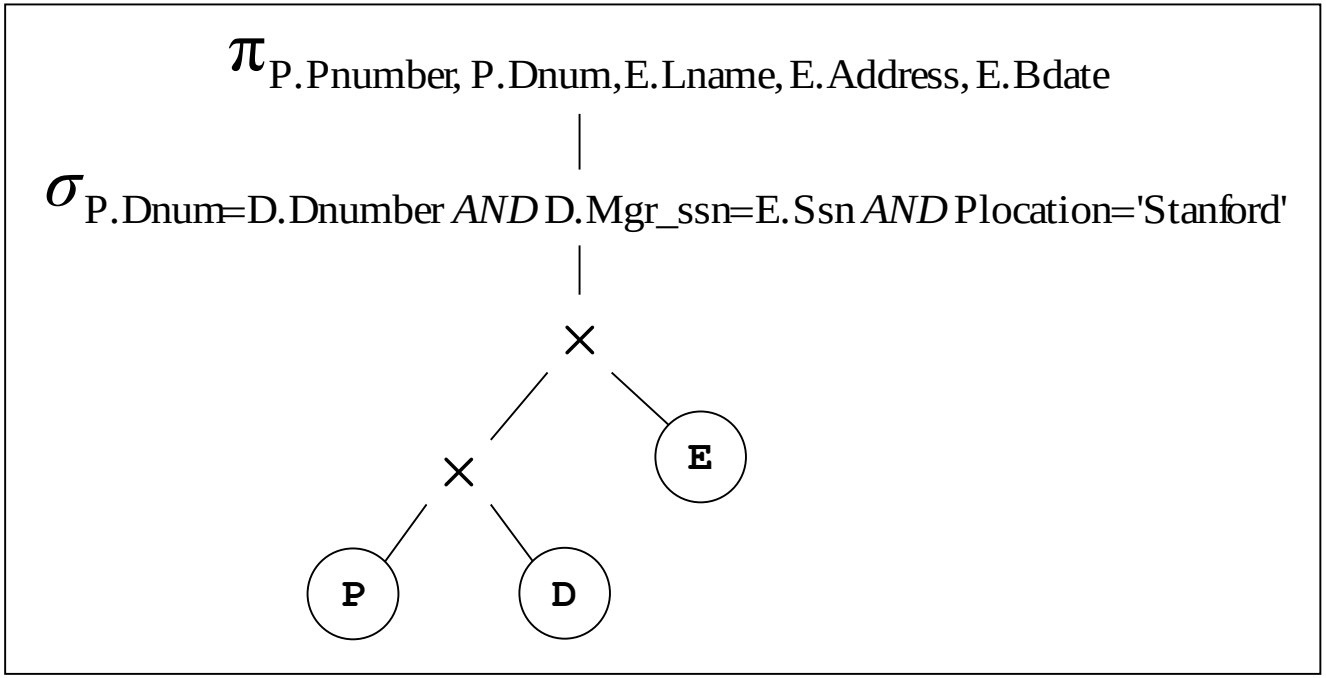
استخدام الإرشادات في تحسين الاستعلامات
Heuristics for query optimization

- يقوم محلل الاستعلام (query parser) بتوليد شجرة استعلام مبدئية قياسية لتوافق استعلام SQL بدون القيام بأي تحسين

- في هذه الشجرة المبدئية :
  - أولاً نطبق الجداء الديكارتي للعلاقات في عبارة FROM
  - ثم نطبق شروط القسر و الربط في عبارة WHERE
  - أخيراً نطبق الإسقاط على الواصفات في عبارة SELECT

استخدام الإرشادات في تحسين الاستعلامات
Heuristics for query optimization

```
SELECT Pnumber, Dnum, Lname, Address, Bdate
  FROM Project, Department, Employee
 WHERE Plocation = 'Stanford'
   AND Dnum = Dnumber
   AND Mgr_ssn = Ssn
```



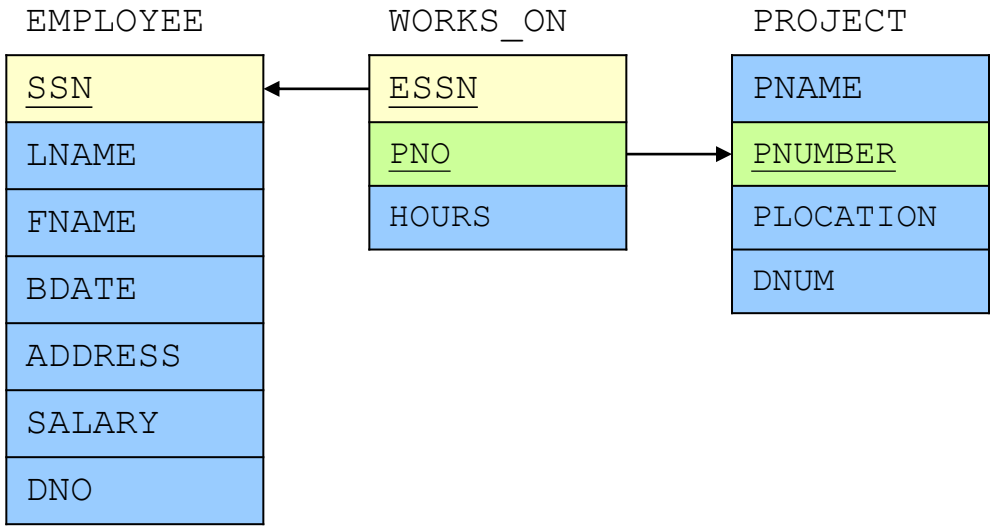
استخدام الإرشادات في تحسين الاستعلامات
Heuristics for query optimization

- تنفيذ هذه الشجرة المبدئية غير فعال كثيراً بسبب عمليات الجداء الديكارتي
- لنفرض أن حجوم السجلات و عدد الأنساق في العلاقات Project و Department و Employee هي:

	Record Size	Tuples
Project	100 byte	100
Department	50 byte	20
Employee	150 byte	5000

- نتيجة الجداء الديكارتي ستحوي 10ملايين نسق حجم كل منها 300 بايت !!

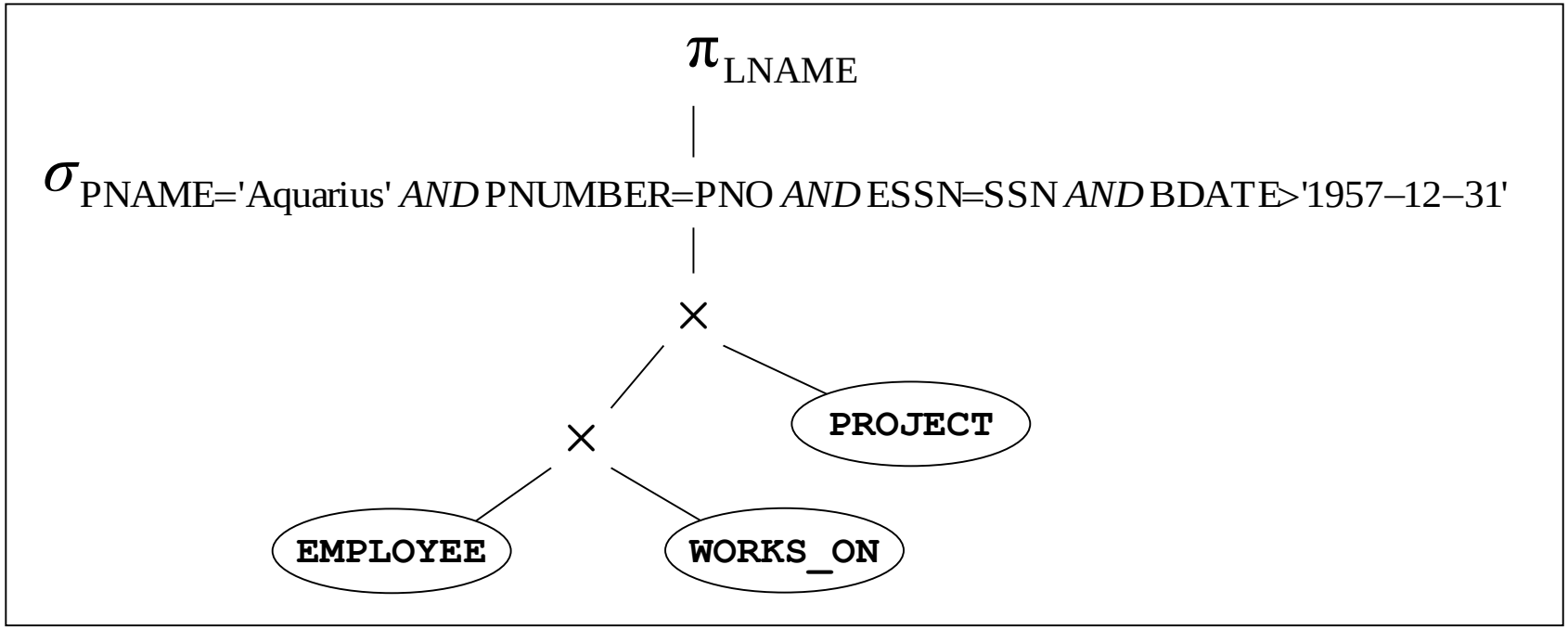
مثال



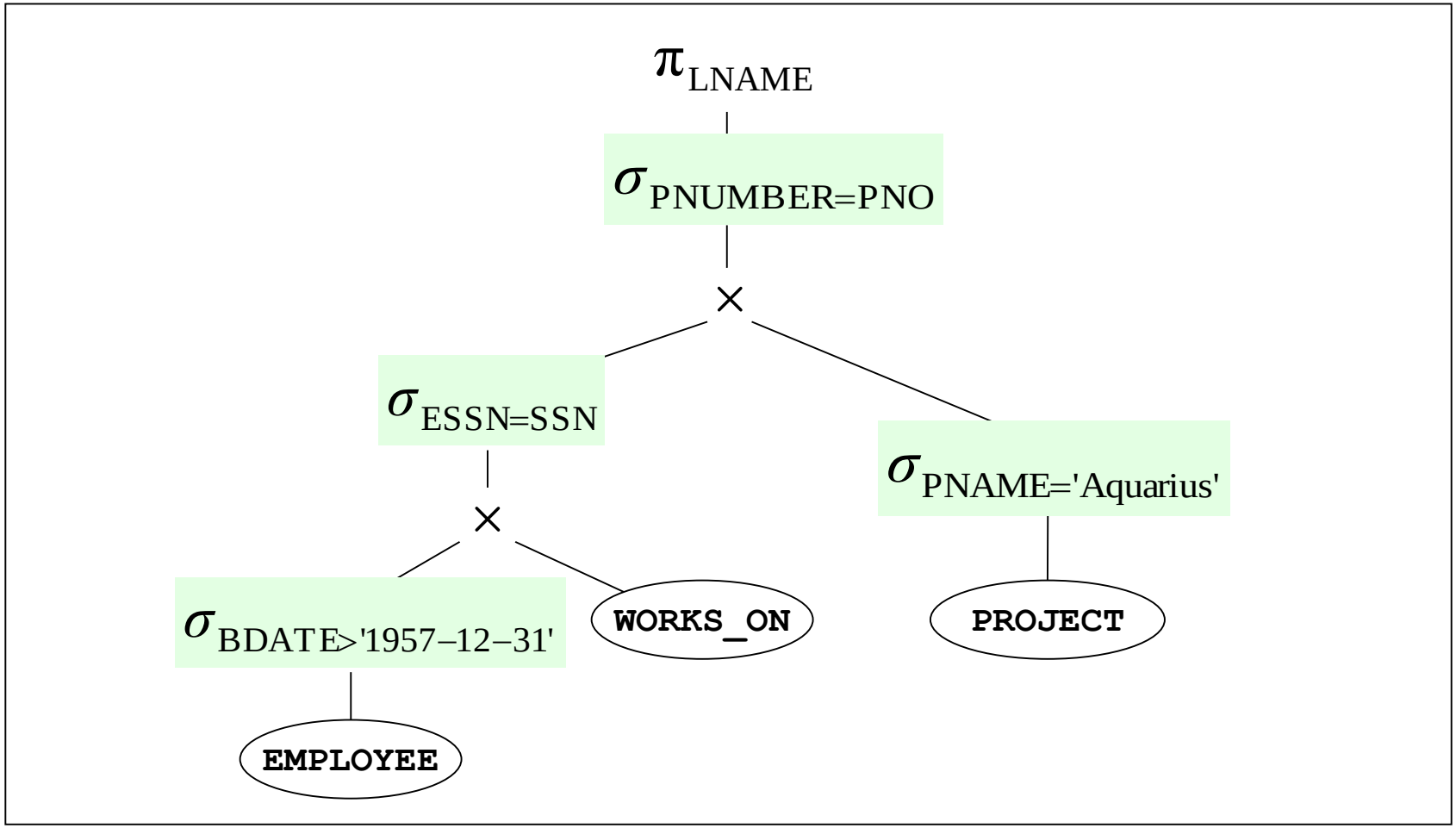

```
SELECT LNAME
  FROM EMPLOYEE, WORKS_ON, PROJECT
 WHERE PNAME='AQUARIUS' AND PNUMBER=PNO
   AND ESSN=SSN AND BDATE > '1957-12-31';
```

• الشجرة المبدئية
Initial query tree

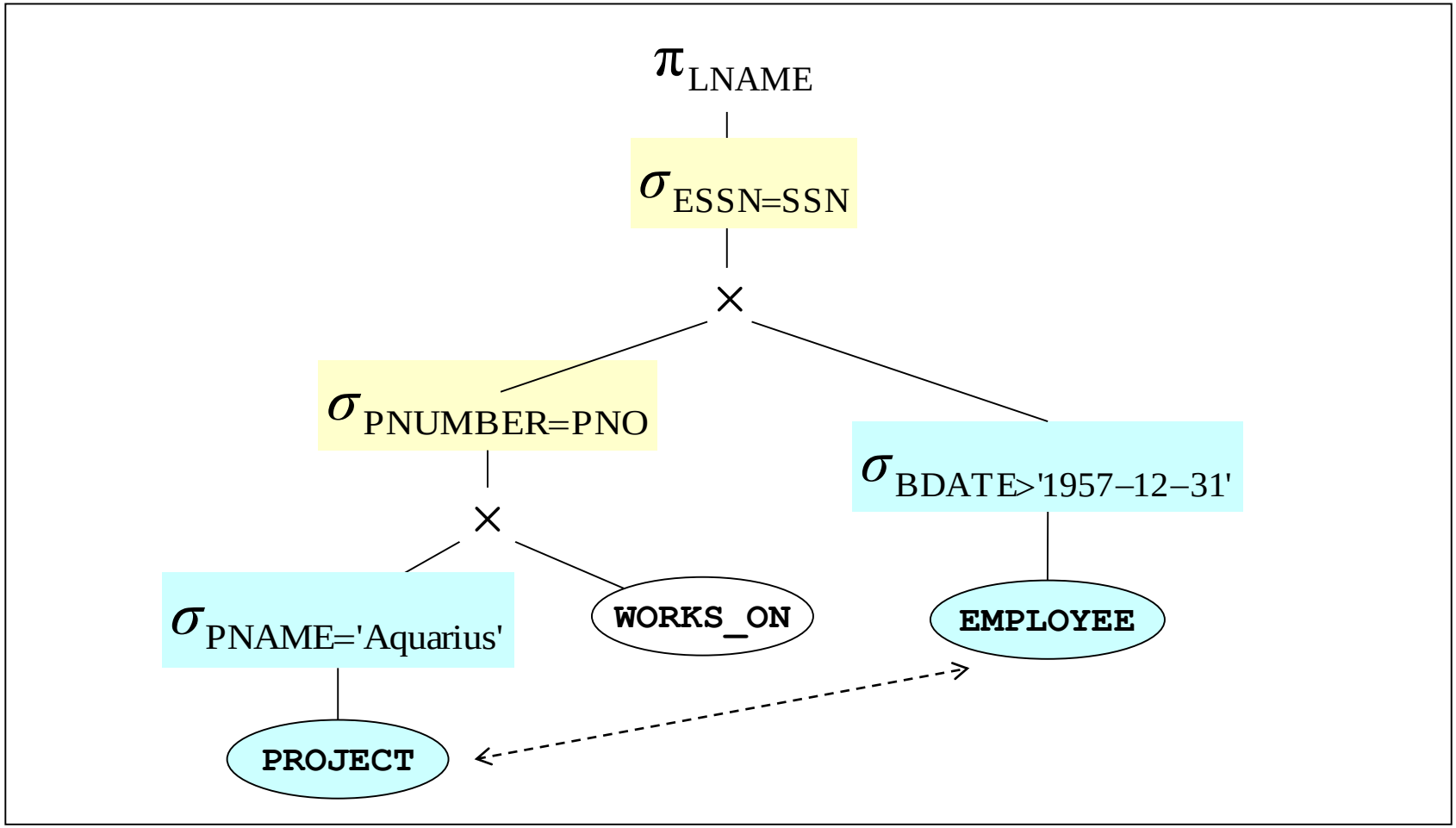
```
SELECT LNAME
  FROM EMPLOYEE, WORKS_ON, PROJECT
 WHERE PNAME='AQUARIUS' AND PNUMBER=PNO
   AND ESSN=SSN AND BDATE > '1957-12-31';
```



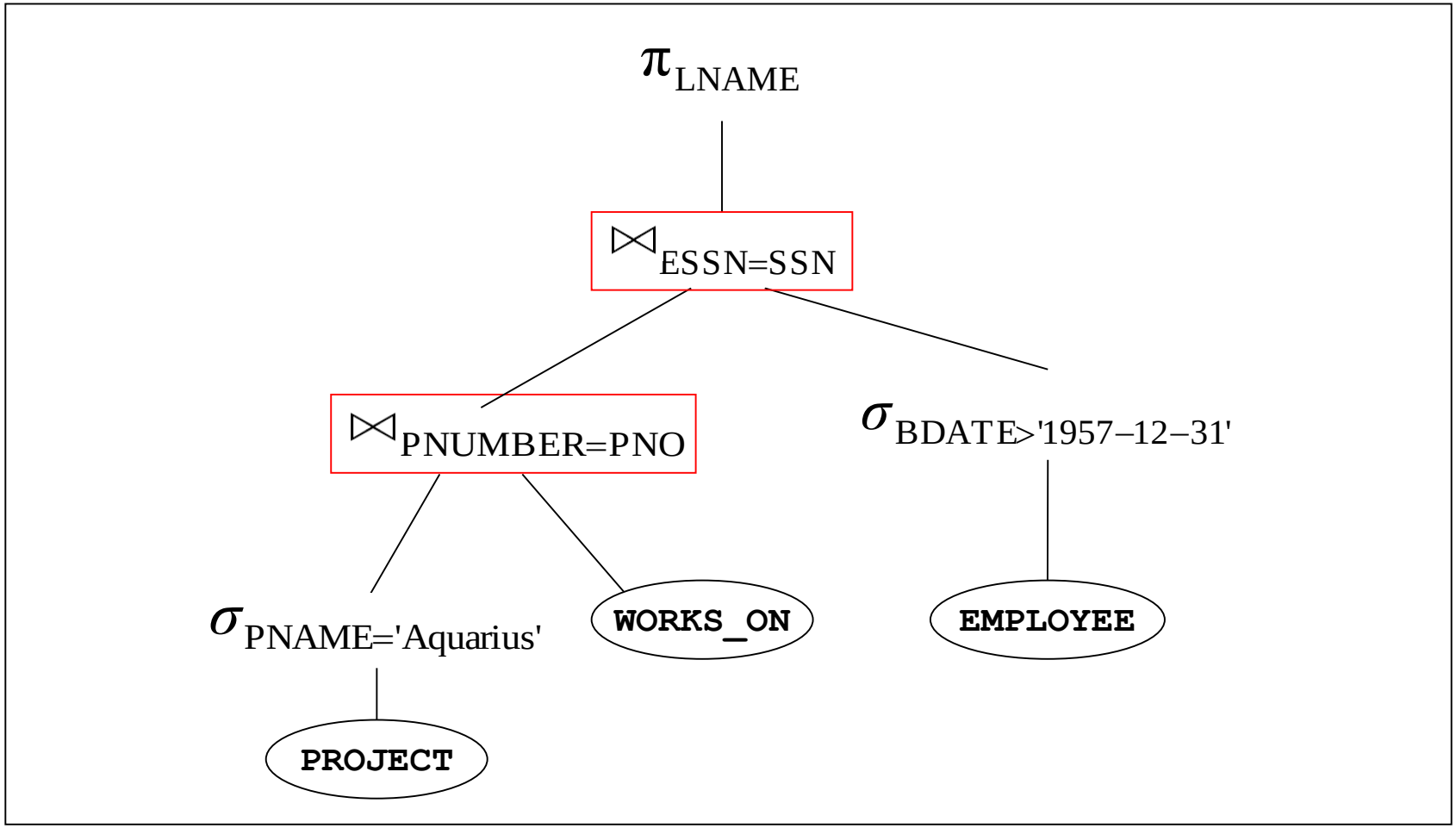
١ ، ٢ تفكيك عمليات القسر و تحريكها هبوطاً في الشجرة



٣ تطبيق عمليات القسر الأكثر تقييداً أولاً

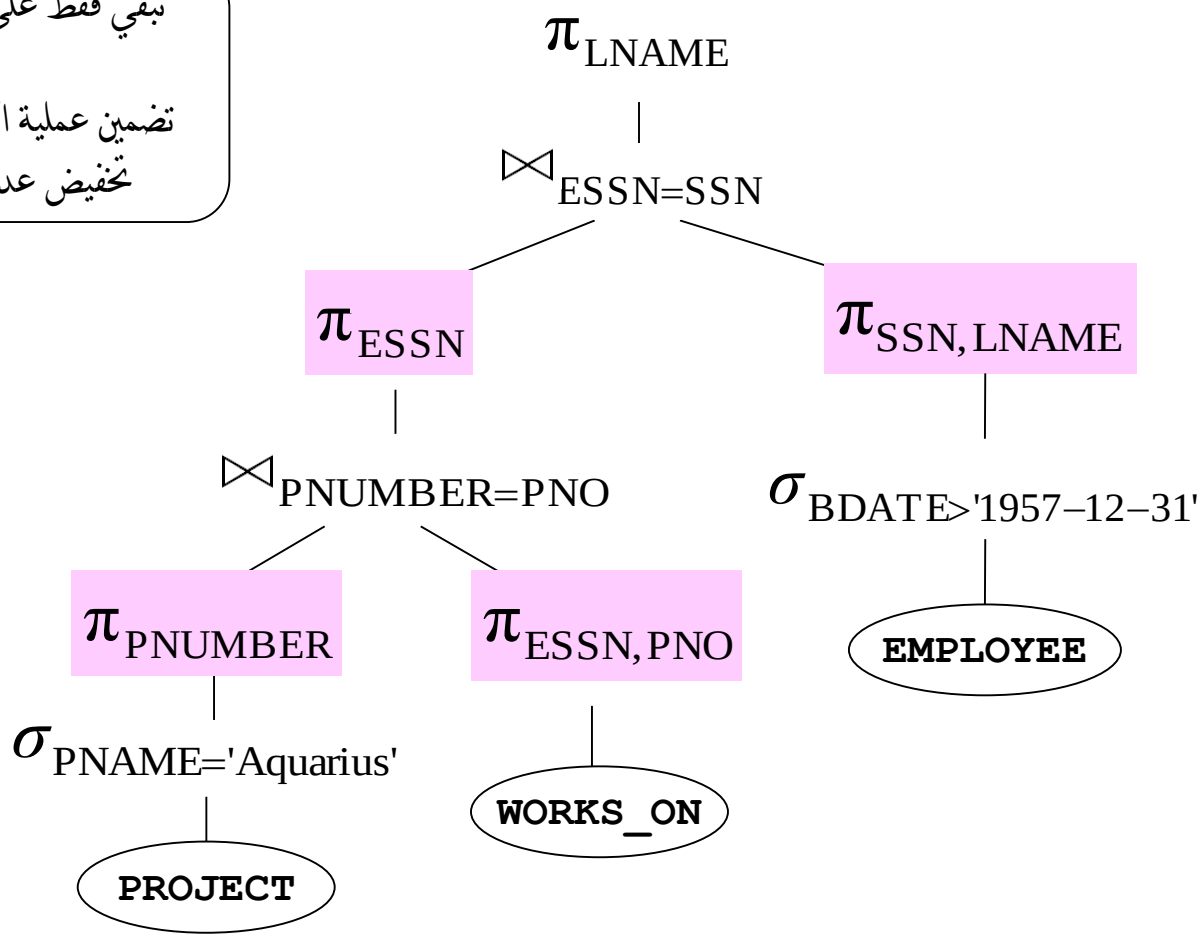


٤ تحويل عملية الجداء الديكارتي المتبوعة بعملية قسر إلى عملية ربط



٥ تحريك عمليات الإسقاط هبوطاً في الشجرة

نبقي فقط على الواصفات التي نحتاجها في العمليات اللاحقة
تضمين عملية الإسقاط أبكر ما يمكن في الشجرة←تخفيض عدد الواصفات

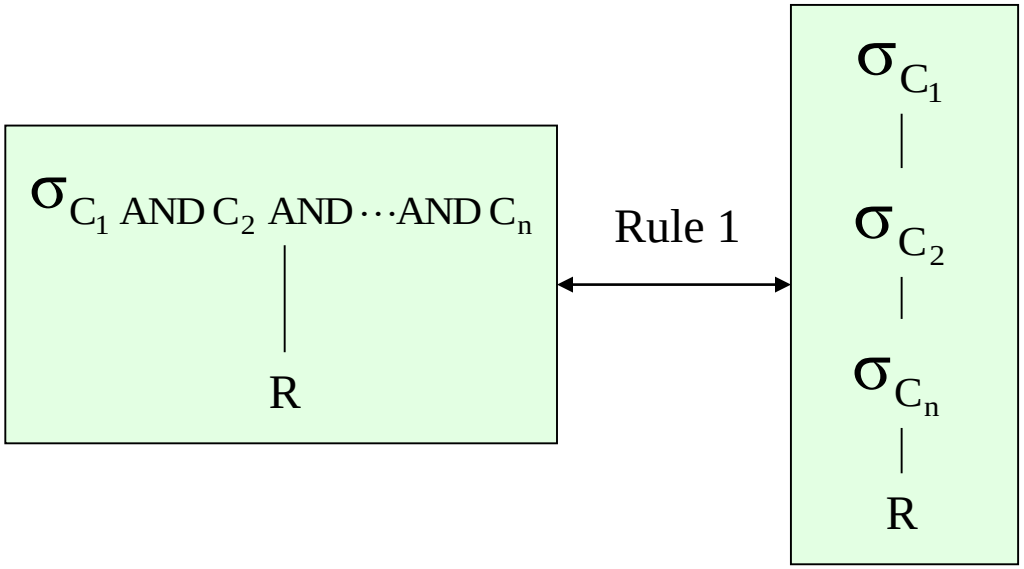


قواعد التحويل العامة لعمليات الجبر العلائقي

General transformation rules for relational algebra operations

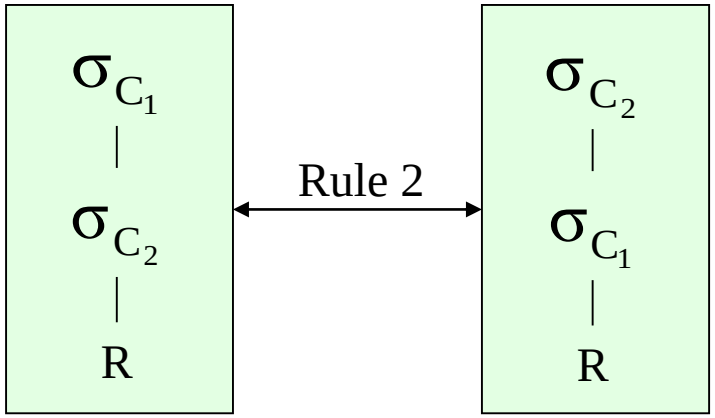
١. القسر σ تتابعي

$$\sigma_{C_1 \text{ AND } C_2 \text{ AND } \cdots \text{AND } C_n}(R) \equiv \sigma_{C_1}(\sigma_{C_2}(\cdots(\sigma_{C_n}(R))\cdots))$$



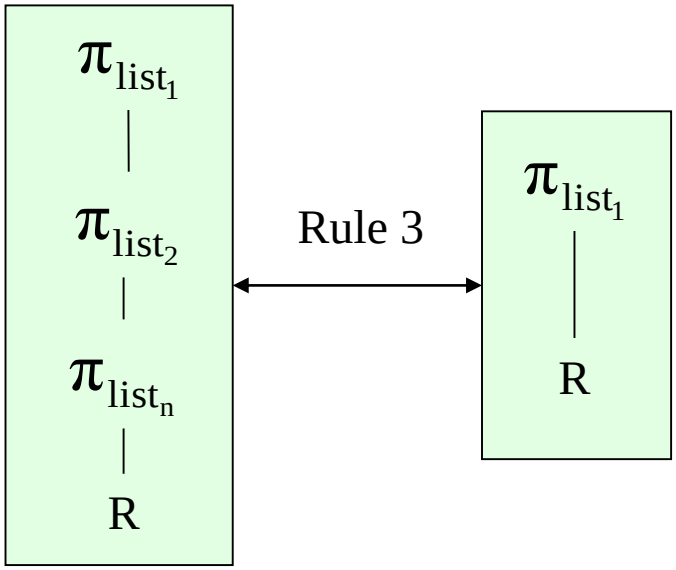
٢. القسر σ تبديلي

$$\sigma_{c_1}(\sigma_{c_2}(R)) \equiv \sigma_{c_2}(\sigma_{c_1}(R))$$



٣. الإسقاط π تتابعي

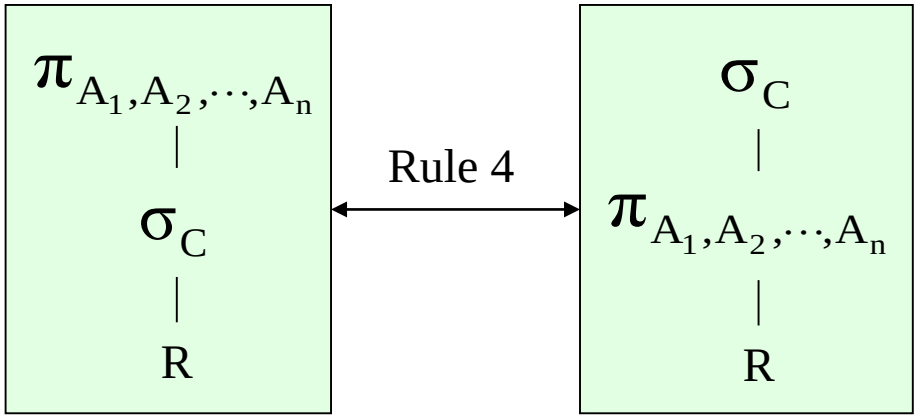
$$\pi_{list_1}(\pi_{list_2}(\cdots(\pi_{list_n}(R))\cdots)) \equiv \pi_{list_1}(R)$$



٤. القسر σ تبديلي مع الإسقاط π

- إذا كان شرط القسر C يتضمن فقط واصفات الإسقاط

$$\pi_{A_1,A_2,\cdots,A_n}(\sigma_C(R)) \equiv \sigma_C(\pi_{A_1,A_2,\cdots,A_n}(R))$$

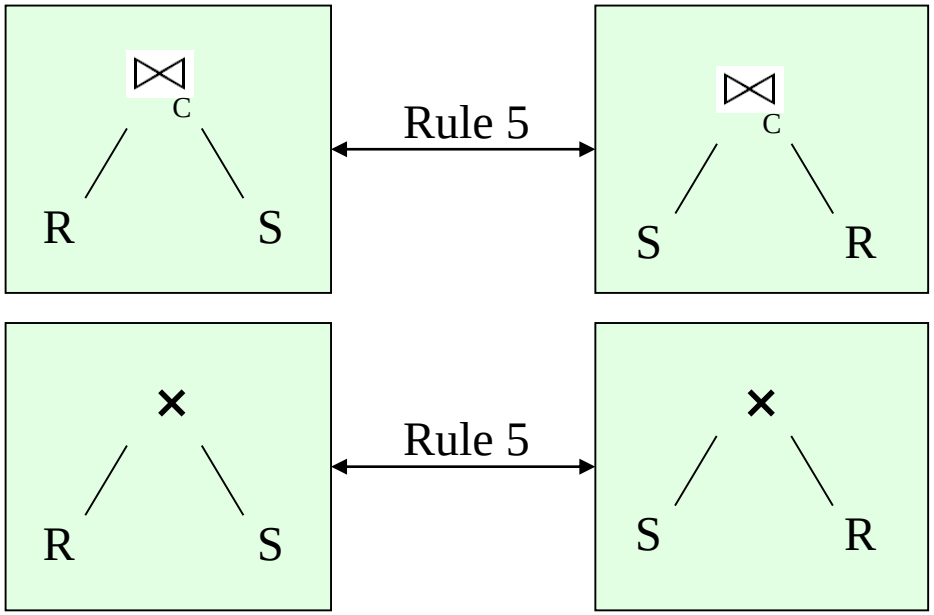


قواعد التحويل العامة لعمليات الجبر العلائقي

General transformation rules for relational algebra operations

٥. الربط $\bowtie$ تبديلي و الجداء الديكارتي $\times$ تبديلي

$$R \bowtie_C S \equiv S \bowtie_C R \qquad R \times S \equiv S \times R$$

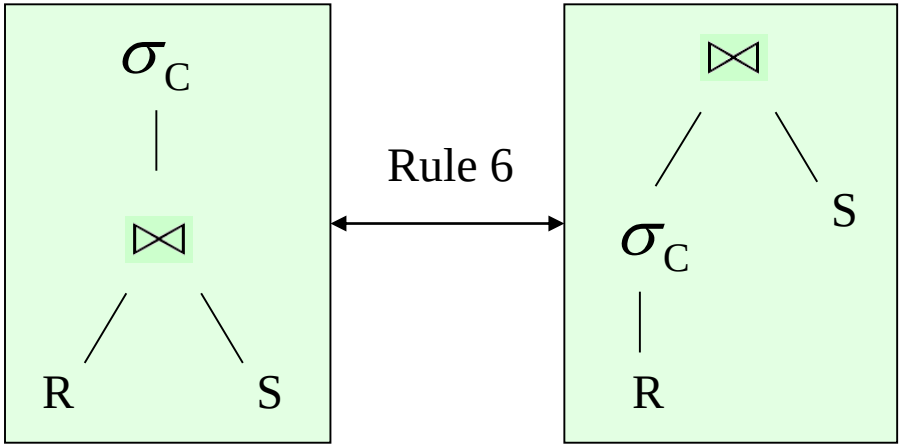


قواعد التحويل العامة لعمليات الجبر العلائقي
General transformation rules for relational algebra operations

٦. القسر σ تبديلي مع الربط ⋈ (أو الجداء الديكارتي ×)

إذا كانت جميع الواصفات في شرط القسر C تتضمن فقط الواصفات لأحد علاقتي الربط (و لتكن R)

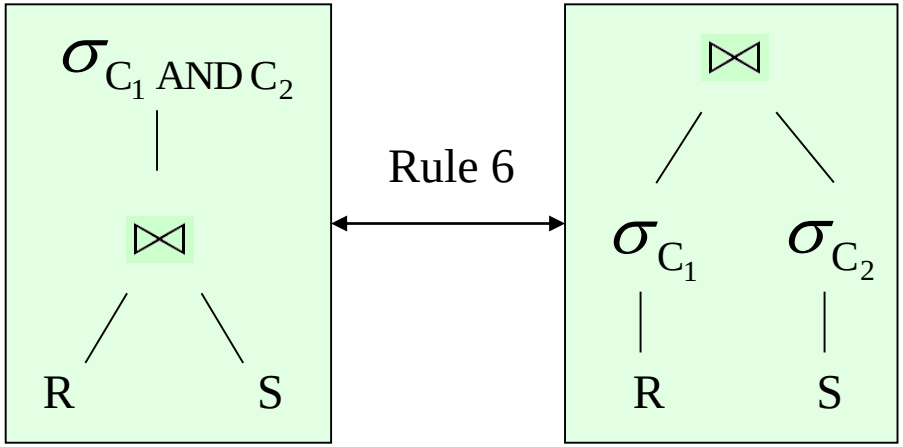
$$\sigma_C(R \bowtie S) \equiv (\sigma_C(R)) \bowtie S$$



٦. القسر σ تبديلي مع الربط ⋈ (أو الجداء الديكارتي ×)

إذا كان شرط القسر C يكتب بالشكل C_1 AND C_2 حيث الشرط C_1 يحوي فقط واصفات R و الشرط C_2 يحوي فقط واصفات S

$$\sigma_{C_1 \text{ AND } C_1}(R \bowtie S) \equiv (\sigma_{C_1}(R)) \bowtie (\sigma_{C_2}(S))$$



قواعد التحويل العامة لعمليات الجبر العلائقي
General transformation rules for relational algebra operations

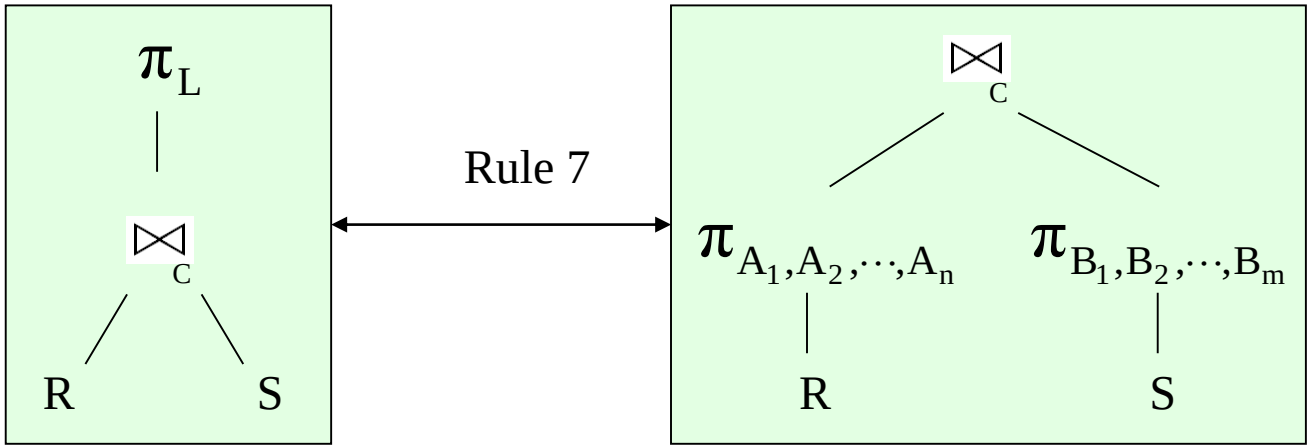
٧. الإسقاط π تبديلي مع الربط $\bowtie$ (أو الجداء الديكارتي $\times$)

إذا كانت قائمة الإسقاط هي $L=\{A_1, A_2, ..., A_n, B_1, B_2, ..., B_m\}$

حيث $A_1, A_2, ..., A_n$ هي واصفات R و $B_1, B_2, ..., B_m$ هي واصفات S

إذا كان شرط الربط C يتضمن فقط الواصفات في L فيمكن تبديل العمليتين

$$\pi_L(R \bowtie_C S) \equiv (\pi_{A_1,A_2,\cdots,A_n}(R)) \bowtie (\pi_{B_1,B_2,\cdots,B_m}(S))$$

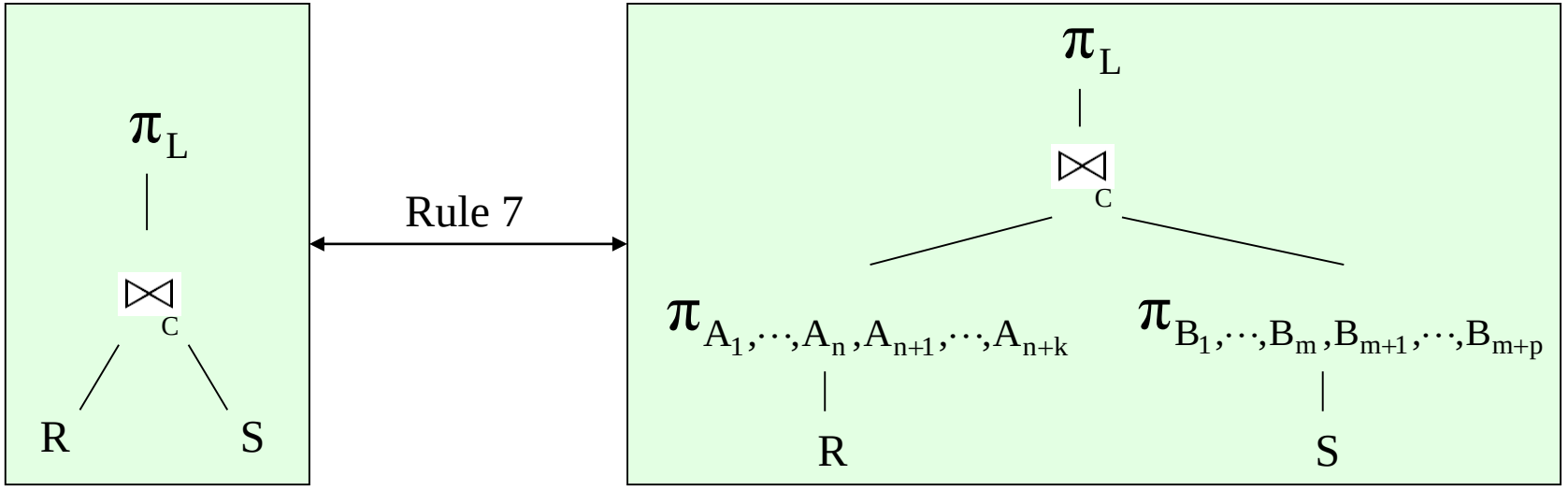


قواعد التحويل العامة لعمليات الجبر العلائقي

General transformation rules for relational algebra operations

- إذا كان شرط الربط C يحوي واصفات إضافية غير موجودة في L، فيجب إضافة هذه الواصفات إلى قائمة الإسقاط ، و نحتاج عملية إسقاط إضافية.
- مثلاً، إذا كانت الواصفات $A_{n+1}, \ldots, A_{n+k}$ من R و الواصفات $B_{m+1}, \ldots, B_{m+p}$ من S موجودة في شرط الربط C و لكن غير موجودة في قائمة الإسقاط L :

$$\pi_L(R \bowtie_C S) \equiv \pi_L((\pi_{A_1,\cdots,A_n,A_{n+1},\cdots,A_{n+k}}(R)) \bowtie_C (\pi_{B_1,\cdots,B_m,B_{m+1},\cdots,B_{m+p}}(S)))$$



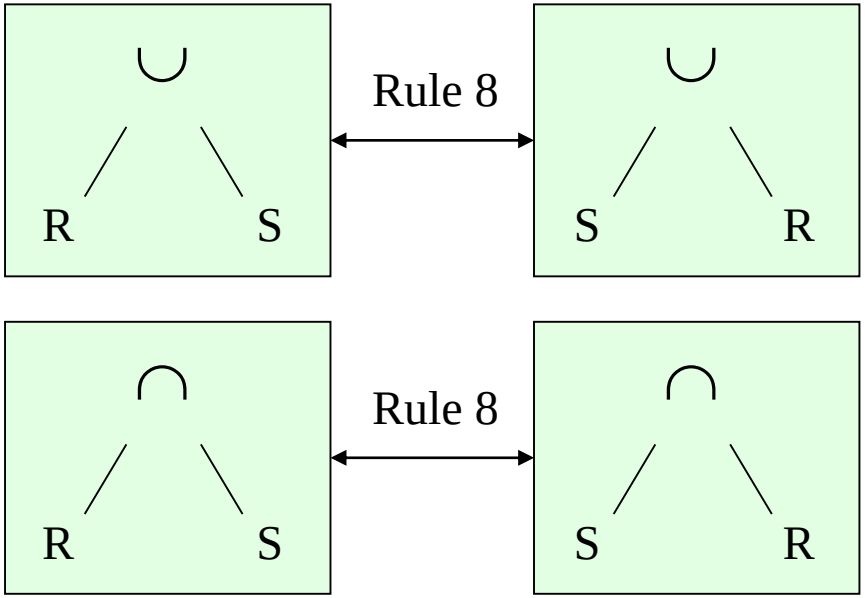
٨. عمليات الاجتماع ∪ و التقاطع ∩ تبديلية

- عملية الفرق - غير تبديلية

$$R \cup S \equiv S \cup R$$

$$R \cap S \equiv S \cap R$$

$$R - S \neq S - R$$



قواعد التحويل العامة لعمليات الجبر العلائقي

General transformation rules for relational algebra operations

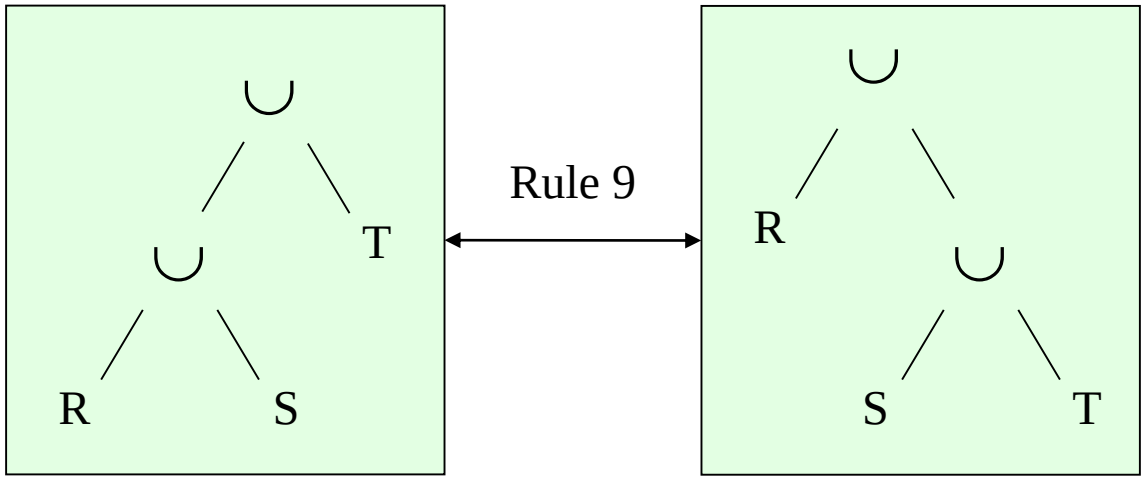
٩. عمليات الربط $\bowtie$ و الجداء الديكارتي $\times$ و الاجتماع $\cup$ و التقاطع $\cap$ تجميعية

$$(R \times S) \times T \equiv R \times (S \times T)$$

$$(R \bowtie S) \bowtie T \equiv R \bowtie (S \bowtie T)$$

$$(R \cup S) \cup T \equiv R \cup (S \cup T)$$

$$(R \cap S) \cap T \equiv R \cap (S \cap T)$$



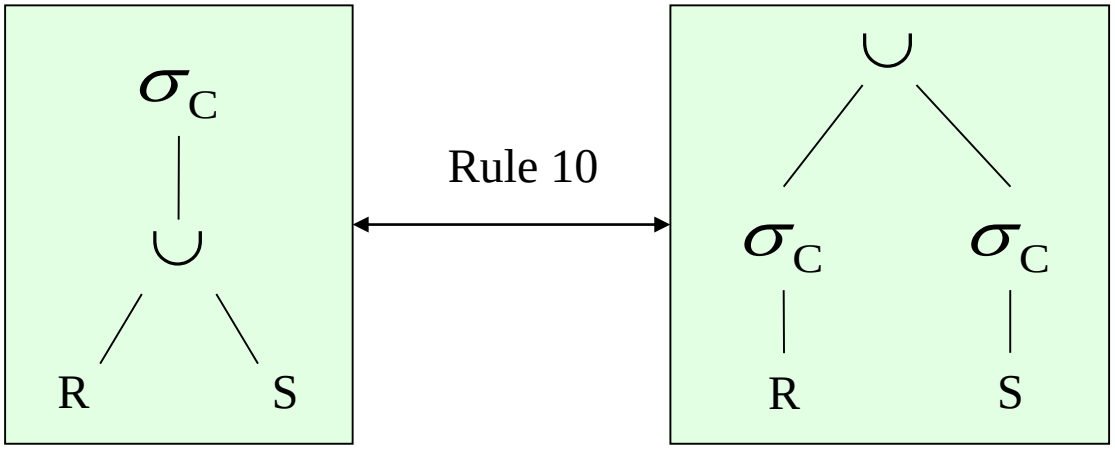
قواعد التحويل العامة لعمليات الجبر العلائقي
General transformation rules for relational algebra operations

١٠. عملية القسر σ توزيعية على عمليات المجموعات (الاجتماع ∪ و التقاطع ∩ و الفرق -)

$$\sigma_C(R \cup S) \equiv \sigma_C(R) \cup \sigma_C(S)$$

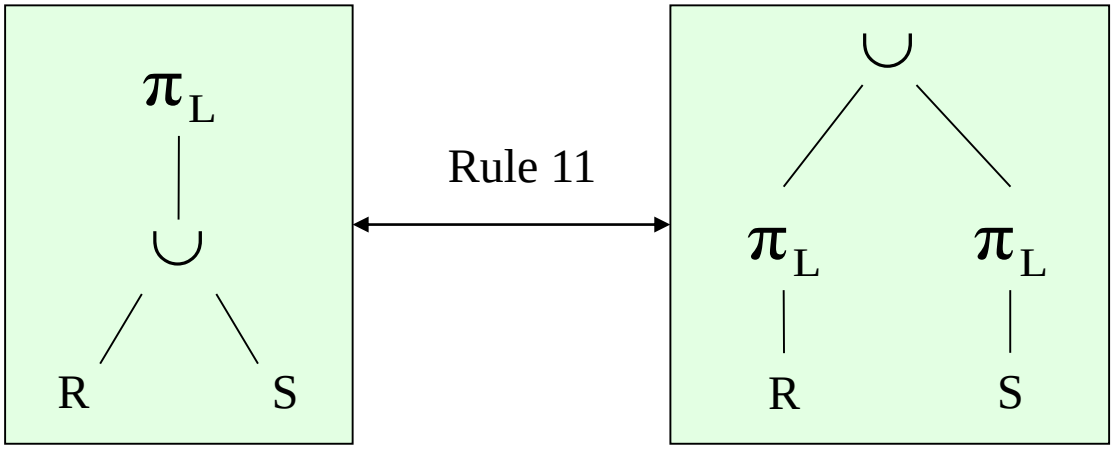
$$\sigma_C(R \cap S) \equiv \sigma_C(R) \cap \sigma_C(S)$$

$$\sigma_C(R - S) \equiv \sigma_C(R) - \sigma_C(S)$$



١١. عملية الإسقاط π توزيعية على عملية الاجتماع $\cup$

$$\pi_L(R \cup S) \equiv \pi_L(R) \cup \pi_L(S)$$

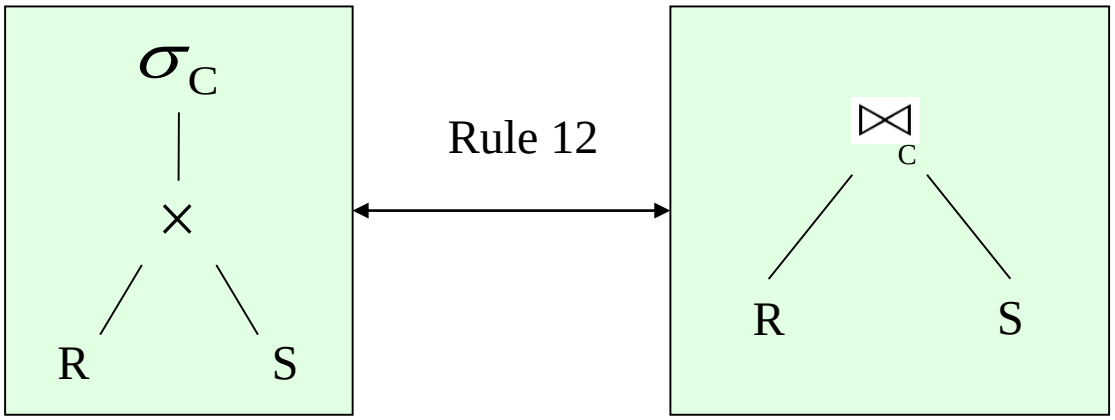


قواعد التحويل العامة لعمليات الجبر العلائقي
General transformation rules for relational algebra operations

٢١. عملية القسر σ التي تتبع عملية جداء ديكارتي × تتحول إلى عملية ربط ⋈

- إذا كان شرط القسر C يوافق شرط ربط

$$\sigma_C(R \times S) \equiv R \bowtie_C S$$



قواعد التحويل العامة لعمليات الجبر العلائقي
General transformation rules for relational algebra operations

- قانونا دومورغان (DeMorgan's rule)

$$\text{NOT}\,(C_1\ \text{AND}\ C_2) \equiv (\text{NOT}\ C_1)\ \text{OR}\ (\text{NOT}\ C_2)$$

$$\text{NOT}\,(C_1\ \text{OR}\ C_2) \equiv (\text{NOT}\ C_1)\ \text{AND}\ (\text{NOT}\ C_2)$$

الخطوط العريضة لخوارزمية تحسين جبري إرشادي
Outline of a heuristic algebraic optimization algorithm

الخطوة الأولى

- باستخدام القاعدة 1 ، فكك عمليات القسر ذات الشروط الموصولة (باستخدام AND) إلى تتابع من عمليات القسر.
- هذا يسمح بدرجة أكبر من الحرية في تحريك عمليات القسر هبوطاً في فروع مختلفة من الشجرة

Rule 1:

$$\sigma_{c1 \text{ and } c2 \text{ and } \ldots \text{ And } cn}(R) \equiv \sigma_{c1}(\sigma_{c2}(\ldots(\sigma_{cn}(R))\ldots))$$

الخطوط العريضة لخوارزمية تحسين جبري إرشادي
Outline of a heuristic algebraic optimization algorithm

الخطوة الثانية

- باستخدام القواعد ٢ و ٤ و ٦ و ١٠ المتعلقة بتبديلية القسر مع العمليات الأخرى، حرك كل عملية قسر هبوطاً في الشجرة بقدر ما هو مسموح بالواصفات الموجودة في شرط القسر

Rule 2: $\sigma_{c1}(\sigma_{c2}(R)) \equiv \sigma_{c2}(\sigma_{c1}(R))$

Rule 4: $\pi_{A1, \ldots, An}(\sigma_c(R) \equiv \sigma_c(\pi_{A1, \ldots, An}(R))$

Rule 6: $\sigma_c(R \bowtie S) \equiv (\sigma_c(R)) \bowtie S$

$\sigma_c(R \bowtie S) \equiv (\sigma_{c1}(R)) \bowtie (\sigma_{c2}(S))$

Rule 10: $\sigma_c(R \ \theta \ S) \equiv \sigma_c(R) \ \theta \ \sigma_c(S)$

الخطوط العريضة لخوارزمية تحسين جبري إرشادي
Outline of a heuristic algebraic optimization algorithm

الخطوة الثالثة

- باستخدام القاعدتين ٥ و ٩ المتعلقة بكون العمليات الثنائية تبديلية و تجميعية، أعد ترتيب العقد الطرفية (الأوراق) للشجرة باستخدام المعايير التالية:
  - أولاً، ضع العلاقات الأوراق ذات عمليات القسر الأكثر تقييداً بحيث يتم تنفيذها أولاً في تمثيل شجرة الاستعلام.
    - (العمليات التي تنتج أقل عدد من الأنساق)
  - ثانياً، تأكد أن ترتيب العقد الأوراق لا يسبب عمليات جداء ديكارتي
    - مثلاً إذا كانت العلاقتان ذاتا القسر الأكثر تقييداً لا تملكان شرط ربط مباشر بينهما
    - قد يكون من المرغوب تغيير ترتيب العقد الورقية لتجنب الجداءات الديكارتية

Rule 5: $R \bowtie_c S \equiv S \bowtie_c R$

$R \times S \equiv S \times R$

Rule 9: $(R\ \theta\ S)\ \theta\ T \equiv R\ \theta\ (S\ \theta\ T)$

الخطوط العريضة لخوارزمية تحسين جبري إرشادي
Outline of a heuristic algebraic optimization algorithm

الخطوة الرابعة

- باستخدام القاعدة ١٢، حول عملية جداء ديكارتي مع عملية قسر تالية في الشجرة إلى عملية ربط إذا كان الشرط يمثل شرط ربط.

Rule 12:

$$\sigma_c(R \times S) \equiv R \bowtie_c S$$

الخطوط العريضة لخوارزمية تحسين جبري إرشادي
Outline of a heuristic algebraic optimization algorithm

الخطوة الخامسة

- باستخدام القواعد ٣ و ٤ و ٧ و ١١ المتعلقة بكون الإسقاط تتابعي و تبديلي مع العمليات الأخرى ، فكك و حرك قوائم واصفات الإسقاط هبوطاً في الشجرة أعمق ما يمكن من خلال إنشاء عمليات إسقاط جديدة عند اللزوم.
- الواصفات التي يجب الإبقاء عليها بعد كل عملية إسقاط هي :
  - الواصفات اللازمة في العمليات التالية في شجرة الاستعلام
  - و الواصفات اللازمة في نتيجة الاستعلام

Rule 3: $\pi_{list1}(\pi_{list2}(\ldots(\pi_{listn}(R))\ldots)) \equiv \pi_{list1}(R)$
if $list1 \subseteq list2 \subseteq \ldots \subseteq listn$.

Rule 4: $\pi_{A1, \ldots, An}(\sigma_c(R) \equiv \sigma_c(\pi_{A1, \ldots, An}(R))$

Rule 7: $\pi_L(R \bowtie_C S) \equiv (\pi_{A1, \ldots, An}(R)) \bowtie_c (\pi_{B1, \ldots, Bm}(S))$

Rule 11: $\pi_L(R \cup S) \equiv (\pi_L(R)) \cup (\pi_L(S))$