Habib University
CS343: Graph Data Science
Spring 2024 - Class Activity

January 16, 2024

Designing a Graph Database Schema

The most important factor when designing a graph database is the act of defining its use-case. Different requirements can lead to different information structures, edges between two entities, and implementations.

The domain, described in the form of requirements written down by people who will be using the graph database, is the starting point for graph models. The underlying structure in the given requirements gives the graph database designer an idea about what are the key entities (nodes) that will be part of the model, how they will be categorized (labels), how they will be related to other entities in the model (relationships), and what data of the node or relationship should be stored (properties).

A good way to identify the labels for entities and the relationships between them is to identify the nouns and verbs from the requirements provided to you.

- Nouns represent nodes or labels.
- Verbs represent relationships.

After identifying these, the next task is to piece them together in the form of a data model.

Task

You are provided with a structure below that you are familiar with, a relational database. It contains information on the different courses taught to a student in the Computer Science program at Habib. This data is stored in tables, which consist of rows and columns. Your task is to convert this relational database into a graph database which consists of nodes, labels, properties, and relationships.

CourseNumber	CourseName	CreditHours	Type	DepartmentID	PreRequisiteNumber	FacultyLead	MinorID	SemesterNumber
CS101	Algorithmic Problem Solving	3	Foundation	1	1	Dr. Waqar Saleem	1	1
CS101L	Algorithmic Problem Solving Lab	1	Foundation	1	1	Dr. Waqar Saleem	1	1
CS102, CE171	Data Structures and Algorithms	3	Foundation	1	1	Dr. Qasim Pasta	1	2
CS102L, CE171L	Data Structures and Algorithms Lab	1	Foundation	1	1	Dr. Qasim Pasta	1	2
CS113, MATH113	Discrete Mathematics	3	Foundation	1, 3	1	Dr. Waqar Saleem	1, 2	2
CS355, CE373	Database Systems	3	Kernel	1, 2	2	Dr. Qasim Pasta	1	3
CS355L, CE373L	Database Systems Lab	1	Kernel	1, 2	3	Dr. Qasim Pasta	1	3
CS224, CE272	Object Oriented Programming	3	Kernel	1, 2	2	Nadia Nasir	1	2
CS224L, CE272L	Object Oriented Programming Lab	1	Kernel	1, 2	3	Nadia Nasir	1	2
CS201	Data Structures II	3	Kernel	1	2, 3, 4	Dr. Faisal Alvi	1	4
CS232, CE324	Operating Systems	3	Kernel	1, 2	2, 3, 5, 6	Dr. Muhammad Mobeen Movania	1	5
CS232L, CE324L	Operating Systems Lab	1	Kernel	1, 2	2, 3, 5, 6	Dr. Muhammad Mobeen Movania	1	5
CS212	Nature of Computation	3	Kernel	1	4	Dr. Waqar Saleem	1	5
CS412	Algorithms: Design and Analysis	3	Kernel	1	7, 8	Dr. Shah Jamal Alam	1	6

Table 1: A table with information of the courses taught in Computer Science at Habib. **Note: the data is not standardized in this relational database.**

DepartmentID	DepartmentName
1	Computer Science
2	Computer Engineering
3	Mathematics

Table 2: A table with information of relevant departments of the courses taught in Computer Science at Habib. **Note: the data is not standardized in this relational database.**

PreRequisiteNumber	PreRequisiteCourseName
1	None
2	Data Structures and Algorithms
3	Data Structures and Algorithms Lab
4	Discrete Mathematics
5	Computer Architecture
6	Computer Architecture Lab
7	Data Structures II
8	Probability and Statistics

Table 3: A table with information of the prerequisites of the courses taught in Computer Science at Habib. **Note: the data is not standardized in this relational database.**

MinorID	MinorName
1	Computer Science
2	Mathematics

Table 4: A table with information of minors that the courses taught in Computer Science at Habib are a part of. **Note: the data is not standardized in this relational database.**