

Επεξεργασία και Βελτιστοποίηση Ερωτήσεων



- ③ Ο βελτιστοποιητής ερωτήσεων (query optimizer)
- ③ Μετασχηματισμός εκφράσεων σχεσιακής άλγεβρας
- ③ Υπολογισμός μεγεθών πράξεων σχεσιακής άλγεβρας (επιλογή σύνδεση, άλλες πράξεις)
- ③ Τρόποι εκτέλεσης (και μέτρηση κόστους) πράξεων σχεσιακής άλγεβρας
- ③ Βελτιστοποίηση βάσει κόστους vs. Ευριστική βελτιστοποίηση

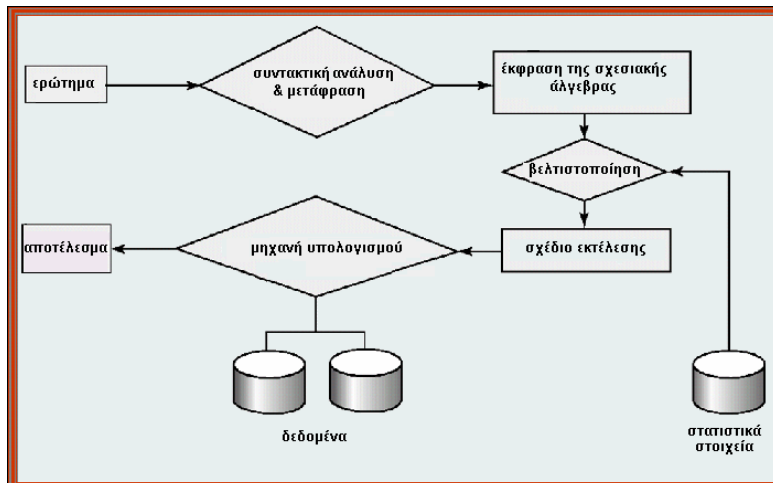
Βασική πηγή διαφανειών: Silberschatz et al., "Database System Concepts", 4/e

Επεξεργασία ερωτήσεων



- 🕒 Όταν ο client υποβάλει μια ερώτηση στο server (συνήθως σε δηλωτική γλώσσα, π.χ. SQL), ο server εκτελεί τα εξής βήματα:
 1. **Συντακτική ανάλυση** (parsing) και **μετάφραση** (translation)
 - 🕒 Μετατροπή σε έκφραση σχεσιακής άλγεβρας
 2. **Βελτιστοποίηση** (optimization)
 - 🕒 Σύνθεση εναλλακτικών πλάνων εκτέλεσης (ισοδύναμες εκφράσεις σχεσιακής άλγεβρας – χρήση ή μη ευρετηρίων για τις επιλογές – τρόποι υλοποίησης συνδέσεων)
 3. **Εκτέλεση πλάνου** (query plan execution) και **επιστροφή αποτελέσματος**

Η επεξεργασία ερωτήσεων σε μια εικόνα



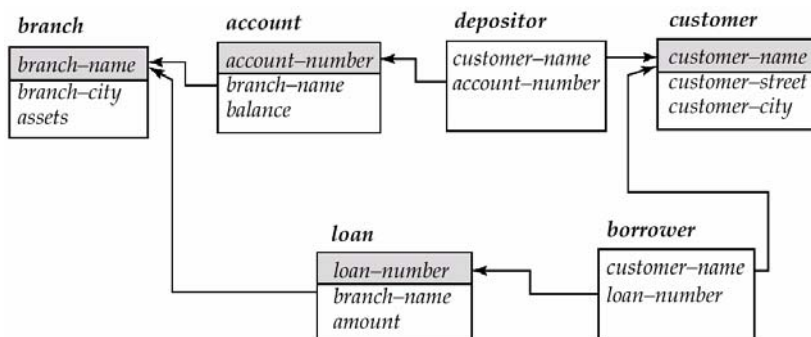
3

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ. – Σιούτας Σπύρος

Διάγραμμα Σχεσιακού σχήματος



🕒 Βάση δεδομένων τράπεζας
(πελάτες – υποκαταστήματα – καταθετικοί λογαριασμοί – δάνεια)



4

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ. – Σιούτας Σπύρος

Επεξεργασία ερωτήσεων (Συν.)



③ Συντακτική ανάλυση και Μετάφραση

③ Η ερώτηση SQL

```
SELECT customer-name
FROM account JOIN depositor
WHERE balance < 2500
```

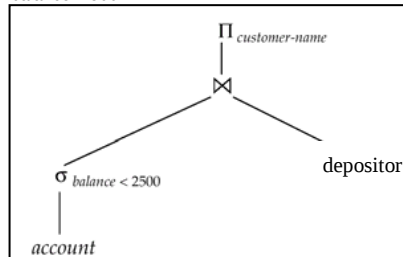
- ③ ελέγχεται για συντακτικά / σημασιολογικά σφάλματα και στη συνέχεια μεταφράζεται στην ισοδύναμη έκφραση σχεσιακής άλγεβρας:

$\Pi_{\text{customer-name}}(\sigma_{\text{balance} < 2500}(\text{account} \bowtie \text{depositor}))$

③ Βελτιστοποίηση .. $\Pi_{\text{customer-name}}(\sigma_{\text{balance} < 2500}(\text{account}) \bowtie \text{depositor})$

③ Εκτέλεση πλάνου και επιστροφή αποτελέσματος

- ③ Η μηχανή υπολογισμού δέχεται ως είσοδο ένα **σχέδιο εκτέλεσης** (query execution plan - QEP), το εκτελεί και επιστρέφει τις απαντήσεις στον client.



5

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ. – Σιούτας Σπύρος

Βελτιστοποίηση ερωτήσεων



- ③ Μια έκφραση της σχεσιακής άλγεβρας ενδέχεται να έχει πολλές ισοδύναμες εκφράσεις

③ Π.χ. η έκφραση $\sigma_{\text{balance} < 2500}(\Pi_{\text{balance}}(\text{account}))$ ισοδυναμεί με την έκφραση $\Pi_{\text{balance}}(\sigma_{\text{balance} < 2500}(\text{account}))$

- ③ Κάθε πράξη της σχεσιακής άλγεβρας μπορεί να υλοποιηθεί με χρήση κάποιου από πολλούς διαφορετικούς αλγορίθμους

- ③ Άρα, δεν αρκεί μόνο ο προσδιορισμός της πράξης, αλλά χρειάζεται να προσδιοριστεί και ο αλγόριθμος που θα χρησιμοποιηθεί για την υλοποίησή της.

- ③ Ο καθορισμός μιας λεπτομερούς στρατηγικής για τον υπολογισμό μιας πράξης ονομάζεται **σχέδιο εκτέλεσης** (execution plan).

6

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ. – Σιούτας Σπύρος

Βελτιστοποίηση με ευριστικούς αλγορίθμους



- ③ Η βελτιστοποίηση βάσει κόστους είναι ακριβή, ακόμα και όταν υλοποιείται μέσω δυναμικού προγραμματισμού. Για το λόγο αυτό, πολλά συστήματα χρησιμοποιούν ευριστικούς αλγορίθμους.
- ③ Στη βελτιστοποίηση με ευριστικούς αλγορίθμους το δέντρο της ερώτησης μετασχηματίζεται μέσω ενός συνόλου κανόνων που τυπικά βελτιώνει την απόδοση της εκτέλεσης:
 - ③ **Εκτέλεση των επιλογών νωρίς** (έτσι μειώνεται το πλήθος των πλειάδων)
 - ③ **Εκτέλεση των προβολών νωρίς** (έτσι μειώνεται το πλήθος των γνωρισμάτων)
 - ③ **Εκτέλεση των πιο περιοριστικών επιλογών και συνδέσεων πριν από άλλες παρόμοιες λειτουργίες**
- ③ Άλλα συστήματα χρησιμοποιούν μόνο ευριστικούς αλγορίθμους, ενώ άλλα συνδυάζουν ευριστικούς με εν μέρει βελτιστοποίηση βάσει κόστους.

Ένας ευριστικός βελτιστοποιητής ερωτήσεων



- ③ Κανόνες
 - ③ Κανόνας 1: «εκτέλεσε τις επιλογές όσο το δυνατό νωρίτερα»
 - ③ Κανόνας 2: «αντικατέστησε τα καρτεσιανά γινόμενα με συνδέσεις όπου είναι δυνατό»
 - ③ Κανόνας 3: «αν υπάρχουν αρκετές συνδέσεις, εκτέλεσε πρώτα τις πιο περιοριστικές»
 - ③ Κανόνας 4: «με προβολή, απομάκρυνε τα άχρηστα γνωρίσματα όσο το δυνατό νωρίτερα»
- ③ Όλοι οι κανόνες έχουν στόχο την δημιουργία μικρών ενδιάμεσων σχέσεων

Ένας ευριστικός βελτιστοποιητής ερωτήσεων (συν.)



③ Μεθοδολογία

- ③ Μετασχηματισμός SQL ερώτησης στην ("τετριμμένη") ισοδύναμη έκφραση Σχεσιακής Άλγεβρας
- ③ Εφαρμογή αλγεβρικών μετασχηματισμών (Κανόνες 1, 2, 3, 4) → εναλλακτικά πλάνα εκτέλεσης
- ③ Επιλογή του 'βέλτιστου πλάνου' και εκτέλεσή του

③ Π.χ. δίδεται η SQL:

```
SELECT DISTINCT D.customer-name
FROM   account A, depositor D
WHERE  A.account-number = D.account-number
AND    balance < 2500
```

... που ισοδυναμεί με την ακόλουθη έκφραση σχεσιακής άλγεβρας:

$\Pi_{\text{customer-name}}(\sigma_{\text{balance} < 2500 \wedge \text{A.account-number} = \text{D.account-number}} (A \times D))$

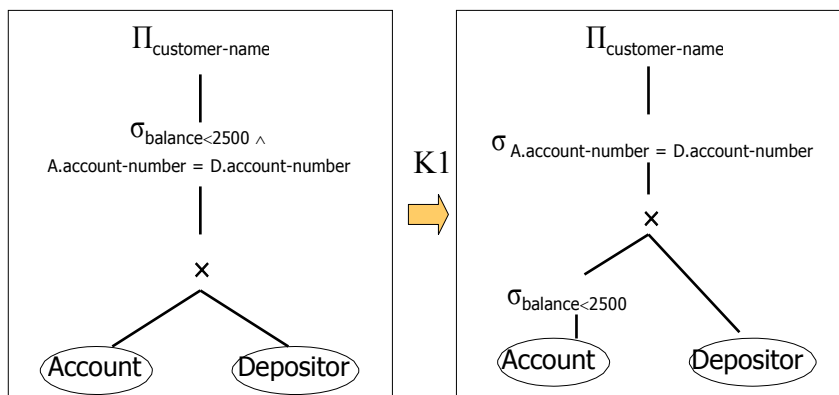
57

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ. – Σιούτας Σπύρος

Παράδειγμα ευριστικής βελτιστοποίησης



«εκτέλεσε τις επιλογές όσο το δυνατό νωρίτερα»



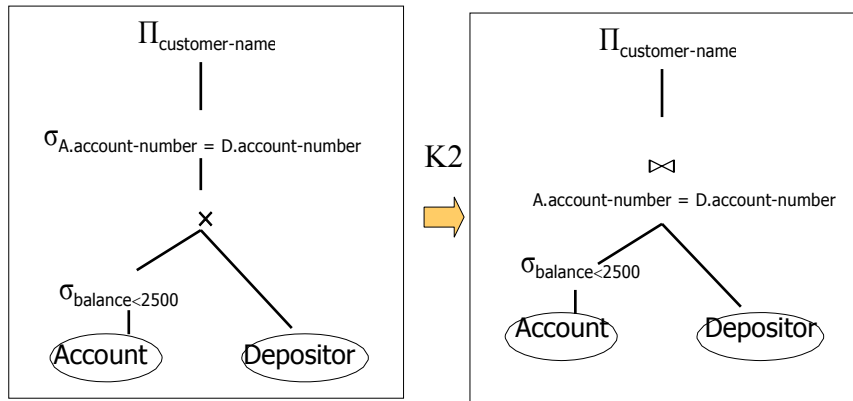
58

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ. – Σιούτας Σπύρος

Παράδειγμα ευριστικής βελτιστοποίησης (συν.)



«αντικατέστησε τα καρτεσιανά γινόμενα με συνδέσεις, όπου είναι δυνατό»



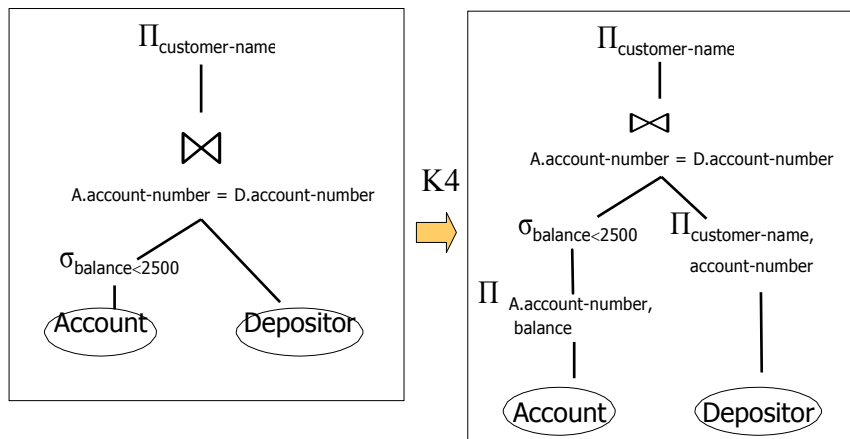
59

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ. – Σιούτας Σπύρος

Παράδειγμα ευριστικής βελτιστοποίησης (συν.)



«με προβολή, απομάκρυνε τα άχρηστα γνώρισματα όσο το δυνατό νωρίτερα»



60

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ. – Σιούτας Σπύρος

Η βελτιστοποίηση στα εμπορικά ΣΔΒΔ



	IBM DB2	IBM Informix	Microsoft SQL Server	Oracle	Sybase
Τεχνικές Βελτιστοποίησης	?	?	?	?	?

- ③ Όλα τα εμπορικά ΣΔΒΔ κρατάνε ως «επτασφράγιστα μυστικά» τις λεπτομέρειες κατασκευής του βελτιστοποιητή ερωτήσεων που έχουν υλοποιήσει. Πάντως:
 - ③ Όλα χρησιμοποιούν ιστογράμματα για την εκτίμηση του μεγέθους του αποτελέσματος
 - ③ Όλα ευνοούν τη χρήση ευρετηρίων σε πλάνα εκτέλεσης
 - ③ Όλα αναζητούν δέντρα αριστερού βάθους χρησιμοποιώντας δυναμικό προγραμματισμό
- ③ Η ανάπτυξη ενός τέτοιου λογισμικού αντιστοιχεί σε εργασία δεκάδων ανθρωποετών.



Τέλος ενότητας!