DBMS I — Homework 1 — Solution

September 5, 2007

Consider a simplified version of the NMSU library database. We are interested in the following relations:

- book(ISBN, Title, Publisher_Code, Published_Year, Purchase_Year, Cost, Copy_No, Shelf_id)
- publisher(Publisher_Code, Name, Address)
- location(Shelf_id, location)

The primary key for each relation is underlined (e.g., the key for the relation book is the set {ISBN, Copy_No}; this is because the library might have more than one copy of the same book). The *foreign keys* are: Shelf_id in book refers to the primary key of location; and Publisher_Code in book refers to the primary key of publisher. Write the relational algebral expression for the following queries:

1. List all the books that the library has more than 5 copies.

   $$\sigma_{\texttt{Copy\_No}>5}(book)$$

   You can also use

   $$\pi_{\texttt{ISBN,Title}}(\sigma_{\texttt{Copy\_No}>5}(book))$$

   This works only under the assumption that the copies are numbered with 1,2,... and the numbers are used in order 1,2,... (for example, there is no copy #3 if copy #2 does not exist). A better way would be

   $$book_1 \leftarrow \ {}_{\texttt{ISBN}}\mathcal{G}_{\texttt{ISBN, Title, count(Copy\_No) AS Total}}(book)$$

   and the answer is:

   $$\sigma_{\texttt{total}>5}(book_1)$$

2. A computer science book is a book whose title contains the string "Computer Science." List all computer science books.

   $$\sigma_{'ComputerScience'\ in\ \texttt{Title}}(book)$$

   Here, in is the function checking for the occurrence of 'Computer Science' in Title.

3. An expensive book costs more than $US 100.00. List all the expensive books.

   $$\sigma_{\texttt{Cost}>100}(book)$$

4. A new cheap book is the one that is published in 2007 and costs less than $US 50.00. List all the new cheap books.

   $$\sigma_{\texttt{Published\_Year}=2007\wedge\texttt{Cost}<50}(book)$$

5. List all computer science books which are expensive.

$$\sigma_{'ComputerScience' \; in \; \texttt{Title} \wedge \texttt{Cost} > 100}(book)$$

6. List all new cheap computer science books.

$$\sigma_{'ComputerScience' \; in \; \texttt{Title} \wedge \texttt{Published\_Year} = 2007 \wedge \texttt{Cost} < 50}(book)$$

7. List the location of the book whose ISBN is "1234-67-8901."

$$\pi_{\texttt{location}}(\sigma_{\texttt{ISBN}='1234-67-8901'}(book \triangleright\!\triangleleft location))$$

8. A recent addition to the library is a book purchased in 2007. List all recent additions to the libary that were published by "Prentice Hall."

$$\sigma_{\texttt{Purchased\_Year}=2007 \wedge \texttt{Name}='PrenticeHall'}(book \triangleright\!\triangleleft publisher)$$

In all of the above, you might add a projection to limit the answers to the `ISBN` and `Title`. The answer is correct as well because the question **did not specifically ask** for some attributes. If the first question were "List the `ISBN` and `Title` of all the books that the library has more than 5 copies" then the first answer would not be correct.

Grading schema: 10 points for question 1–6, 20 for 7, and 20 for 8.