Exam revision

Database Systems Lecture 19
Natasha Alechina

In This Lecture

- Exam format
- Main topics
- How to revise

Exam format

- Answer three questions out of five
- Each question is worth 25 points
- I will only mark three questions in the order you answer them
- (So cross out any answers you don't want marked)
- Final mark for the module is your coursework mark (at most 25) plus your exam mark (at most 75).

Main topics

- What is a database, what is a DBMS, data manipulation language, data definition language, data control language
- Relational model
  - Relations, attributes, domains
  - candidate keys, primary keys, foreign keys, entity integrity, referential integrity)
- Relational algebra
- SQL (the kind of questions you did in cw2, cw3, cw6)
- Normalisation (1NF, 2NF, 3NF, BCNF)
- Security, privileges (how to grant and revoke them)
- Transactions, recovery
- Concurrency

How to revise

- Do all the exercises, then check the model solutions
- Remember SQL syntax – you will have to write SQL queries in the exam
- Look at the previous exam papers (for G51DBS06-07, G51DBS07-08, and G52DBS)
- Exam for last year and answers are now on the web
- If you get stuck with some previous exam paper questions, send me an email – I will either answer by email or, if I get a lot of similar questions, arrange a tutorial

Particular topics

Since I did not get any requests for revision questions, I will go through some topics I already had slides on:

- Normalisation
- Relational Algebra

Normalisation

- General idea: given relation R
  - Find all candidate keys in R
  - Find all non-trivial functional dependencies in R
  - Decomposing to XNF: for every functional dependency $A \rightarrow B$ which is "bad" with respect to XNF, decompose R into π_{AB} (R) and π_{AC} (R) where C is the rest of R's attributes.

Candidate keys

- A set of attributes A (can be a singleton set) is a *candidate key* for relation R if it has properties of
  - Uniqueness: no two different tuples in R can have the same values for attributes in A
  - Minimality: no subset of A has the uniqueness property
- A set of attributes is a *superkey* if it includes a candidate key
- We call an single attribute a *key attribute* if it is part of a candidate key.

Functional dependencies

- A, B are sets of attributes of relation R.
- A *functional dependency* $A \rightarrow B$ holds for R if for every two tuples in R, if they have the same values for A, then they have the same values for B.
- *Non-trivial* functional dependency: $A \rightarrow B$ is non-trivial if B is not a subset of A.

"Bad" functional dependencies

- For 2NF: $A \rightarrow B$ where A is a part of a candidate key and B is a non-key attribute
- For 3NF: $A \rightarrow B$, $B \rightarrow C$ where C is a not a key attribute

(alternative definition of 3NF: bad fd $A \rightarrow B$ is where A is not a superkey and B is non-key attribute)

- For BCNF: non-trivial $A \rightarrow B$ where A is not a superkey.
- Example of 3NF but not BCNF: R(A,B,C), candidate keys (A,B) and (A,C), fd $B \rightarrow C$.

Revision of relational algebra

- Operations: $\cup$ (union), $-$ (difference), $\times$ (product), π (projection), σ (selection)
- Other operations are definable using the ones above: $\cap$ (intersection), $\bowtie$ (natural join – can be defined using $\times$, σ and π)

Revision of relational algebra

- Union-compatible relations: same number of attributes/columns, with the same domains
- For *named* relations (as SQL tables, where columns have names), also the names of the attributes/columns should be the same

Example: not union-compatible

Name	email	telephone
Bob	bbb	222222
Chris	ccc	333333

Name	DOB
Sam	1985
Steve	1986

Example: not union-compatible

(different domains for the second column)

Name	Email (domain: char(3))
Bob	bbb
Chris	ccc

Name	DOB (domain: int)
Sam	1985
Steve	1986

Example: union-compatible

Name	DOB
Bob	1971
Chris	1972

Name	DOB
Sam	1985
Steve	1986

Union of two relations

- Let R and S be two union-compatible relations. Then their union $R \cup S$ is a relation which contains tuples from both relations:

$R \cup S = \{x: x \in R \text{ or } x \in S\}$.

Example: shopping lists

R

Name	Price
Milk	0.80
Bread	0.60
Eggs	1.20
Soap	1.00

S

Name	Price
Cream	5.00

$R \cup S$

Name	Price
Milk	0.80
Bread	0.60
Eggs	1.20
Soap	1.00
Cream	5.00

Difference of two relations

Let R and S be two union-compatible relations. Then their ***difference*** $R - S$ is a relation which contains tuples which are in R but not in S:

$R - S = \{x: x \in R \text{ and } x \notin S\}$.

Example

R

Name	Price
Milk	0.80
Bread	0.60
Eggs	1.20
Soap	1.00

S

Name	Price
Soap	1.00

R – S

Name	Price
Milk	0.80
Bread	0.60
Eggs	1.20

(Extended Cartesian) product of relations

A relation containing all tuples of the form:

<tuple from R, tuple from S>:

$R \times S = \{<c_1,\ldots,c_n, c_{n+1},\ldots,c_{n+m}>:$

$<c_1, \ldots,c_n> \in R, \ <c_{n+1},\ldots,c_{n+m}> \in S\}$

(this assumes R has n columns and S has m columns)

Example

R

Name	Price
Milk	0.80
Bread	0.60
Eggs	1.20
Soap	1.00

S

Name	Calories
Milk	200
Bread	300

R×S

Name	Price	Name	Cal
Milk	0.80	Milk	200
Bread	0.60	Milk	200
Eggs	1.20	Milk	200
Soap	1.00	Milk	200
Cheese	1.34	Bread	300
Milk	0.80	Bread	300
Bread	0.60	Bread	300
Eggs	1.20	Bread	300
Soap	1.00	Bread	300

Projection

- Let R be a relation with n columns, and X is a set of column names. Then projection of R on X is a new relation $\pi_X(R)$ which only has columns from X.

Example: $\pi_{Name,Telephone}(R)$

R:

Name	Email	Telephone
Bob	bbb@cs.nott.ac.uk	0115222222
Chris	ccc@cs.nott.ac.uk	0115333333

Example: $\pi_{Name,Telephone}(R)$

$\pi_{Name,Telephone}(R)$:

Name	Telephone
Bob	0115222222
Chris	0115333333

Selection

- Let R be a relation and α is a property of tuples from R.
- ***Selection from R subject to condition α*** is defined as follows:

$\sigma_{\alpha}(R) = \{<a_1,\ldots,a_n> \in R: \alpha(a_1,\ldots,a_n)\}$

Example: selection

- $\sigma_{\text{Year} < 2002 \text{ and Director} = \text{Nolan}}(R)$

R

Title	Director	Year
Insomnia	Nolan	2002
Magnolia	Anderson	1999
Insomnia	Skjoldbjaerg	1997
Memento	Nolan	2000
Gattaca	Niccol	1997

Example: selection

- $\sigma_{\text{Year} < 2002 \text{ and Director} = \text{Nolan}}(R)$

Title	Director	Year
Memento	Nolan	2000

Example (small) exam question

- What is the result of $\pi_{\text{R.Name,S. Name}}(\sigma_{\text{R.Tel} = \text{S.Tel}}(R \times S))$, where R and S are:

R

Name	Tel
Anne	111111
Bob	222222

S

Name	Tel
Chris	333333
Dan	111111

Example exam question

$R \times S$

R.Name	R.Tel	S.Name	S.Tel
Anne	111111	Chris	333333
Anne	111111	Dan	111111
Bob	222222	Chris	333333
Bob	222222	Dan	111111

Example exam question

$\sigma_{\text{R.Tel} = \text{S.Tel}}(R \times S)$

R.Name	R.Tel	S.Name	S.Tel
Anne	111111	Dan	111111

Example exam question

$\pi_{R.Name,S.Name}(\sigma_{R.Tel = S.Tel}(R \times S))$

R.Name	S.Name
Anne	Dan

Any questions?