



# 数据库系统原理课程

## 第二次作业

学院：	计算机与信息技术学院
专业：	计算机科学与技术
班级：	计科 1804
姓名：	王子龙
学号：	18281218
教师：	徐薇

2020 年 9 月 24 日

1. 尝试标识该业务场景中的数据模式，即列出该业务涉及到的所有实体型(包括实体名及属性(可以适当简化)。可以采用如下形式列出：

实体型名 1(属性 1, ..., 属性 n) 尝试在你选定的数据库平台中建立实体型所对应的表，并想办法往其中插入数据，验证你所设计的模型的合理性。

12306 业务涉及道德所有实体型：

用户（用户名、登录密码、姓名、手机号、证件号、证件类型、旅客类型）

列车（车次、出发站点、到达站点、历时、是否为高铁/动车）

途径站（车次，站点 1，到达时间 1，出发时间 1，停留时间 1，. . . . . 站点 38，到达时间 38，出发时间 38，停留时间 38）

余票（车次，出发站，到达站，商务座特等座余票，商务座特等座票价，一等座余票，二等座余票，基础票价）

注：下划线为主键、斜体为外键

随后在四个表中插入数据：

username	password	name	phoneNumber	emailAddress	ID	Type	IDType
wang	wangzilong2020	王子龙	+8617610100201	wangzilong@bjtu.edu.cn	110112200002014610	0	0
wang2	wangzilong2020	Zilong Wang	+8617610100201	wangzilong@bjtu.edu.cn	EF7845586	0	1

图 1 用户信息数据

在 12306 中，用户需登录购票，登录可选择使用证件号码登录，同时考虑到 12306 有多种购票方式：

- a. 购票类型不同（儿童、学生...）
- b. 购票证件不同（身份证、永久居留证、护照...）

如在我插入的两个数据中，虽然是同一人的信息，但是因为主键（证件号）不同，分别为身份证号和护照号，因此导致了王子龙一人拥有两行数据。

TrainNumber	DepartureStation	ArrivalStation	Duration	isGaotie
1 C2551	BJN	BH	56	1
2 G1	BJN	SHHQ	268	1
3 G666	XAB	BJX	329	1

图 2 列车信息数据

在 12306 中，高铁买票查询是需要考虑是否为高铁/动车，因此加入了一个属性 isGaotie。不同班次列车之间最主要的不同就是列车号不同，因此选用 TrainNumber 列车号作为主键。同时每班高铁还需记录始发站和到达站以及总时间，因为在 12306 中查询列车需要得知这三个属性。

车次	出发站 到达站	出发时间 到达时间	历时
G415	北京南 上海	06:19 12:23	06:04 当日到达

TrainNumber	Station1	ArrivalTime1	DepartureTime1	Station2	ArrivalTime2	DepartureTime2	Station3	Arriv
1 C2551	BJN	2020-10-04 05:47	2020-10-04 06:02:00	TJ	2020-10-04 06:32	2020-10-04 06:36:00	BH	2020-10
2 G1	BJN	2020-10-04 08:45	2020-10-04 09:08:00	JNN	2020-10-04 10:22	2020-10-04 10:24:00	NJN	2020-10
3 G666	XAB	2020-10-04 15:45	2020-10-04 16:08:00	WNB	2020-10-04 16:17	2020-10-04 16:19:00	XMXN	2020-10

图 3 途径站信息数据

每次列车都经过数个途径站，同时途径站具有抵达时间和发车时间两个时间，因此此表格再次将 TrainNumber 作为主键作为列车之间的区分，同时前两个车站是 Not Null 的因为一辆车势必会经过至少一个始发站和一个终点站。此外还预留了 38 个车站。

	Tr...	...	A...	SpecialClassRemaining	FirstClassRemaining	SecondClassRemaining	BaseFare
1	G1	BJN	JNX	0	2	12	184.5
2	G1	BJN	NJN	1	1	10	443.5
3	G1	BJN	SHHQ	10	20	200	553
4	G1	JNX	NJN	0	0	2	279
5	G1	JNX	SHHQ	0	2	10	398.5
6	G1	NJN	SHHQ	1	2	100	134.5

图 4 余票信息数据

每次列车一般情况下其每个途径站到之后任一个后序的途径站之间均出售车票，因此在设计余票 Table 的时候，要考虑到讲列车号 TrainNumber、上车站 DepartureStation、抵达站 ArrivalStation 共同设为主键。同时还要记录两个车站区间的不同等级舱位的余票情况，并记录该区间之间的不同舱位的票价（一等座票价，以及商务座（特等座）票价）。

在上述表中，车站名均为中文车站名的缩写（如：北京南→BJN）。

通过验证查询，发现设计模型工作正常、合理。

2. 请自学了解 Axure RP 等快速原型设计工具的概念及功能，并尝试下载一个试用版。请你充当产品经理和产品界面模型设计工程师，模仿或创新设计两个用户界面(UI、用户视图或外模式的子模式)。请提交原始设计文档以证明你做了此项工作。

打开 12306，可以发现该网站的主要功能为查询车票，界面如下：

出发地

北京

到达地

上海

出发日期

2020-10-03

学生

高铁/动车

查 询

此时通过 Axure RP 即可创建一个类似的界面（UI1.rp）。

出发车站

到达车站

出发时间

学生票

动车/高铁

立即查询余票

三个文字通过三个 Box 来实现，而选择区即可设置为一个 TextField，同时加入学生票以及动车、高铁的选项。最后加入一个 Button “立即查询余票” 来实现 12306 的查询。

此外，当查询了某一天的车票之后，就会进入如下界面：

10-03	10-04 周日	10-05	10-06	10-07	10-08	10-09	10-10	10-11	10-12	10-13	10-14	10-15	10-16	10-17	10-18	10-19	10-20	10-21	10-22
车次类型: <input checked="" type="checkbox"/> G-高铁/城际 <input type="checkbox"/> D-动车 <input type="checkbox"/> Z-直达 <input type="checkbox"/> T-特快 <input type="checkbox"/> K-快速 <input type="checkbox"/> 其他																			
出发车站: <input checked="" type="checkbox"/> 北京南 <input type="checkbox"/> 北京		发车时间: 00:00--24:00																	
北京 --> 上海 (10月4日 周日) 共计51个车次 您可使用快捷换乘功能，查询途中换乘一次的部分列车余票情况。																			
<input type="checkbox"/> 显示积分兑换车次 <input type="checkbox"/> 显示全部可预订车次																			
车次	出发站 到达站	出发时间 到达时间	历时	商务座 特等座	一等座	二等座 二等包座	高铁 软卧	软卧 一等卧	动卧	硬卧 二等卧	软座	硬座	无座	其他	备注				
G415	北京南 上海	06:19 12:23	06:04 当日到达	5	候补	有	--	--	--	--	--	--	--	--	预订				
G101	北京南 上海虹桥	06:36 12:40	06:04 当日到达	10	有	有	--	--	--	--	--	--	--	--	预订				
G103	北京南 上海虹桥	06:43 12:48	06:05 当日到达	10	10	19	--	--	--	--	--	--	--	--	预订				
	北京南	07:00	04:40																

通过 Axure 设计类似界面（UI2.rp），通过 CheckBox，Table 以及 Label 实现了该界面：

车次类型: ☒全部 ☐GC-高铁/城际 ☐D-动车 ☐Z-直达 ☐T-特快 ☐K-快速 ☐其他

出发车站: ☒全部 ☐北京南 ☐北京

北京 -> 上海 (10月4日 周日) 共计51个车次

车次	出发站	到达站	出发时间	到达时间	历时	商务座	一等座	二等座
G415	北京南	上海	06:19	12:23	06:04	5	候补	有
G101	北京南	上海虹桥	06:36	12:40	06:04	5	有	有

3. 请分析你所设计的用户界面所对应的外模式子集，以及该子集与前面的模式间的映射关系，思考并叙述数据库系统从内模式、模式、外模式并最终到用户界面的双向映射(mapping)过程，分析过程中可能会涉及到的数据加工环节。分别指出一般需要由程序员写程序实现的环节和由数据库管理系统实现的环节。

余票查询涉及到的查询条件主要是时间及始发站，终点站。  
查询结果主要是符合筛选日期和始发站、终点站条件的车票余量、始发时间、到达时间、路程时长、途径车站等数据。

外模式与模式之间的映射主要保证了在数据库中增加实体关系或改变实体属性时，不必要改变应用程序与外模式间的调用关系，只需要改变模式与外模式之间的映射即可。  
用户应用视图根据外模式进行数据操作，通过外模式—模式映射，定义和建立某个外模式与模式间的对应关系，将外模式与模式联系起来，当模式发生改变时，只要改变其映射，就可以使外模式保持不变，对应的应用程序也可保持不变；另一方面，通过模式—内模式映射，定义建立数据的逻辑结构(模式)与存储结构(内模式)间的对应关系，当数据的存储结构发生变化时，只需改变模式—内模式映射，就能保持模式不变，因此应用程序也可以保持不变。

当数据库模式发生变化时，例如关系数据库系统中增加新的关系、改变关系的属性数据类型等，可以调整外模式/模式间的映像关系，保证面向用户的各个外模式不变。应用程序是依据数据的外模式编写的，从而应用程序不必修改，保证了数据与应用程序的逻辑独立性。当数据库中数据物理存储结构改变时，即内模式发生变化，例如定义和选用了另一种存储结构，可以调整模式/内模式映像关系，保持数据库模式不变，从而使数据库系统的外模式和各个应用程序不必随之改变。这样就保证了数据库中数据与应用程序间的物理独立性。

在这一过程中，涉及到许多的加工过程。  
从用户输入搜索条件到查询结果需要一个从内模式-外模式的一个两级映射过程。用户输入的搜索条件经过映射过程到了模式中，经过数据管理系统的加工处理再映射到内模式中，在内模式中进行数据的逻辑分析比较，比较完后，内模式就会把结果通过两级映射反馈回用户界面，从而产生一系列的响应。这反馈过程需要数据管理系统和物理配件的紧密配合。

主要需要程序员完成外模式与用户界面的设计以及概念模式表的建立。  
数据库系统能够帮助程序员完成内模式包括：数据的物理存储、方便的建立索引、按照模式提供增删改查等操。

4. 请尝试学习掌握数据库系统数据规模(条或记录数)估算方法，估计所选业务场景的数据规模，针对这种规模，请说明这种规模的数据可能对内模式的实现或表示存在什么要求。

数据库的规模，需考虑综合指标，即数据库对象的数量和属性、数据库对象所占用的物理空间等诸多因素。因为“表”是数据库的核心和基础，通过表估算数据库规模能很好地代表其它数据库对象（视图、索引、角色等）。

不同类型的表对数据库规模的影响不同。在字典表、操作表、报表表中，操作表对数据库规模的影响应该最大。表结构对数据库规模也有影响，主要体现在属性数量和属性数据类型上。属性数量越多，数据库规模越大；数据类型越复杂，数据库规模越大。表中元组的数

量对数据库规模也有影响。表中元组越多，数据库规模越大。

因此结合这三个因素，可知，对于 12306 余票查询系统来说，客户的查询，订票业务和客户的数量有关，我国人口众多，出行需求大，用户数量以万亿计，据统计，2019 年全国铁路完成旅客发送量 36.60 亿人，比上年增长 8.4%，完成旅客周转量 14706.64 亿人公里[1]。面对如此庞大的客运需求，我们需要内模式存储的优化，使得查询效率提高。

对于 12306 余票系统来说，最严峻的考验莫过于春运，从国家统计局的数据中简单统计得：2020 年春运持续 40 天，单日最高发送旅客 1300 万，总共发送 4.1 亿人次，日均售票 1320 万张，日均 4787 对旅客列车，而中国约有 2100 个左右客运站。

这种庞大的数据体量和查询的实时性需求导致内模式需要采取很多易于理解，结构稳定，减少表连接的代价，提高查询效率和保证安全性的策略，如存储方式、数据结构、组织索引等，对于用户的敏感信息还需要加密存储等。

[1]2019 年中国营业性客运量及周转量分析, [www.chyxx.com/industry/202005/864938.html](http://www.chyxx.com/industry/202005/864938.html)