CS341

Exercise 1

User

Id	Name	Age	Gender	OccupationId	CityId
1	John	25	Male	1	3
2	Sara	20	Female	3	4
3	Victor	31	Male	2	5
4	Jane	27	Female	1	3

Occupation

OccupationId	OccupationName
1	Software Engineer
2	Accountant
3	Pharmacist
4	Library Assistant

City

CityId	CityName
1	Halifax
2	Calgary
3	Boston
4	New York
5	Toronto

1. Solve the following relational expressions for above relations.

a. $\Pi_{\text{Name}}(\sigma_{\text{Age} > 25}(\text{User}))$

b. $\sigma_{\text{id} > 2 \vee \text{Age} != 31}(\text{User})$

c. $\sigma_{\text{User.OccupationId} = \text{Occupation.OccupationId}}(\text{User X Occupation})$

d. User $\bowtie$ Occupation $\bowtie$ City

e. $\Pi_{\text{Name, Gender}}(\sigma_{\text{CityName} = \text{"Boston"}}(\text{User} \bowtie \text{City}))$

2.

r(author)

author_id	first_name	last_name
1	John	McCarthy
2	Dennis	Ritchie
3	Ken	Thompson
4	Claude	Shannon
5	Alan	Turing
6	Alonzo	Church
7	Perry	White
8	Moshe	Vardi
9	Roy	Batty

r(author_pub)

author_id	pub_id	author_position
1	1	1
2	2	1
3	2	2
4	3	1
5	4	1
5	5	1
6	6	1

r(book)

book_id	book_title	month	year	editor
1	CACM	April	1960	8
2	CACM	July	1974	8
3	BST	July	1948	2
4	LMS	November	1936	7
5	Mind	October	1950	NULL
6	AMS	Month	1941	NULL
7	AAAI	July	2012	9
8	NIPS	July	2012	9

r(pub)

pub_id	title	book_id
1	LISP	1
2	Unix	2
3	Info Theory	3
4	Turing Machines	4
5	Turing Test	5
6	Lambda Calculus	6

a. How many tuples will the following relational algebra query return?

$$\Pi_{\text{book_title}} \text{Book}$$

b. What question does the following expression answer?

$$\Pi_{\text{author_id}} (\text{author}) - \Pi_{\text{editor}} (\text{book})$$

c. Write a relational algebra expression that returns the full names of all authors who are book editors.

d. Write a relational algebra expression that returns the full names of all authors who are **NOT** book editors.

3. Compute the theta-join of R and S with the condition R.B = S.B AND R.A < S.C

A	**B**	
1	a	
7	t	
2	g	**D**
4	c	
9	t	
B		
c	5	6
a	7	8
t	8	9

4. Use the following relation to answer the question that follow

Actor

actorId	name	nationality	age
LDC21	Leonardo DiCaprio	American	40
KW871	Kate Winslet	British	39
CB379	Christian Bale	British	40
MKE12	Michael Keaton	American	63
JGL81	Joseph Gordon-Levitt	American	33
EMG32	Ewan McGregor	British	43
HBC54	Helena Bonham Carter	British	48

Director

directorId	name	nationality
CN345	Christopher Nolan	British
JC212	James Cameron	Canadian
SM521	Sam Mendes	British
SCO78	Martin Scorsese	American
BUR34	Tim Burton	American

Performance

actorId	filmId	character
LDC21	INC10	Dominic Cobb
LDC21	TIT97	Jack Dawson
KW871	TIT97	Rose DeWitt Bukater
LDC21	RR008	Frank Wheeler
KW871	RR008	April Wheeler
LDC21	SHI10	Teddy Daniels
CB379	DK008	Bruce Wayne
CB379	DKR12	Bruce Wayne
JGL81	INC10	Arthur
MKE12	BAT92	Bruce Wayne
EMG32	FISH4	Ed Bloom
HBC54	FISH4	Jenny

Film

filmId	title	year	directorId
INC10	Inception	2010	CN345
TIT97	Titanic	1997	JC212
RR008	Revolutionary Road	2008	SM521
SKF12	Skyfall	2012	SM521
SHI10	Shutter Island	2010	SCO78
DK008	The Dark Knight	2008	CN345
DKR12	The Dark Knight Rises	2012	CN345
BAT92	Batman Returns	1992	BUR34
FISH4	Big Fish	2003	BUR34

a. For each of the following queries in relational algebra, calculate the output table and give a brief statement of what query it answers.

(a) $\sigma_{age>45}(\text{Actor})$

(b) $\pi_{title}(\text{Film})$

(c) $\pi_{title}(\sigma_{year<2000}(\text{Film}))$

(d) $\sigma_{year=2012}(\text{Film}) \times \sigma_{nationality\neq'American'}(\text{Director})$

(e) $\sigma_{year=2012}(\text{Film}) \bowtie \sigma_{nationality\neq'American'}(\text{Director})$

(f) $\pi_{title}(\text{Film} \bowtie \sigma_{nationality='British'}(\text{Director}))$

(g) $\sigma_{year<2000}(\text{Film}) \cup \sigma_{year>2010}(\text{Film})$

(h) $\sigma_{year\geq2000}(\text{Film}) \cap \sigma_{year\leq2010}(\text{Film})$

b. For each of the following questions, formulate the specified queries using relational algebra.

i. Retrieve details of all films that were released in 2010. The output schema should be the same as that of the Film table.
ii. Retrieve details of all actors that are not in their thirties. The output schema should be the same as that of the Actor table.
iii. Retrieve the names of all directors.
iv. Retrieve the names of all American directors.
v. Find out the names of all British actors above the age of 40.
vi. Retrieve the name of each actor together with the titles of the films he/she has performed in.
vii. Find out the names of all actors that have played the character of Bruce Wayne (Batman).
viii. Retrieve the names of all actors that have played the character of Bruce Wayne, together with the year the corresponding films were released.
ix. Retrieve all actors from the film Inception. The output schema should be the same as that of the Actor table.
x. Find out the names of all actors that have performed in a film directed by Christopher Nolan.
xi. Retrieve the titles of all films in which Leonardo Di Caprio and Kate Winslet have co-acted.
xii. Assuming that the actorId and directorId values for actors and directors are consistent across the tables, retrieve details of all actors that have directed a film.