Database Systems, CSCI 4380-01
Homework # 3 Answers
Due Wednesday February 7, 2018 at 11:00:00 PM

Question 1 (20 points). You are given the following:

$\mathcal{F} = \{AB \rightarrow CD, AE \rightarrow G, GD \rightarrow H, HB \rightarrow ED, D \rightarrow B\}$ for relation $R(A, B, C, D, E, F, G, H)$

Use 3NF decomposition algorithm to find a set of relations that satisfy the third normal form.

For each resulting relation, list all functional dependencies projected over the relation. Also show whether the relation is in BCNF or not.

Answer.

This is a minimal cover. So, we can apply the 3NF decomposition.

Keys: ABEF, ABFH, ADEF, AGDF, ABGF and AFDH

Relations:

$R1(A, B, C, D)$, $\mathcal{F}_1 = \{AB \rightarrow CD, D \rightarrow B\}$

Keys: AB, AD. Not in BCNF because the second f.d. does not have a superkey on the left.

$R2(A, E, G)$, $\mathcal{F}_2 = \{AE \rightarrow G\}$

Key: AE, in 3NF and BCNF

$R3(G, D, H)$, $\mathcal{F}_2 = \{GD \rightarrow H\}$

Key: GD, in 3NF and BCNF

$R4(H, B, E, D)$, $\mathcal{F}_2 = \{HB \rightarrow ED, D \rightarrow B\}$

Key: HB, HD in 3NF and but not in BCNF

$R5(A, B, E, F)$, $\mathcal{F}_5 = \{\}$

Key: all attributes, in 3NF and BCNF

This last relation is generated for one of the keys. Any of the keys can be used to generate this relation.

Question 2 (30 points). You are given the following:

$\mathcal{F} = \{AB \to C, BD \to E, E \to F, AF \to G\}$ for relation $R(A, B, C, D, E, F, G)$

Use BCNF decomposition to find a set of relations that satisfy the Boyce-Codd normal form.

For each resulting relation, list all functional dependencies projected over the relation. Also show whether the resulting decomposition is functional dependency preserving or not.

Answer.

Key: ABD

All functional dependencies violate it. We can use any of them. My ordering is arbitrary.

Decompose using $AB \to C$ ($AB^+ = \{A, B\}$)

$R1(A, B, C)$ $\mathcal{F}_1 = \{AB \to C\}$. Key: AB, In BCNF.

$R2(A, B, D, E, F, G)$, $\mathcal{F}_2 = \{BD \to E, E \to F, AF \to G\}$.

Key: ABD. Still all f.d.s violate it.

Decompose using $BD \to E$, ($BD^+ = \{B, D, E, F\}$)

$R21(B, D, E, F)$ $\mathcal{F}_{21} = \{BD \to E, E \to F\}$

Key: BD, not in BCNF

$R22(A, B, D, G)$ $\mathcal{F}_{22} = \{ABD \to G\}$

Key: ABD, in BCNF.

We must decompute $R21$ further, using $E \to F$:

$R211(E, F)$ with $\mathcal{F}_{211} = \{ErightarrowF\}$

Key: E, in BCNF

$R212(B, D, E)$ with $\mathcal{F}_{212} = \{BDrightarrowE\}$

Key: BD, in BCNF

Final decomposition:

$R1(A, B, C)$ $\mathcal{F}_1 = \{AB \to C\}$
$R211(E, F)$ with $\mathcal{F}_{211} = \{ErightarrowF\}$
$R212(B, D, E)$ with $\mathcal{F}_{212} = \{BDrightarrowE\}$
$R22(A, B, D, G)$ $\mathcal{F}_{22} = \{ABD \to G\}$

Decomposition preserving? We can look at the union of the set of functional dependencies: $\mathcal{F}_{all} = \{AB \to C, ErightarrowF, BDrightarrowE, ABD \to G\}$. According to this:

$AF^+ = \{A, F\}$. Hence, $AF \to G$ is not implied by $\mathcal{F}_{all}$. As a result, this decomposition is not dependency preserving.

Question 3 (50 points). In this database, there are users who store their username, email, password, address. Furthermore, the database holds information about the different devices that they use, make, year, OS while exercising (phone, fitbit, watch, etc.). A person can have many such personal devices.

Users may also friend other users. Friendship is mutual, i.e. both sides agree. Friends are first requested by one side (only requests at this point) and if approved by the other side, it becomes stored friendships. Both requests and friendships must be stored.

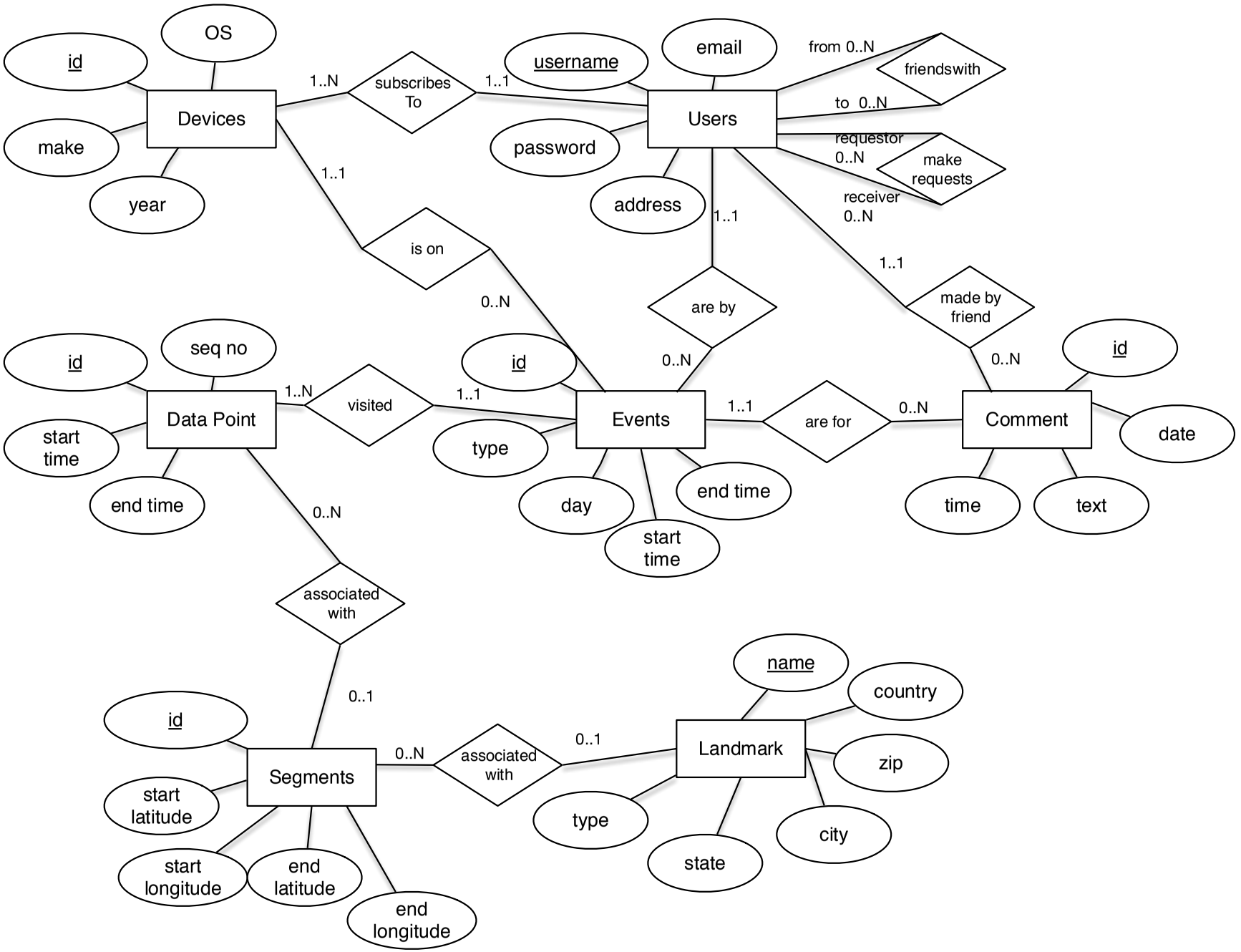
The database holds a (very large) number of road segments, each segment has a starting and end point, both are latitude and longitude values. Plus, each segment may have an associated geographic landmark. Each landmark is given by a name, type, state, city, zip and country attached. Landmarks may be associated with multiple segments.

Each user stores multiple events. Each event has an id, type (running or cycling), day, start time and end time, recording device and starting segment. The event also has a sequence of data points. Each data point has an associated segment, associated event, the exact times the start and end points of the segment were visited, and the sequence number in the event (an integer, for example 3 means this was the third data point in the segment). Note that the sequence number is needed because people can be running in a circle, hitting the same segment many times.

The events also have comments, made by the friends of the user. Each comment has the author, associated event, date, time and text.

We stand ready to analyze the database to see if it reveals these important secret locations!

Answer.


OS
id
Devices
make
year
1..N
subscribes To
1..1
username
email
password
address
Users
from 0..N
friendswith
to 0..N
requestor 0..N
make requests
receiver 0..N
1..1
is on
0..N
1..1
are by
0..N
1..1
made by friend
0..N
id
seq no
Data Point
start time
end time
1..N
visited
1..1
id
Events
type
day
start time
end time
1..1
are for
0..N
id
Comment
date
time
text
0..N
associated with
0..1
id
Segments
start latitude
start longitude
end latitude
end longitude
0..N
associated with
0..1
name
country
Landmark
zip
city
state
type