In **family_history** database, there is a **person** table, where people are identified by their name, and always have their gender, date of birth (**dob**) and place of birth (**born_in**) recorded. In addition, each person may optionally have recorded the name of their father, and the name of their mother. If the person has died, then the date of death **dod** must be present. Note that only a fragment of the data held in the database is listed below.

person						
<u>name</u>	gender	dob	dod?	father?	mother?	born_in
Alice	F	1885-02-25	1969-12-05	null	null	Windsor
Andrew	M	1960-02-19	null	Philip	Elizabeth II	London
Andrew of Greece	M	1882-02-02	1944-12-03	George I of Greece	null	Athens
Anne (Princess)	F	1950-08-15	null	Philip	Elizabeth II	London
Charles	M	1948-11-14	null	Philip	Elizabeth II	London
Elizabeth II	F	1926-04-21	null	George VI	Elizabeth	London

$$\vdots$$

$$\text{person(father)} \stackrel{fk}{\Rightarrow} \text{person(name)}$$
$$\text{person(mother)} \stackrel{fk}{\Rightarrow} \text{person(name)}$$

Questions

In the following questions you can test for a value v being null using the predicate isNull(v), and v being not null using isNotNull(v). You may use subcripts on relation names to creat aliases of relations such that person_a, person_b, *etc*. are aliases for person.

1. Describe how you would enhance the database schema (with additonal tables, columns, primary keys o   a monarch, and ensure that we record f   one as **succ_year**, and (ii) the nam

2. Write a query in each of the following languages that returns the scheme (name,born_in) containing the name and place of birth of all people known to ha   born in the same place as their mother.

   (a) RA

   (b) Datalog

3. Write a query in each of the following languages that returns the scheme (name) containing names of all people known to be parents.

   (a) RA

   (b) Datalog

4. Write a query in each of the following languages that returns the scheme (name) containing the names of all men not known to be fathers.

   (a) RA

   (b) Datalog

5. Write a query in each of the following languages returning the scheme (name) listing those people that have had at least one child of each gender that appears in the database.

(a) RA

Note that the above answer does make the assumption that each person can be just a father, or just a mother. If you want to capture the concept of a person sometimes being a mother, and sometimes being a father, you would need a union on the LHS of each division.

(b) Datalog

6. Suppose the following RA query q has been executed:

$$\pi_{\mathsf{person_b.name,person_a.father}} \, \sigma_{\mathsf{person_a.name=person_b.mother} \wedge \mathsf{isNotNull(person_a.father)}}(\mathsf{person_a} \times \mathsf{person_b})$$

(a) If the query has been executed at one point in time, since which Δ row have been inserted into person to give $\mathsf{person'} = \mathsf{person} \cup \Delta_\mathsf{p}$, give an RA query that returns any additional answers to the original query.

(b) If the enhancements requested in Question 1(a) have been added, how could your answer be improved?

7. Write an SQL query that returns the scheme (name, born_in) ordered by name containing the name and place of birth of all people known to have been born in the same place as their mother.

8. Write an SQL query that returns the scheme (name) ordered by name containing the name of all people known to be parents.

9. Write an SQL query that returns the scheme (name) ordered by name that lists parents for whom all known children are of the same sex.

10. Suppose you have to design a new database to hold the following information about the companies that do business with ACME Computing Ltd. These companies may be customers, suppliers, or both. For all companies we record their name and contact email address, and for companie ... umber.

For supplier ... ... and record a number of typ ... to each supplier all the stock items currently being supplied. It is company policy that each stock item may come from only one supplier, and some stock items are manufactur ... E Computing Ltd itself, and therefore have no supplier. Each stock item has a part ... er and description, and some stock items have their colour recorded.

For customers, we record the sales manager that deals with the customer, and a credit limit. We also record all orders made by the customer. Each order is on a particular date, is given an order number, and has a reference number given by the customer. An order may have any number of stock items, with the quantity and price for each stock item recorded.

When orders are sent, we record the tracking number of each parcel the order was sent in, and the date on which the parcel was sent. We also wish to record how much of each stock item was put in each parcel, in case the parcel gets lost and an insurance claim must be made. We record the delivery address for each order, from a record of delivery addresses, for which we record the customer name, postcode and address. We identify an address record by the combination of the customer name and postcode.

(a) Design an ER$^{\mathcal{ADHKLMNOSVW}}$ schema to represent this new database.

(b) Map the ER schema you designed in (i) into a relational schema.

11. Suppose that a relation $R(A, B, C, D, E, F, G, H)$ has the functional dependency set $S = \{A \to BE, AC \to G, AFG \to E, B \to ACG, CF \to D, D \to G, DEG \to FCD, G \to H\}$

(a) Compute a minimum cover S_c of S.

(b) Identify and justify all the candidate keys of R.

(c) Decompose the relation R into 3NF.

(d) Decompose the relation R into BCNF.

12. The following histories describe the sequence of operations performed respectively by four transactions T_1–T_4.

$H_1 = r_1[c_{CZ}], r_1[c_R], r_1[c_B], r_1[c_{GB}], c_1$
$H_2 = r_2[c_B], r_2[c_R], r_2[c_{CZ}], w_2[c_{CZ}], c_2$
$H_3 = r_3[c_B], w_3[c_B], r_3[c_{CZ}], w_3[c_{CZ}], c_3$
$H_4 = w_4[c_R], w_4[c_B], w_4[c_{GB}], c_4$

(a) Briefly explain if the following concurrent execution is serialisable and recoverable. If non-serialisable, explain what anomaly occurs.

$H_a = r_1[c_{CZ}], r_1[c_R], r_3[c_B], w_3[c_B], r_3[c_{CZ}], w_3[c_{CZ}],$
$\quad r_1[c_B], r_1[c_{GB}], c_1, c_3,$

(b) Briefly explain if the following concurrent execution is serialisable and recoverable. If non-serialisable, explain what anomaly occurs.

$H_b = r_3[c_B], w_4[c_R], w_4[c_B], w_4[c_{GB}], w_3[c_B],$
$\quad r_3[c_{CZ}], w_3[c_{CZ}], c_4, c_3$

(c) Briefly explain if the following concurrent execution is serialisable and recoverable. If non-serialisable, explain what anomaly occurs.

$H_c = r_1[c_{CZ}], w_4[c_R], r_1[c_R], w_4[c_B], r_1[c_B], w_4[c_{GB}], c_4, r_1[c_{GB}], c_1$

(d) Briefly explain which (if any) pair of the transactions taken from T_1–T_4 will be serialisable for all concurrent executions of the pair, and also briefly explain which (if any) pair of the transactions taken from T_1–T_4 will be recoverable for all concurrent executions of the pai

(e) Give a conc   uces a deadlock involving all fou   ck state.

13. The table below lists the contents of a database log, which keeps   DO records on a table country.

UNDO	$w_1[c_R, \text{population} = 148, 179, 000]$
UNDO	$w_4[c_{TR}, \text{population} = 62, 481, 123]$
UNDO	$w_2[c_{CH}, \text{population} = 7, 392, 444]$
UNDO	$w_2[c_{GB}, \text{population} = 58, 543, 111]$
UNDO	$w_3[c_{CH}, \text{population} = 7, 312, 222]$
UNDO	$w_5[c_{CH}, \text{population} = 7, 210, 000]$
LOG	c_4
UNDO	$w_3[c_B, \text{population} = 11, 020, 000]$
LOG	c_3

na			area	population
Czech Republic	CZ	Prague	78,703	10,321,120
Switzerland	CH	Bern	41,290	7,207,060
Russia	R	Moscow	17,075,200	148,178,487
Belgium	B	Brussels	30,510	10,170,241
Turkey	TR	Ankara	780,580	62,484,478
United Kingdom	GB	London	244,820	58,489,975
Egypt	ET	Cairo	1,001,450	63,575,107
			$\vdots$	

(a) If the country table has the contents illustrated above, describe the actions performed by the recovery procedure using the above log, and what population figures will be left after recovery.

(b) If an additional LOG record were added for a_5 to record the completion of aborting T_5, would your answer to (a) change, and if so, how does it change?

(c) Considering the time just after when c_4 occurs, describe and justify which updates from the above log must have been written to disc, which might have been written to disc, and which must not have been written to disc.