

# 第5章 数据检索与宣询文件

♀?知识导入

#### 大海捞针——高致的数据检索

笛ぃ音知辺 小南选修了一门通识课——信息与情报检索。在课堂上,老师让小南用网络搜索引擎



点拼图

检索了几个关键词,分别统计了它们的检索时长。小南平时经常使用网络检索功能,却从来 没想过,在浩如烟海的网络数据中,如何才能在零点几秒的时间内,从整个 Internet 世界中 找到那几个关键的字眼呢? 小南隐隐约约地觉得,一定是谷歌这种搜索引擎在数据库的检 索上使用了特殊的方法。

用户建立数据库的目的就是为了存储和提取信息,信息提取的关键在于方便快速地查 询数据。因此,查询便成了数据库操作的主要内容。除了直接的查询操作外,对数据的追 加、更新、删除等操作也常常要首先找到需要处理的数据,所以,这些操作也通常以查询为基 础。本章将介绍如何使用 Access 提供的查询工具检索数据。

第5章动 画导入:小 南的"大海 5.1 - 掛针

## 数据检索方法

### 5.1.1 检索时长是否与 N 无关

数据检索——把数据库中存储的数据根据用户的需求提取出来,数据检索的结果会生 成一个数据表,既可以放回数据库,也可以作为进一步处理的对象。

在人类的常识中,数据总量 N 越大,要从中检索到需要的信息,需要的时间越长。检索 的时间复杂度 O(N),永远是一个以 N 为自变量的函数。

如果  $N \to \infty$ ,那么  $O(N) \to \infty$ 

但如果是这样的话,要在 Google 中检索一个关键词,检索范围是整个 Internet,那么检 索的时长或许将是无穷无尽的。这显然让人无法忍受。因此我们要想办法,让检索的时间 复杂度 O(N)等于一个与 N 无关的常量。

### 1. 顺序检索

一个无序数据列 N 中,找到目标数据 66。如果进行顺序检索,需要将 N 个数据一一查找,才能保证找到所有的 66,如图 5-1 所示。



O(N) = N如果  $N \rightarrow \infty$ ,那么  $O(N) \rightarrow \infty$ 

#### 2. 折半检索

对于一个**有序**序列,可以采用折半查找的检索策略,每次从数据列的中间位置开始比 对,比较中间值与目标值的大小,就可以一次排除掉一半数据,再用剩下的一半数据进行折 半检索。这也是经典算法分治算法的典型思想,如图 5-2 所示。





这种检索方法大大降低了时间复杂度,尤其是 N 值很大时,这种降低愈加明显。但遗憾的是,当 N 趋于无穷时, $\log_2 N$  仍然趋于无穷。

3. 索引检索

**索引**(Index)是对数据库表中一列或多列的值进行排序的一种结构,使用索引可快速访问数据库表中的特定信息。

我们在设计数据库的物理结构时,曾为某些数据字段设置过索引。主键索引是无重复 索引,也有些其他字段是有重复索引。索引就能将检索的时间复杂度 O(N)变成一个与 N 无关的常量。

在表中使用索引就如同在书中使用目录一样:要想查找某些特定的数据,先在索引中 查找数据的位置。为一个字段创建的索引会生成一个索引序列,这个序列存储在数据库的 索引文件中。一个字段的索引序列在逻辑上可以表示为一个有两列的表,分别存储索引内 容和地址指针,如图 5-3 所示。

			系号	学号	姓名	性别		照片
			01	1901011	李晓明	男	•••	
·			02	1901012	王民	男		
索引内容	指针		01	1901013	马玉红	女		
01		$\langle$	03					
02		$\langle \rangle$	04					
			01					
06			02	•••				
LI		'		•••				
			06	1901032	崔一楠	女		

图 5-3 设置索引和索引文件

当我们要检索 01 号系的学生时,不用直接去检索学生表,只要找到系号字段的索引序 列就行了。可以按照索引中 01 系号后面的地址指针,直接找到这个系学生在数据硬盘中的 位置。要知道,一个真正运行的商业数据库往往存储了海量的数据,经常整个内存都放不下 一张表,动辄就对数据表检索是非常浪费时间和空间的,不亚于大海捞针,而索引文件往往 都非常小,并且索引是一个有序序列,可以使用折半查找,检索起来很容易。

> O(N) = 1 $u \not\equiv N \rightarrow \infty, O(N) \rightarrow 1$



扩展阅读 5.1.1: 二叉树查找算法



扩展阅读 5.1.1: 哈希查找算法



扩展阅读 5.1.1: 海量数据查找算法



扩展阅读 5.1.1: 数据库多级索引

# ②思考

如果数据量实在太大呢?就连数据库的目录——索引文件都大到无法装入内存?那就 为索引文件建立索引,称为二级索引,甚至三级索引,四级索引……索引文件有几级,那就查 找几次,即便 N 趋于无穷,索引文件的级数也仅仅是一个常量罢了。

### 5.1.2 索引不是万能的

#### 1. 索引的优势

(1) 通过创建无重复索引,可以保证数据库表中每一行数据的唯一性。

(2) 大大加快数据的检索速度。

(3) 可以加速表和表之间的连接,特别是在实现数据的参照完整性方面特别有意义。

#### 2. 索引的劣势

(1) 创建索引和维护索引要耗费时间,这种时间随着数据量的增加而增加。

(2) 索引需要占物理空间,当数据量不大时,谈论索引是毫无意义的。

(3) 对表中的数据进行增加、删除和修改时,索引也要动态维护,降低操作速度。

数据量小时建立索引毫无必要,为所有字段建立索引也不明智,通常应该为主键、外键 和经常需要搜索的列创建索引。

## 5.2 Access 的数据检索

👷 问题导入

使用 Access 提供的筛选器、选择、高级筛选等工具可以对表格中的特定数据进行筛选。 那是不是就不需要其他的检索方式了呢?

虽然使用这些工具可以找到所需要的记录,但最主要的问题是,筛选是一种实时的屏幕 交互式操作,随着筛选结果的输出,筛选的操作也就结束了,整个过程不能以文件的形式存 储在数据库磁盘上,如果下次要执行同样的或类似的筛选任务,只能将这一操作过程一一重 现。如果需要反复查找大量数据,就要不断重复筛选操作,这显然非常烦琐而浪费,并且无 法将其开发成应用程序发布给用户使用。

为了便于用户在数据库中检索自己需要的数据, Access 提供了一种能以文件形式存储 的检索工具——查询文件。

#### 5.2.1 什么是查询文件

所谓查询就是找到用户所需数据库子集的过程。Access 根据用户定义的查询条件,在

91

#### 数据库与信息系统设计(慕课版)-

数据库中的一张或多张表中检索出满足条件的一组记录,这些记录只显示用户指定的所需 字段。因此,用户建立查询时可以定义需要显示的字段及筛选条件,当运行查询时,只有那 些指定的字段和符合筛选条件的记录才被检索出来。Access 在磁盘上建立一个查询文件 来存储这些检索需求。不论何时运行这个查询文件,Access 都根据文件中保存的需求将相 关数据组合起来建立一个动态数据集,也就是查询结果。这个动态数据集看起来像一张表, 但它不是真正的表,不存储在数据库磁盘上,只在内存中临时存储和显示。当用户关闭这个 动态数据集后,内存中的存储就可以清除了。因为这个动态数据集(查询结果)来源于数据 库中的数据,当数据源发生改变后,再运行查询文件,查询结果就发生改变;反过来,当用户 修改查询结果中的数据,查询结果从内存写回数据库磁盘时,同样也会改变数据源。

实际上,用户在前台运行了查询文件后,DBMS 自动在后台按照查询文件中的查询要求生成一条查询命令,该命令用数据库标准语言 SQL 写成,再通过执行这条 SQL 命令来实现查询操作。

Access 中的查询一旦生成,可以作为窗体、报表,甚至是另一个查询的数据源。数据库 查询过程如图 5-4 所示。



#### 5.2.2 查询文件的分类

为了完成上述查询功能, Access 主要提供了 两种查询方式:一种是屏幕操作方式,通过建立查 询文件的可视化方法存储查询条件;另一种是程 序方式,通过直接书写 SQL 命令的方式实现查询, 如图 5-5 所示。

本章着重介绍第一种方式,即查询文件。SQL 的语法和使用将在第6章介绍。每一个查询文件 都能转换成 SQL 命令来编译执行,但并不是所有 直接书写的 SQL 命令都能用查询文件显示。

Access 的查询文件有多种形式,包括选择查询、



92

第5章 数据检索与查询文件 -

参数查询、操作查询(动作查询)、交叉表查询、重复项查询、不匹配查询等。总结起来有4大 类:选择查询、参数查询、操作查询和特殊用途查询。具体的分类和功能说明如表5-1所示。

查询类型	查询方式	功 能 说 明	举 例
选择查询	选择查询	最基本的查询方式,指定记录和字段并 对查询结果排序、分组、统计汇总	查询学生表中男生的学号和 姓名,并按姓名升序排列
参数查询	参数查询	执行查询时提供参数的输入接口,实现 用户交互式查询,本质上也是选择查询	按用户输入的系号查找学生 信息
	生成表查询	查询结果生成一张新的基本表	用学生表中的男生记录生成 新数据表学生1
操作查询	追加查询	将查询结果插入至另一张基本表	将学生表中的所有女生记录 插入学生1表
	更新查询	对查询结果进行更新,存入原基本表	给 102 号课成绩 80 分以下的 学生每人加 5 分
	删除查询	将查询结果从数据源中删除	删除大学英语的选课记录
特殊用途查询	交叉表查询	用交叉表的形式组织查询结果,本质上 也是一种选择查询	横纵字段名分别为学号和课 程号,交叉位置显示成绩
	重复项查询	查找指定字段的重复项	统计每个系的人数
	不匹配项查询	在一张表中查询和另一张表不相关的 记录	查找没有人选修的课程

表 5-1 Access 查询文件类型的分类和功能说明

## 🖉 注 意

Access 提供了两种方法建立查询:一种是使用"查询向导";另一种是使用"查询设计" 视图(查询设计器)建立查询。查询向导可以按照一定的模式引领用户创建查询,实现基本 的查询操作,不需要使用者具备过多的数据库查询知识,最简单易行。但向导的功能比较单 一,要想完成丰富多变的查询任务,必须使用"查询设计"视图。

## 5.2.3 查询的视图

Access 的查询文件主要提供了三种模式的视图:设计视图、SQL 视图、数据表视图。

- 设计视图是创建一个查询文件的主要方式,它提供了完善的设计手段去实现各种条件设置。
- SQL视图是输入 SQL语句实现查询的界面,仅提供了命令文本的录入功能。
- 数据表视图是查询结果的显示界面,无论是用设计视图还是 SQL,查询的结果都以 二维表的形式在数据表视图中显示。

视图的切换如图 5-6 所示。

数据库与信息系统设计(慕课版)



图 5-6 视图切换

## 5.3 选择查询

选择查询是一种最基本的查询方式,其功能包括指定记录和字段的查询条件和对查询 结果的排序、分组、统计汇总。

#### 5.3.1 利用向导创建简单查询

【例 5-1】 利用简单查询向导创建查询:学生选课成绩单。

建立"查询向导",可以看到如图 5-7(a)所示的几种向导。其中,使用第一种"简单查询 向导"即可创建一个选择查询文件。在"表/查询"下拉列表框中选择数据源,如果所需字段 来自多个表,可以先后选定多张表格来一一添加。本例最终选取的字段是系名、学号、姓名、 性别、课程名、学分和成绩,分别来源于四张表,如图 5-7(b)所示。

单击"下一步"按钮,确定采用明细查询还是汇总查询。如果选择明细查询,将查出所选数据源中所有满足条件的记录,并且不对查询结果做统计汇总,如图 5-7(c)所示。

最后一步,为查询指定标题,在标题文本框中输入"学生选课成绩单",如图 5-7(d)所示。单击"完成"按钮,可以看到查询结果,如图 5-7(e)所示。

如果要查询学生成绩的统计汇总,如查询学生选课的总学分和平均成绩。首先应该在 字段选取的阶段去掉"课程名"字段,因为汇总成绩是要考察每个学生考试的总体情况,不关

#### 第5章 数据检索与宣词文件 -



系之 .	1 114					
AL H	学号 •	姓名 •	性别・	课程名 -	学分 🔸	成绩 •
信息系	1901011	李晓明	男	高等数学	5	95
信息系	1901011	李晓明	男	大学英语	3	70
信息系	1901011	李晓明	男	数据库应用	3	82
人力资源系	1901012	王民	男	大学英语	3	88
人力资源系	1901012	王民	男	数据库应用	3	85
人力资源系	1901012	王民	男	邓小平理论	2	81
信息系	1901013	马玉红	女	第二外语	2	85
国际经济与贸易	1901014	王海	男	软件基础	8	90
国际经济与贸易	1901014	王海	男	高等数学	5	88
计算机技术与科学	1901015	李建中	男	第二外语	2	80
计算机技术与科学	1901015	李建中	男	软件基础	8	80
信息系	1901016	田爱华	女	高等数学	5	83
信息系	1901016	田爱华	女	数据库应用	3	98
信息系	1901016	田爱华	女	第二外语	2	75
人力资源系	1901017	马萍	女	大学英语	3	85
人力资源系	1901017	马萍	女	数据库应用	3	89
国际经济与贸易	1901018	王刚	男	邓小平理论	2	92
国际经济与贸易	1901018	王刚	男	第二外语	2	81
计算机技术与科学	1901019	刘伟	男	软件基础	8	85
计管机技术与科学	1901019	刘佳	里	對捏度应用	3	86



心每门课程的细节,如图 5-8(a)所示。接下来在如图 5-7(c)所示界面中选择"汇总",单击下 方的"汇总选项"按钮,出现如图 5-8(b)所示的界面。给学分字段选择"汇总",即求和;给成 绩字段选择"平均",即求均值。最后查询的结果如图 5-8(c)所示。



图 5-8 汇总查询向导

可以看出,用向导建立选择查询确实非常简单,但建立的查询文件的形式也极其单一。 如果要想为查询结果按照考试平均分排序,或者只想查询计算机系或中文系的学生成绩,诸 如此类的查询要求,在以上的向导中都没有体现。因此,向导的方式只适合初步建立查询文 件,其余的设置要在接下来的查询设计视图模式中完成。下面将介绍使用查询文件设计视 图创建选择查询。

## 5.3.2 利用设计视图创建选择查询

查询设计视图可以独立地创建查询文件,也可以对向导创建的查询文件进行修改。

#### 1. 查询设计视图界面

查询设计视图由上下两部分组成,上半部分是数据源区,显示查询数据源使用的数据表 和关联关系,下半部分是设计网格区,负责设计查询的主要内容,如图 5-9 所示。

设计网格默认有以下几个主要内容。

- 字段: 查询设计中所使用的字段名,从数据源区的表中选取。
- 表:说明上方对应的该字段来自哪个数据源表。

第5章 数据检索与宣词文件 -

	4	学生			
		*			
		系号 『学号 姓名 性别			
		山土山知	-		
m)		<u></u>		-	
www. 字段: 表:					
字段:       排序:       显示:       条件:	•				

图 5-9 查询设计视图界面

- 排序:查询结果是否按该字段排序,如果排序,是升序还是降序。
- 显示:该字段在查询结果表中是否显示。
- 条件:限定该字段的查询条件。
- 或: 当查询条件多于一个,且多个条件之间采用逻辑或运算时,将用到该网格。

在设计网格区域的过程中,将用到的工具按钮会显示在 Access 上方的查询工具选项卡中,如图 5-10 所示。

	5	· c · ·			查询工具		Databa	ase : 数据库- C:\Use	ers\DELL\D	esktop\Database.a	ccdb (Access 2007 -	2016 文件格式) - Access
文件	形	治 创建	外部数据	B 数据库工具	1. 设计 🖓 🕄							
SQL 视图	<b>!</b> 运行	选择 生成:	<b>十二</b> 表 追加	ビリン (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)	● 传递 ● 传递 単数据定义	显示表	<ul> <li>         • 插入行         <ul> <li>                  ● 揃除行</li></ul></li></ul>	*** 插入列 ※ 删除列 2回返回: All	<b>、</b> 汇总	[?] 目 屋性表 答数		
结理	果			查询类型			查道	的设置		显示/隐藏		

图 5-10 查询工具选项卡

具体的按钮和功能如表 5-2 所示。

表 5-2 查询工具选项卡具体的按钮和功能

工具类别	按 钮	功 能 说 明					
结果	视图	弹出"视图"菜单,选择数据表视图、SQL视图、设计视图三种模式					
	运行	运行设计好的查询文件,单击此按钮查看结果					
查询类型	选择/生成表/追加	查询文件类型的选择,默认共有选择、生成表、追加、更新、交叉表、 删除、联合、传递、数据定义9种					
查询设置	显示表	数据源选择窗口,可选择基本表或已有的查询作为数据源					
	生成器	弹出"表达式生成器"对话框,在这里可以选择 Access 内置函数、数据库对象、常量、操作符、通用表达式等内容,当用户不熟悉 Access 表达式的写法时,可以用该生成器辅助生成表达式,减少语法错误的概率					
	插入行/删除行	在设计网格的最下方增加或删除一行					
	插入列/删除列	在设计网格的最右侧增加或删除一列					
	返回	指定查询结果返回多少条数据,可以按照行数或是百分比指定					

97