Challenge: an SQL exercise

Maarten Fokkinga; feb 23th, 2003 (version of March 5, 2003, 11:19)

In an examination for the 2nd year course Databases I've given the exam question below. None of the students produced a correct answer. I present the exercise here as a challenge for members of the database group, and other interested people.

Consider the database schema:

```
Battle (name, ...)
Class (name, country, ...)
Ship (name, classname, ...)
         classname references Class(name)
Outcome (shipname, battlename, result)
         shipname references Ship(name)
         battlename references Battle(name)
```

Attributes belonging to the primary key have been underlined.

All ships of a class come from one country, the country mentioned in the class of the ship. Tabel `Outcome` indicates whether and how a ship survived a battle: `result` = *sunk*, *damaged*, or *ok*.

Exercise (8 points out of 100) Formulate in Set notation and in SQL:

Find the countries of which *all* ships have sunk.

Hint (not given at the exam): in the lectures and in additional written course material I've stressed the fact that for the more difficult "SQL questions" you really need to apply set and predicate notation and manipulations in order to construct a correct SQL formulation.

By the way, the preceding question read: Find the countries of which *some* ships have sunk.

On the next page I explain my set and predicate notation, and the translation to SQL.
An elaborate solution is on page 3.
Frequently given wrong answers are on page 4.
So, don't look at pages 3 and 4 if you want to take up the challenge.

Set and predicate notation

We use the following set and predicate notation (known as the 'Z notation'):

$\{D \mid P \bullet E\}$: the set of all values expressed by E, where the free variables range over the domains as specified by D but only insofar as they satisfy P

$\exists\, D \mid P \bullet P'$: "there exists values as declared by D and satisfying P, for which P' holds"

$\forall\, D \mid P \bullet P'$: "for all values as declared by D and satisfying P, it is true that P' holds"

In all three cases the part '$\mid P$' may be omitted if P is *true*.

Apart from the usual rules of logic and set theory, we also have the following equations and equivalences:

$$\exists\, D \mid P \bullet P' \quad \Leftrightarrow \quad \exists\, D \bullet P \wedge P'$$
$$\forall\, D \mid P \bullet P' \quad \Leftrightarrow \quad \forall\, D \bullet P \Rightarrow P'$$

$$\neg\ \exists\, D \mid P \bullet P' \quad \Leftrightarrow \quad \forall\, D \mid P \bullet \neg\ P'$$
$$\neg\ \forall\, D \mid P \bullet P' \quad \Leftrightarrow \quad \exists\, D \mid P \bullet \neg\ P'$$

$$\{D \mid (\exists\, D' \bullet P') \wedge P \bullet E\} \quad = \quad \{D; D' \mid P' \wedge P \bullet E\}$$

In the latter line, it is assumed that no name clash occurs (that is, the variables declared in D' do not occur free in P and E).

A rule of logic that is used (and that occurs frequently) is the 'one-point rule':

$$(\exists\, x : S \bullet x = val) \quad \Leftrightarrow \quad val \in S$$

Translation to SQL

SQL queries may be considered as a restricted form of set and predicate notation. The restriction in set notation is that the declarations D must have a special form (tables only), and not all set operations are allowed. The restriction in predicate notation is that $\Rightarrow$ and $\forall$ (in its full generality) are not allowed. These restrictions have to be met by suitable manipulations (rules of logic and set theory). Furthermore, the lexical notation is slightly different. In general, a translation looks as follows:

$\{D \mid P \bullet E\}$ = `select` E' `from` D' `where` P'

$\exists\, D \mid P \bullet Q$ = `exists (select * from` D' `where` P' `and` Q'`)`

where D' is the SQL formulation of D, and similarly for P, Q, and E. For example:

The countries of which *some* ships have sunk
$= \{c : C \mid (\exists\, s : S;\ o : O \bullet c.n = s.c \wedge s.n = o.s \wedge o.r = sunk) \bullet c.c\}$
$= \{c : C;\ s : S;\ o : O \mid c.n = s.c \wedge s.n = o.s \wedge o.r = sunk \bullet c.c\}$
= `select c.c from C c, S s, O o where c.n=s.c and s.n=o.s and o.r=sunk`

Since SQL works with tables rather than sets, occasionally a '`distinct`' may be needed when duplicates are to avoided.

Solutions

Abbreviating the identifiers to their first letter, the schema reads:

```
B (n, ...)
C (n, c, ...)      -- c abbreviates: country
S (n, c, ...)      -- c abbreviates: classname
O (s, b, r)
```

The answer 1

$= $ "the countries of which all ships have sunk" 2

$= \{t : \mathit{Text} \mid$ "t is a country" $\wedge (\forall s : S \mid$ "s comes from country t" $\bullet$ "s has sunk") $\bullet\ t\}$ 3

$= \{t : \mathit{Text} \mid (\exists c : C \bullet t = c.c) \wedge (\forall s : S \mid$ "s comes from country t" $\bullet$ "s has sunk") $\bullet\ t\}$ 4

$= \{t : \mathit{Text};\ c : C \mid t = c.c \wedge (\forall s : S \mid$ "s comes from country t" $\bullet$ "s has sunk") $\bullet\ t\}$ 5

$= \{t : \mathit{Text};\ c : C \mid t = c.c \wedge (\forall s : S \mid$ "s comes from country $c.c$" $\bullet$ "s has sunk") $\bullet\ c.c\}$ 6

$= \{c : C \mid (\exists t : \mathit{Text} \bullet t = c.c) \wedge (\forall s : S \mid$ "s comes from country $c.c$" $\bullet$ "s has sunk") $\bullet\ c.c\}$ 7

$= \{c : C \mid (\forall s : S \mid$ "s comes from country $c.c$" $\bullet$ "s has sunk") $\bullet\ c.c\}$ 8

$= \{c : C \mid \neg\ (\exists s : S \mid$ "s comes from country $c.c$" $\bullet \neg$ "s has sunk") $\bullet\ c.c\}$ 9

$= \{c : C \mid \neg\ (\exists s : S \bullet$ "s comes from country $c.c$" $\wedge \neg$ "s has sunk") $\bullet\ c.c\}$ 10

$= \{c : C \mid \neg\ (\exists s : S \bullet (\exists c' : C \bullet s.c = c'.n \wedge c'.c = c.c) \wedge \neg$ "s has sunk") $\bullet\ c.c\}$ 11

$= \{c : C \mid \neg\ (\exists s : S;\ c' : C \bullet s.c = c'.n \wedge c'.c = c.c \wedge \neg$ "s has sunk") $\bullet\ c.c\}$ 12

$= \{c{:}C \mid \neg\ (\exists s{:}S;\ c'{:}C \bullet s.c = c'.n \wedge c'.c = c.c \wedge \neg\ (\exists o : O \mid o.r = \mathit{sunk} \bullet o.s = s.n)) \bullet c.c\}$ 13

```
= select distinct c.c from C c where not exists (                          14
     select * from S s, C c1 where s.c=c1.n and c1.c=c.c and not exists (   15
       select * from O o where o.r=sunk and o.s=s.n ))                     16
```

$= \{c{:}C \mid \neg\ (\exists s{:}S;\ c'{:}C \bullet s.c = c'.n \wedge c'.c = c.c \wedge (\forall o : O \mid o.r = \mathit{sunk} \bullet o.s \neq s.n)) \bullet c.c\}$ 17

$= \{c{:}C \mid \neg\ (\exists s{:}S;\ c'{:}C \bullet s.c = c'.n \wedge c'.c = c.c \wedge s.n \notin \{o : O \mid o.r = \mathit{sunk} \bullet o.s\}) \bullet c.c\}$ 18

```
= select distinct c.c from C c where not exists (                          19
     select * from S s, C c1 where s.c=c1.n and c1.c=c.c and               20
     s.n NOT IN (select o.s from O o where o.r=sunk ))                     21
```

$= \{c{:}C \mid \neg\ (\exists s{:}S;\ c'{:}C \mid s.c = c'.n \wedge s.n \notin \{o : O \mid o.r = \mathit{sunk} \bullet o.s\} \bullet c'.c = c.c) \bullet c.c\}$ 22

$= \{c : C \mid c.c \notin \{s : S;\ c' : C \mid s.c = c'.n \wedge s.n \notin \{o : O \mid o.r = \mathit{sunk} \bullet o.s\} \bullet c'.c\} \bullet c.c\}$ 23

$= \{c : C \mid c.c\} \setminus \{s : S;\ c : C \mid s.c = c.n \wedge s.n \notin \{o : O \mid o.r = \mathit{sunk} \bullet o.s\} \bullet c.c\}$ 24

$= $ "all countries" $\setminus$ "countries of ships that have not sunk" 25

```
= (select c.c from C c)                                                    26
  EXCEPT                                                                   27
  (select c.c from S s, C c where                                          28
     s.c=c.n and s.n NOT IN (select o.s from O o where o.r=sunk ))         29
```

Line 13 can be written with an $\forall$-quantifier, as follows:

$\{c{:}C \mid \neg\ (\exists s{:}S;\ c'{:}C \bullet s.c = c'.n \wedge c'.c = c.c \wedge \neg\ (\exists o : O \mid o.r = \mathit{sunk} \bullet o.s = s.n)) \bullet c.c\}$

$= \{c{:}C \mid \neg\ (\exists s{:}S;\ c'{:}C \mid s.c = c'.n \wedge c'.c = c.c \bullet \neg\ (\exists o : O \mid o.r = \mathit{sunk} \bullet o.s = s.n)) \bullet c.c\}$

$= \{c{:}C \mid \quad (\forall s{:}S;\ c'{:}C \mid s.c = c'.n \wedge c'.c = c.c \bullet \quad (\exists o : O \mid o.r = \mathit{sunk} \bullet o.s = s.n)) \bullet c.c\}$

but a translation to SQL should first eliminate the $\forall$ by going back to line 13.

Frequently given wrong answers

$\{c : C;\ s : S;\ o : O \mid c.n = s.c \wedge s.n = o.s \wedge \ldots \bullet c.c\}$

= countries with a ship that participated in a battle and ....

$\{c : C;\ s : S \mid c.n = s.c \wedge (\forall\, o : O \mid o.s = s.n \bullet o.r = sunk) \bullet c.c\}$

$= \{c : C;\ s : S \mid c.n = s.c \wedge \neg\, (\exists\, o : O \mid o.s = s.n \bullet o.r \neq sunk) \bullet c.c\}$

= countries with a ship that survived each battle in which it participated.

$\{c : C \mid (\forall\, s : S \mid c.n = s.c \bullet \forall\, o : O \bullet s.n \neq o.s \vee o.r \neq sunk) \bullet c.c\}$

$= \{c : C \mid (\forall\, s : S \mid c.n = s.c \bullet \forall\, o : O \mid s.n = o.s \bullet o.r \neq sunk) \bullet c.c\}$

$= \{c : C \mid (\forall\, s : S \mid c.n = s.c \bullet \neg\ \exists\, o : O \mid s.n = o.s \bullet o.r = sunk) \bullet c.c\}$

$= \{c : C \mid (\forall\, s : S \mid c.n = s.c \bullet \neg\ \exists\, o : O \bullet s.n = o.s \wedge o.r = sunk) \bullet c.c\}$

$= \{c : C \mid \neg\, (\exists\, s : S \mid c.n = s.c \bullet \exists\, o : O \bullet s.n = o.s \wedge o.r = sunk) \bullet c.c\}$

$= \{c : C \mid \neg\, (\exists\, s : S;\ o : O \bullet c.n = s.c \wedge s.n = o.s \wedge o.r = sunk) \bullet c.c\}$

= countries of classes of which all ships have not sunk.

$\{c : C \mid (\forall\, s : S \mid c.n = s.c \bullet \exists\, o : O \bullet s.n = o.s \wedge o.r = sunk) \bullet c.c\}$

$= \{c : C \mid \neg\, (\exists\, s : S \mid c.n = s.c \bullet \neg\ \exists\, o : O \bullet s.n = o.s \wedge o.r = sunk) \bullet c.c\}$

= countries of classes of which all ships have sunk

$\{c : C \mid (\forall\, s : S \mid c.n = s.c \bullet \forall\, o : O \bullet s.n \neq o.s \vee o.r = sunk) \bullet c.c\}$

$= \{c : C \mid (\forall\, s : S \mid c.n = s.c \bullet \forall\, o : O \mid s.n = o.s \bullet o.r = sunk) \bullet c.c\}$

$= \{c : C \mid (\forall\, s : S \mid c.n = s.c \bullet \neg\ \exists\, o : O \mid s.n = o.s \bullet o.r \neq sunk) \bullet c.c\}$

$= \{c : C \mid (\forall\, s : S \mid c.n = s.c \bullet \neg\ \exists\, o : O \bullet s.n = o.s \wedge o.r \neq sunk) \bullet c.c\}$

$= \{c : C \mid \neg\, (\exists\, s : S \mid c.n = s.c \bullet \exists\, o : O \bullet s.n = o.s \wedge o.r \neq sunk) \bullet c.c\}$

$= \{c : C \mid \neg\, (\exists\, s : S;\ o : O \bullet c.n = s.c \wedge s.n = o.s \wedge o.r \neq sunk) \bullet c.c\}$

= countries of classes that have no battle-surviving ships

= countries of classes of which each ship sunk in every battle in which it participated.

countries of which each ship has sunk

$\subset$ countries of which each ship has sunk in every battle in which it participated

$= \{c{:}C \mid (\forall\, s{:}S \mid (\exists\, c'{:}C \bullet c.c = c'.c \wedge c'.n = s.c) \bullet \forall\, o{:}O \mid s.n = o.s \bullet o.r = sunk) \bullet c.c\}$

$= \{c{:}C \mid \neg\, (\exists\, s{:}S \mid (\exists\, c'{:}C \bullet c.c = c'.c \wedge c'.n = s.c) \bullet \neg\ \forall\, o{:}O \mid s.n = o.s \bullet o.r = sunk) \bullet c.c\}$

$= \{c{:}C \mid \neg\, (\exists\, c'{:}C;\ s{:}S \mid c.c = c'.c \wedge c'.n = s.c \bullet \neg\ \forall\, o{:}O \mid s.n = o.s \bullet o.r = sunk) \bullet c.c\}$

$= \{c : C \mid \neg\, (\exists\, c' : C;\ s : S \mid c.c = c'.c \wedge c'.n = s.c \bullet \exists\, o : O \mid s.n = o.s \bullet o.r \neq sunk) \bullet c.c\}$

$= \{c : C \mid \neg\, (\exists\, c' : C;\ s : S \bullet c.c = c'.c \wedge c'.n = s.c \wedge \exists\, o : O \mid s.n = o.s \bullet o.r \neq sunk) \bullet c.c\}$

$= \{c : C \mid \neg\, (\exists\, c' : C;\ s : S;\ o : O \bullet c.c = c'.c \wedge c'.n = s.c \wedge s.n = o.s \wedge o.r \neq sunk) \bullet c.c\}$

$= \{c : C \mid c.c \notin \{c' : C;\ s : S;\ o : O \mid c'.n = s.c \wedge s.n = o.s \wedge o.r \neq sunk \bullet c'.n\} \bullet c.c\}$

$= \{c : C \bullet c.c\} \setminus \{c' : C;\ s : S;\ o : O \mid c'.n = s.c \wedge s.n = o.s \wedge o.r \neq sunk \bullet c'.n \bullet c.c\}$

= all countries except for the countries with a ship that survived a battle.

= countries of which no ship survived a battle.

= countries of which each ship has sunk in every battle in which it participated

$\subset$ countries of which each ship has sunk or did not participate in a battle

$\{c : C \mid \neg\, (\exists\, c' : C;\ s : S;\ o : O \bullet c.c = c'.c \wedge c'.n = s.c \wedge s.n = o.s \wedge o.r = sunk) \bullet c.c\}$

$= \{c : C \mid c.c \notin \{c' : C;\ s : S;\ o : O \mid c'.n = s.c \wedge s.n = o.s \wedge o.r = sunk \bullet c'.n\} \bullet c.c\}$

$= \{c : C \bullet c.c\} \setminus \{c' : C;\ s : S;\ o : O \mid c'.n = s.c \wedge s.n = o.s \wedge o.r = sunk \bullet c'.n \bullet c.c\}$

= all countries except for those that have (at least) a sunken ship

= countries of which no ship has sunk.