

Introduction to Database Systems

Problem Set 1

Due: January 19, 2018 at 11:59 PM

1 Database Background

(.1 pts each, .5 pts total)

1. What is an entity?
2. Describe the difference between a weak entity and an entity.
3. What is a schema? What does it do?
4. What are the properties of an ACID transaction? For each property, define it in your own words and explain why it is important.
5. What is a key?

2 E-R

Pat Fitzgerald currently keeps track of his former players who went pro using a binder. Inspired by the modernity of the new fieldhouse, he asks you to organize his data with a database. Fitz comes to you with the following requirements:

- Players have a name, player ID, and years played
 - Northwestern players play a position, this position does not change over the course of their career.
 - Within each position there are "squads". Per position, there is exactly one player per squad. For each position, there are multiple squads. For example, Bob is a Linebacker, First Squad; Mary is a Linebacker, Second Squad.
 - Each position has a name.
 - Players are each members of a single team.
 - Teams have team names, and are located in cities.
 - Teams must be associated with a city, a city may have multiple teams.
 - Teams have a single stadium, majority owner, and mascot.
 - Cities have a single population statistic, and a single calculated GDP.
- (a) List the entity sets and relationship sets associated with the Fitz's database. Do any of these sets need an additional attribute for a primary key identifier? (0.25 pt)
 - (b) Create an entity-relationship diagram for this system. Use at least one weak entity set. Clearly show the relationships as 1:1, 1:many, or many:many. Be sure to underline the primary key in each entity set. (1 pt)
 - (c) Translate your E-R diagram into a set of relations. (0.25 pt)

3 Normalization

3.1 The perils of not being normalized

Consider the following unnormalized schema. The pay of an employee is a function of the service time. The schema has functional dependencies $Fav. Record \rightarrow Fav Rec. Artist$, $Service Time \rightarrow Pay$. Consider the three operations: Update, Insert, and Delete. For each operation give an explicit example of potential anomalies caused by the unnormalized schema. (.2pts for each)

EMPLOYEE ID	FAV. RECORD	SERVICE TIME	PAY	FAV. REC. ARTIST
1	Rocky Mountain High	1 yr	\$10	John Denver
2	Back to the Country	5 years	\$15	Loretta Lynn
3	Jolene	3 years	\$12	Dolly Parton
4	Rocky Mountain High	2 yr	\$11	John Denver
5	22	1 years	\$10	Taylor Swift

3.2 True/False

Are the given set of relations $\{R\}$, and functional dependencies $\{F\}$ in BCNF?

1. $R(A, B, C), F = \{A \rightarrow B, B \rightarrow C\}$ (.2pts)
2. $R_1(A, B), R_2(B, C), F = \{A \rightarrow B, B \rightarrow C\}$ (.2pts)

3.3 BCNF

You want to keep track of your friends' favorite Hamilton song, as well as their location. The lead singer on an individual's favorite Hamilton song implies their favorite Hamilton artist. Consider the following schema:

FRIEND ID	FAV. HAMILTON TRACK	LOCATION	DISTANCE	FAV. HAMILTON ARTIST
1	Alexander Hamilton	Wisconsin	300 miles	Leslie Odom Jr.
2	You'll be Back	California	2200 Miles	Jonathan Groff
3	What'd I miss	Iowa	300 miles	Daveed Diggs
4	You'll be Back	Florida	1200 miles	Jonathan Groff
5	Helpless	California	2200 Miles	Phillipa Soo

- (a) Is the schema for this table in Boyce-Codd Normal Form? Explain why or why not. (.5 pts)
- (b) Decompose the schema into BCNF, if necessary. Show your work, including the functional dependency you use at each step of the decomposition. (.5 pts)