**BOOKS(DocId, Title, Publisher, Year)**

**STUDENTS(StId, StName, Major, Age)**

**AUTHORS(AName, Address)**

**Borrows(DocId, StId, Date)**

**has-written(DocId, AName)**

**describes(DocId, Keyword)**

---

1. List all books published by McGraw-Hill before 1990.

$$\sigma_{publisher='McGraw-Hill' \wedge Year<1990}(BOOKS)$$

**BOOKS(DocId, Title, Publisher, Year)**

**STUDENTS(StId, StName, Major, Age)**

**AUTHORS(AName, Address)**

**Borrows(DocId, StId, Date)**

**has-written(DocId, AName)**

**describes(DocId, Keyword)**

---

2. List the name of students who are older than 30 and who are not studying CS.

$$\pi_{StName}(\sigma_{Age>30 \wedge Major \neq CS}(STUDENTS))$$

**BOOKS(DocId, Title, Publisher, Year)**

**STUDENTS(StId, StName, Major, Age)**

**AUTHORS(AName, Address)**

**Borrows(DocId, StId, Date)**

**has-written(DocId, AName)**

**describes(DocId, Keyword)**

---

2. List the name of students who are older than 30 and who are not studying CS.

$$\pi_{StName}(\sigma_{Age>30}(STUDENTS))$$
$$-\pi_{StName}(\sigma_{Major='CS'}(STUDENTS))$$

Which one is more efficient?

---

$$\pi_{StName}(\sigma_{Age>30 \wedge Major \neq CS}(STUDENTS))$$

$$\pi_{StName}(\sigma_{Age>30}(STUDENTS))$$
$$-\pi_{StName}(\sigma_{Major='CS'}(STUDENTS))$$

First one is more efficient, because it only requires one pass through the data and the filtering process is much faster.

**BOOKS(DocId, Title, Publisher, Year)**

**STUDENTS(StId, StName, Major, Age)**

**AUTHORS(AName, Address)**

**Borrows(DocId, StId, Date)**

**has-written(DocId, AName)**

**describes(DocId, Keyword)**

---

3. List the names of all students who have borrowed a book and who are CS majors.

$$\pi_{StName}(\sigma_{STUDENTS.StId=Borrows.StId}(\sigma_{Major='CS'}(STUDENTS) \times Borrows))$$

Any other way?!

$$\pi_{StName}(\sigma_{STUDENTS.StId=Borrows.StId}(\sigma_{Major='CS'}(STUDENTS) \times Borrows))$$

$$\pi_{StName}(\sigma_{Major='CS'}(STUDENTS) \bowtie_{STUDENTS.StId=Borrows.StId} Borrows)$$

Any other way?!

$$\pi_{StName}(\sigma_{STUDENTS.StId=Borrows.StId}(\sigma_{Major='CS'}(STUDENTS) \times Borrows))$$

$$\pi_{StName}(\sigma_{Major='CS'}(STUDENTS) \bowtie_{STUDENTS.StId=Borrows.StId} Borrows)$$

$$\pi_{StName}(\sigma_{STUDENTS.StId=Borrows.StId}(\sigma_{Major='CS'}(STUDENTS)) \cap Borrows)$$

# Which one is the best?

$$\pi_{StName}(\sigma_{STUDENTS.StId=Borrows.StId}(\sigma_{Major='CS'}(STUDENTS) \times Borrows))$$

**Very inefficient**, unless the database is small.

$$\pi_{StName}(STUDENTS \bowtie_{STUDENTS.StId=Borrows.StId} Borrows)$$

Best when both tables are large.

$$\pi_{StName}(\sigma_{STUDENTS.StId=Borrows.StId}(\sigma_{Major='CS'}(STUDENTS)) \cap Borrows)$$

Best when one table is small.

**BOOKS(DocId, Title, Publisher, Year)**

**STUDENTS(StId, StName, Major, Age)**

**AUTHORS(AName, Address)**

**Borrows(DocId, StId, Date)**

**has-written(DocId, AName)**

**describes(DocId, Keyword)**

---

4. List the title of books written by the author 'Silberchatz'.

$$\pi_{Title}(A \bowtie_{BOOKS.DocId=has-written.DocId} B)$$

And other ways…

**BOOKS(DocId, Title, Publisher, Year)**

**STUDENTS(StId, StName, Major, Age)**

**AUTHORS(AName, Address)**

**Borrows(DocId, StId, Date)**

**has-written(DocId, AName)**

**describes(DocId, Keyword)**

---

5. List the title of books written by the author 'Silberchatz', <u>but not</u> books that have the keyword 'database'.

$$A \leftarrow \sigma_{AName='Silberchatz'}(has\_written)$$

$$B \leftarrow BOOKS$$

$$C \leftarrow \sigma_{Keyword='database'}(describes)$$

$$S1 \leftarrow A \bowtie_{BOOKS.DocId=has-written.DocId} B$$

$$S2 \leftarrow C \bowtie_{BOOKS.DocId=describes.DocId} B$$

$$\pi_{Title}(S1 - S2)$$

**BOOKS(DocId, Title, Publisher, Year)**

**STUDENTS(StId, StName, Major, Age)**

**AUTHORS(AName, Address)**

**Borrows(DocId, StId, Date)**

**has-written(DocId, AName)**

**describes(DocId, Keyword)**

---

6. Find the title of the oldest book.

$$\rho_{B1}(BOOKS) \qquad \rho_{B2}(BOOKS)$$

$$\pi_{Title}(BOOKS) - \pi_{B1.Title}(B1 \bowtie_{B1.year>B2.year} B2)$$

**Suppliers(sid: integer, sname: string, address: string)**

**Parts(pid: integer, pname: string, color: string)**

**Catalog(sid: integer, pid: integer, cost: real)**

---

1. Find the *sid* s of supplier who supply some red or green part.

$$\pi_{sid}(\sigma_{color=red \vee color=green}(Parts)) \bowtie (Catalog)$$

**Suppliers(sid: integer, sname: string, address: string)**

**Parts(pid: integer, pname: string, color: string)**

**Catalog(sid: integer, pid: integer, cost: real)**

---

2. Find the *pid s* of part supplied by at least two different suppliers.

$$\rho(R1, C)$$
$$\rho(R2, C)$$
$$R3 \leftarrow R1 \bowtie_{R1.pid=R2.pid \wedge R1.sid \neq R2.sid} R2$$
$$\pi_{R1.pid}(R3)$$