

پایگاه داده تمرین اول

عباس یزدان مهر
۹۹۲۴۳۰۷۷

مهندسی کامپیوتر، دانشگاه شهید بهشتی

۲۶ مهر ۱۴۰۱

۱

به دلایل متفاوت یک فایل یک پایگاه داده محسوب نمی شود:

اطلاعات اضافه و ناپیوسته از جایی که برنامه نویس های متفاوت در طول زمان زیاد برنامه های متفاوت توسعه می دهند پس فایل هایی که درست می شود ساختارها و شکل های متفاوتی دارند. یک مشکل دیگر ناپیوسته بودن و تکراری بودن اطلاعات در فایل است چون ممکن است در مکان های متفاوتی به اطلاعات نیاز داشته باشیم و برای هر مکان باید یک فایل داشته باشیم پس برای تغییر یک اطلاعات کوچک باید تمام فایل هایی که در جاهای متفاوت داریم را بصورت تک به تک تغییر دهیم که خیلی کار سخت و بی فایده ای است و علاوه بر آن بدلیل تکراری بودن این اطلاعات داریم فضای حافظه ای زیادی از دست می دهیم.

سختی دسترسی به اطلاعات خیلی مواقع نیاز است که اطلاعاتی با ویژگی خاص را بدست آوریم مثلا لیستی از دانش آموزانی که معدل بالای ۱۷ دارند یا ... در ذخیره سازی فایل برای این کار باید برنامه نویس را برای هر نیاز خاص بکار بگیریم تا برنامه ای بنویسد که اطلاعات درخواستی را به ما بدهد و این اصلا جالب نیست پس باید برنامه ای ویژه داشته باشیم که این کار را صورت دهد.

ایزوله بودن اطلاعات وقتی اطلاعات را در فایل ذخیره می کنیم نوع نوشتن اطلاعات در فایل جزئی از اطلاعات فایل می شود به اصطلاح اطلاعات ایزوله نیست و این قضیه می تواند مشکلاتی ایجاد کند.

۲

۱.۲

DDL(Data-Defenition Language) : We specify a database schema by a set of definitions expressed by a special language called a data-definition language (DDL). The DDL is also used to specify additional properties of the data.

DML(Data-Manipulation Language) : A data-manipulation language (DML) is a language that enables users to access or manipulate data as organized by the appropriate data model. The types of access are: Retrieve, Delete, Insert, Modify(Update).

۲.۲

DDL: CREATE TABLE students (name char(20), age numeric(2, 0));

DML: INSERT INTO students(name, age) VALUES ("ali", 20);

DML: SELECT * FROM students;

۳

استقلال فیزیکی داده به توانایی تغییر دادن ساختار پایین ترین لایه پایگاه داده بدون تاثیر گذاری روی لایه های بالاتر شماتیک گفته می شود.(هر پایگاه داده ۳ سطح معماری طرحواره دارد: سطح فیزیکی، سطح منطقی و سطح نما.

دلیل اینکه ما بخواهیم در سطح فیزیکی داده را تغییر بدهیم این است که ممکن است بخواهیم فایل ها یا فهرست هایی را اضافه یا کم کنیم تا عملکرد سیستم پایگاه داده را افزایش داده و سریع تر کنیم و این فقط وقتی ممکن است که استقلال فیزیکی داده برقرار باشد.

۴

- a. این دستور ابتدا تمام take هایی که سال آنها از ۲۰۰۹ بیشتر مساوی است را انتخاب می کند و سپس با جدول دانشجو پیوند می دهد. در واقع تمام دانشجو هایی را نشان میدهد که پس از سال ۲۰۰۸ در دوره ای شرکت کرده اند.
- b. این دستور ابتدا جداول *take, student* را باهم پیوند می دهد و سپس از بین آنها تمام سطر هایی که سال آنها بیشتر مساوی ۲۰۰۹ است را انتخاب می کند. در واقع خروجی با مثال a فرقی نمی کند.
- c. این دستور ابتدا دو جدول *takes, student* را پیوند می دهد و سپس ستون های نوشته شده $(ID, name, course_id)$ را انتخاب می کند.

۵

خیر - چون ممکن است دو *advisor* دقیقا مقادیر یکسان داشته باشند آنگاه دیگر *s_id* کلید کاندیدا نیست پس کلید اصلی هم نیست. کلید کاندیدای جدید مجموعه (s_id, i_id) است که تنها انتخاب برای کلید اصلی است.

۶

۱. برای مقادیری که وجود دارند ولی نامشخص هستند به null نیاز داریم.
۲. برای مقادیری که اصلا وجود ندارند به null نیاز داریم.

پایان!