CSE 233 **Spring, 2018**

Problem Set #1

Due on Thursday , April 19

Please typeset your answer (latex recommended).

1. The movie database consists of the following two relations

movie	title	director	actor

schedule	theater	title

The first relation provides titles, directors, and actors of various movies. Assume a movie is uniquely identified by its title. Do not assume that a movie has a unique director (so a movie by Hitchcock is one for which *one of the directors* is Hitchcock). The second relation provides the titles of currently playing movies and the theaters where they are being shown. Express the following queries in relational calculus and relational algebra.

(a) (1 *point*) List the theaters showing some movie by Hitchcock.

(b) (3 *points*) List the theaters showing only movies by Hitchcock.

2. The *division* binary operator $\div$ on relations is defined as follows. Given relations P and Q for which $att(Q) \subset att(P)$, $P \div Q$ is a relation with attributes $att(P) - att(Q)$ containing the tuples t for which $\{t\} \bowtie Q \subseteq P$. For example, if $att(P) = \{A, B\}$ and $att(Q) = \{B\}$, $P \div Q$ is the relation with attribute $\{A\}$ containing the tuples $\langle a \rangle$ for which $\langle a, b \rangle \in P$ for every tuple $\langle b \rangle \in Q$. Intuitively, $\div$ is a direct implementation of universal quantification.

(i) (2 *points*) Use $\div$ (and standard algebra operators) to express the query "List the theaters showing *every* movie by Hitchcock".

(ii) (4 *points*) Show how $P \div Q$ can be expressed using the standard relational algebra operators (you can assume, for simplicity, that $att(P) = \{A, B\}$ and $att(Q) = \{B\}$).

3. Consider the following query on the above *schedule* relation:

Find the theaters showing more than one title

(i) (2 *points*) Express this query in relational calculus and relational algebra.

(ii) (6 *points*) Prove that every relational algebra expression defining the above query *must* use the attribute renaming operator δ.

4. (Automorphisms)

(i) (5 *points*) Let σ be a database schema and I an instance of σ. A one-to-one mapping $f : \mathbf{dom} \to \mathbf{dom}$ is an automorphism of I iff $f(I) = I$ (f is applied to tuples componentwise). For example, if I is the binary relation instance

a	c
b	c
c	d

then the mapping given by the table

f			
	a	$\to$	b
	b	$\to$	a
	c	$\to$	c
	d	$\to$	d

is an automorphism of I but the mapping

f			
	a	$\to$	a
	b	$\to$	b
	c	$\to$	d
	d	$\to$	c

is not.

Show that CALC queries without constants are invariant under automorphisms. In other words, if Q is a CALC query without constants over schema σ, then for every instance I over σ, if f is an automorphism of I then f is also an automorphism of $Q(I)$.
Hint: Assume without loss of generality that the CALC formula φ defining Q uses only $\wedge, \exists, \neg$, and use structural induction on the formula.

(ii) (2 *points*) Let σ consist of a binary relation R. Using (i), show that there is no CALC query without constants which on input

R	
a	b
b	c
c	d
d	a

produces as answer

a	c
b	d