CS5200 Database Management Systems

Assignment 2

Name: Jayaraman Venkatesan *Due: 30 May 2023*

1 Normalize a Schema

1.1 Normalized diagram

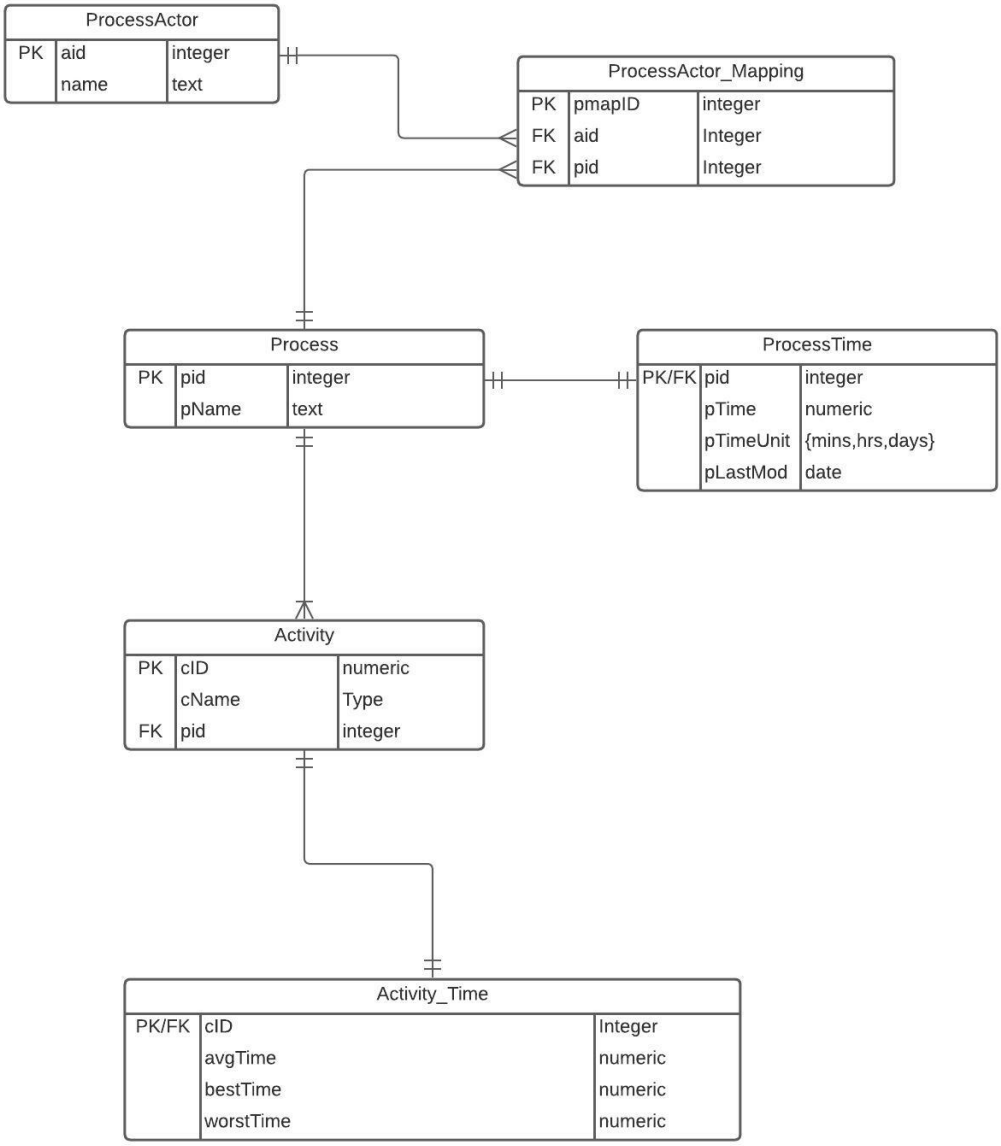

Figure 1: Normalized schema

1.2 Glossary

Entity	Attribute	Definition/Description	Type	Comments/Example
Process	pid	unique identifier	integer	artificial PK, e.g., 332
	pName	the name of the process	text	"Register for a Course"
ProcessTime	pid	Unique identifier / Foreign key	integer	both primary key and belongs to ID of some process. e.g., 332
	pTime	the average flow time for the process, expressed in "time units"	Numeric	120
	pTimeunit	the unit in which the flow time is expressed	value set	mins
	pLastMod	the last time the process has been updated or modified	date	12/09/2022
ProcessActor	aid	unique identifier	integer	artificial PK, e.g., 100
	name	name of the process actor	text	"Accounting", "Registrar"
ProcessActor_Mapping	pmapID	unique Identifier	integer	artificial PK, e.g., 1998
	pid	foreign key	integer	belongs to primary key of process. eg, 332
	aid	foreign key	integer	belongs to some actor e.g,100
Activity	cID	unique identifier (PK)	numeric	artificial PK, e.g., 99876
	cName	name of activity	text	"check metric status"
	pid	foreign key	integer	activity belongs to some process, e.g., 332
ActivityTime	cID	unique identifier / Foreign key	integer	both primary key and belongs to some activity. e.g., 99876
	avgTime	average execution time in secs	numeric	three point estimate. e.g., 28,32
	bestTime	best execution time is secs	numeric	three point estimate. e.g., 28,32
	worstTime	worst execution time in secs	numeric	three point estimate. e.g., 28,32

Figure 2: Glossary

1.3 Reasoning

1. Created *ProcessActor_Mapping* to resolve many-to-many relationships between Process and ProcessActor
2. Created table *ProcessTime* to remove the redundant candidate key information of process name and time being in same table
3. Created table *ActivityTime* to remove the redundant candidate key information of activity name and time being in same table

1.4 Proofs for normalized schema

1.4.1 1 NF

1. The Schema is in 1 NF because all the row in all the entities represents atomic values
2. Each row is unique and identitied by their primary key

1.4.2 2 NF

1. The Schema is in 1 NF
2. A single column in primary key is not funcitonally dependent on any other candidate key to uniquely represent the rows.
    - To satisfy this criteria created new ProcessTime and ActivityTime tables

1.4.3 3 NF

1. The schema is in 2NF
2. There are no transitive functional dependencies in the schema.
3. **PROOF :**
    (a) **Process** : $\{pid \rightarrow pname\}$ is the only functional dependency and hence it is free of any transitive dependencies to remove
    (b) **ProcessActor** : $\{aid \rightarrow name\}$ is the only functional dependency and hence it is free of any transitive dependencies to remove
    (c) **ProcessActor_Mapping** : $\{pmapID \rightarrow pid, aid\}$ is the only functional dependency and hence it is free of any transitive dependencies to remove
    (d) **ProcessTime** : $\{pid \rightarrow pTime, pTimeUnit, pLastMod\}$ is the only functional dependency and hence it is free of any transitive dependencies to remove
    (e) **Activity** : $\{cID \rightarrow cName, pid\}$ is the only functional dependency and hence it is free of any transitive dependencies to remove
    (f) **ActivityTime** : $\{cID \rightarrow avgTime, bestTime, worstTime\}$ is the only functional dependency and hence it is free of any transitive dependencies to remove

1.5 Collaborators :

Acknowledging collaborations with Dhanooram Nagaraj.
email id : nagaraj.d@northeastern.edu