

专业实践总结

信息技术学院 王珊珊

我参加企业实践的龙创科技股份有限公司主要业务范围是智能数据分析、业务软件研发、人才培训，面对的客户主要是各大培训考试机构。研发测试人员外驻在培训公司所在地进行定制研发和测试。龙创公司虽然规模很小，但每年的营收也在一千万元以上。



【企业和项目情况概览】

龙创集团内部的保密控制的是非常严格的,对于非某个部门的员工需要进入办公区,必须经过电子流程审批,而且需要该部门员工亲自来门岗接送。公司招收新员工的标准也比较高。招聘岗位分为部门岗和工人岗。部门岗包括市场、设计、研发这些,招收员工的标准都比较高,一般需要 211 院校本科以上学历。

这次龙创公司为 A 培训公司所做的项目编号为 V439 的考试系统,由几个公司一起合作完成。由于我之前知道过学生进行大数据竞赛,龙创公司几个同事和我之前也有过共事,所以上手并没有什么难度。这次企业实践我还是希望能够了解到项目当中方方面面的开发内容和各个功能的测试需求,所以首先要对项目有个全面的了解。下午翻阅了一下项目计划表、需求任务书、系统概要设计、测试任务书等相关文档。

【产品系统测试环节】

参与了 V439 项目系统测试。V439 项目的开发使用的瀑布模型,其中也包括 3-5 个大的版本迭代和若干个小版本迭代。当前已经进行了 2 个迭代的开发,项目基本功能开发完成,已经开始了基本功能的系统测试。测试主要是采用黑盒测试的方法,从用户使用的角度来进行测试,这也是大多数嵌入式系统通常采用的测试模式。与之相对的白盒测试或者说单元测试一般是在开发阶段由软件开发人员完成,而黑盒功能测试则是主要由专业的测试人员完成。对于大型软件来说,测试是非常重要的。测试人员和开发人员的配比一般是 1:1,或者测试人员比开发人员更多。所以,软件测试技术是一项必备技能,不管对于将来从事测试岗位的人员还是开发岗位的人员,软件测试技术是基本功。

【产品平台测试环节】

随项目进行 V439 平台功能测试,详细的观察了一下当前考试系统的功能和操作体验。

未来考试系统应该会手机、电脑平台拉通，使用相同的系统和开发方式进行研发。这样未来我们计算机专业的学生能够参与进行软件开发的范围会进一步扩展，计算机行业也会随着物联网的发展触及到人们生活中会接触到的各种电子设备，各种设备都需要软硬件的支持，计算机行业的发展也会迎来新的阶段。

关于考试系统的发展和现状，我的想法有以下这些：

1、 考试系统的开发历史是比较久的，20 年前左右美国日本德国的大型教育公司就已经开始进行考试系统研发。考试系统即可以作为嵌入式系统，也可以作为纯软件系统，通常采用小型操作系统，一般是对 Linux 系统进行裁剪，在操作系统之上进行定制化的软件开发。10 年前考试系统的开发规模就已经很大了，整个系统的软件规模估计有几百万行代码。大型考试系统的研发还主要是由美日德所掌握，在我国还未大规模普及。

2、 最近几年考试系统的研发进展好像不大。从所看到的考试系统来说，发展好像并不十分理想，用户体验可以说还停留在早些年水平。我想包括以下几点：一、操作界面比较复杂，用户体验不好。考试系统中包括大量的菜单，操作按键很多。界面中展现的信息过于繁杂，很难让人直观的接受和理解。二、普通考试系统比较独立，与互联网没有很好的融合。检索信息和数据盘不是实时更新的。也不能和网络中的信息联动。同时，手机的考试功能发展很快，操作简单易懂，又能和互联网很好的融合。考试考试的这些问题导致现在宁愿使用手机考试，也不愿意使用复杂的考试系统。前段时间听说苹果也开始投入考试系统的研发，应该是看准了目前考试系统用户体验差的这一点。人们现在越来越希望使用简单易懂的系统，而不再是复杂功能众多的系统。

【产品研发环节】

参与了 V439 项目第二阶段的研发。仔细观察了一下考试系统目前采用的研发模式和方法,可以说现阶段大型考试系统的研发方法还是比较先进的。因为考试系统属于嵌入式系统,以前的开发方式都是采用 C 语言,在一个 Framework 框架中进行系统开发。但现在的开发方法不再使用 C 语言直接编写代码,而是使用工具来完成。

1、 在 V439 项目中,需要定制开发的模块主要有 HMI 界面部分和 Audio System 部分。HMI 模块的开发使用 HMITool 完成, System 模块的开发也是使用 SystemTool。这两个 Tool 并不是我们在市面上能够看到的集成开发环境,而是大众专有的考试系统开发工具,不做大众汽车考试的开发是不会接触到这些 Tool 的。而且,现阶段已发展为绝大多数的代码都是在 Tool 中直接进行编写,代码的语法和 C 语言语法类似,但也有不同之处。以 HMITool 为例,这个工具有些类似于 Excel 表格,系统开发过程就是在表格不同的单元格内填写处理逻辑的功能代码。编写完成后生成一张表格,再使用专有的 Compile 工具对表格进行编译,就能够生成在 Linux 系统下可执行的软件。

2、 这个编码过程感觉很有意思。编写代码的人甚至可以对考试系统的代码框架完全的不了解,只需要知道一个很小的部分,例如一个画面需要实现什么功能和 Tool 的基本语法就可以进行代码开发。Tool 使用的代码语法也是非常简单的,可以说接触过基本编程语言都能够很快上手,对于开发人员的知识储备和技能要求门槛越来越低。这种方式既有利于大量功能的快速开发,也有利于关键资产的保护,便于将项目外包到多个公司去做。这也是 V439 项目能够由包括龙创公司在内的多家公司在共同开发的原因。

3、 软件系统开发门槛的进一步降低,以及物联网大规模电子器件智能化的普及,我想将会带来一次行业的变革。这一变革将会需要大量的普通软件开发“蓝领工人”,就像现代的考试系统开发这种模式,不再对程序员技能有过高的要求,只需要基本的软件开发技能就可以进行系统研发。这一点对未来的高级职业教育是非常有利的。我们要培

养我们的学生具备基本的通用软件开发和系统能力，和良好的职业素养，来应对未来计算机领域工作岗位的要求。从企业的角度来看，良好的职业素养比技术能力更重要。

【产品评审环节】

参与了项目代码的走读和 HMI 部分功能的评审，接触的功能是 Subject Update。这部分功能看似比较简单，但 HMI 部分涉及到的画面数量也有几十个之多。HMI 画面迁移开发的参考主要是迁移式样，这部分之前我也有过接触。迁移式样仍然是非常详细的，可以说对每个画面、功能点介绍非常全面细致，这也是我觉得国内软件开发和国际产品开发比较大的差异。国际软件开发注重文档，注重软件过程，国内软件开发一般不太关注这些。V439 项目因为是大众车型，与德国合作，以德国方面的要求为主。所以，软件开发流程和评审流程非常规范。每一份软件文档的写作都是需要花费很大的心血，甚至大于代码编写本身。德国人的做事严谨态度从软件流程的要求上就能体现出来。这也说明我们的学生如果参与到这类国际软件开发，或者外包项目中，需要的不仅仅是代码编写能力，软件规范、说明文档写作都是基本功。

【IT 高新技术研发】

1、 在现代考试系统中涉及到很多前沿领域的科技。例如谷歌汽车的无人驾驶技术，这就要求汽车对于摄像头捕捉到的画面识别能力很高。这次有幸能够接触到考试系统中一些前沿领域的研发内容。这些内容很多属于基础性研究，也是未来智能化着重发展的方向。其中一个热点研究“局部特征匹配”是通过兴趣点检测、局部特性描述、特征匹配的方法对同一物体的两张尺度、旋转角度不同的图片进行局部特征匹配。

2、 与项目同事商量，我投入了一些时间对考试系统中“局部特征匹配”进行了学习

和研究，我想搞清楚一点这些前沿领域在做什么事情，以及如何转化为实际产品。实际去分析研究的过程中还是有很多困难。今天掌握到“局部特征匹配”分为 3 个方面。第一点是使用 Harris 算法进行角点抽取，在图像中，角点是二维图像亮度变化剧烈的点或图像边缘曲线上曲率极大值的点，是一个重要的局部特征，它决定了图像中关键区域的形状，体现了图像中重要的特征信息，所以在目标识别、图像匹配、图像重构方面角点具有十分重要的意义。角点具有丰富的特征信息，使其在保留图像图形重要特征的同时有效地减少信息的数据量，提高了计算的运算速度，有利于图像的可靠匹配，使得实时处理成为可能。由于角点具有旋转不变性，因此角点几乎不受光照条件的影响，在图像配准与匹配等计算机视觉领域起着非常重要的作用。Harris 算法是研究图像中一个局部窗口在不同方向进行少量的偏移后，考察窗口内图像亮度值的平均变化。

3、继续上周末研究的考试前沿领域“局部特征匹配”的内容。第二方面是，使用 SIFT 算法进行局部特征描述。SIFT 算法是一种提取局部特征的算法，在尺度空间寻找极值点，提取位置，尺度，旋转不变量。关键点抽取完毕后，SIFT 算法还需要实现以下 2 个步骤：1、为每个关键点指定方向参数；2、关键点描述子的生成；

4、通过特征匹配，判断两个图像的匹配关系。将第一个图像中的每个特征点与第二个图像中所有的特征点进行匹配，经过 n^2 次匹配后能够获得所有匹配点。通过调整 distRatio 匹配参数来获得最佳匹配效果。通过“角点抽取、局部特征描述、特征匹配”3 个步骤，完整图像的“局部特征匹配”，这些内容属于程序算法的基本原理，通过 Matlab 代码实现仿真，再使用多组图像进行训练，并调整算法的参数达到最佳效果。之后项目会将实验的成果应用到产品中。这个功能集成到考试系统中，就可以实现用摄像头捕捉到的图像和考试系统中存储的图像进行匹配，达到物体识别的效果，例如交通标志识别，识别的成功率还是非常的高。这是一个基础研究转化为产品的很好例子。

5、 现阶段考试系统中也在逐步应用 IT 领域的先进技术和算法 ,随着网络的普及 ,考试系统中加入无线模块 ,使考试系统能够与网络相连 ,这样就能够通过网络实现很多智能化的功能。这也是考试系统研究的前沿领域之一。设想今后的考试系统应该会借助云计算虚拟化技术 ,安装在机房中呈现给用户的只是一个终端 ,而所有的题目和答案、数据 ,以及其它智能引导功能都是在云端计算完成。即使当前的考试系统中所具备的联网功能 ,也是能够收集到大量的数据。在日本等发达国家 ,早已使用 VICS 等信号源与考试系统进行数据交互了。这两天我计划花时间了解一下考试系统与后台服务器之间的交互 ,这部分内容也正是目前学校主推的云计算和大数据技术相关的内容。

【大数据分析技术在现代企业中的应用】

1、 系统通过后台服务器主要是进行大规模的数据分析 ,应用于考试系统中可以对实时考生答题情况进行跟踪。在实际项目中后台服务器上运行的是目前流行的数据分析 Hadoop 框架。这两天着手在服务器上进行 Hadoop 环境的搭建。Hadoop 环境是基于 Linux 系统搭建的 ,安装环境后又进行很多的配置。环境配置完成后就可以进行数据获取和分析计算了。

2、 本项目中数据获取采用 htmlparse ,这是一个纯的 java 写的 html (标准通用标记语言下的一个应用) 解析的库 ,它不依赖于其它的 java 库文件 ,主要用于改造或提取 html。它能超高速解析 html ,而且不会出错。现在 htmlparser 最新版本为 2.0。htmlparser 是目前最好的 html 解析和分析的工具。htmlparse 的主要工做就是使用正则表达式匹配获取数据 ,然后添加到数据库中。

3、 Hadoop 环境的构建。使用 htmlparse 获取到数据后 ,需要使用 MapReduce 框架进行数据分析。在 Hadoop 中 ,每个 MapReduce 任务都被初始化为一个 Job ,每

个 Job 又可以分为两种阶段：map 阶段和 reduce 阶段。这两个阶段分别用两个函数表示，即 map 函数和 reduce 函数。map 函数接收一个<key,value>形式的输入，然后同样产生一个<key,value>形式的中间输出，Hadoop 函数接收一个如<key,(list of values)>形式的输入，然后对这个 value 集合进行处理，每个 reduce 产生 0 或 1 个输出，reduce 的输出也是<key,value>形式的。

4、 编写代码实现，Map 函数输入 key-value 类型分别是 Text,ArchiveReader 类型。循环每一个 ArchiveRecord 对象，获取到 Server 值，输出中间格式为 (key,value) 为形如 (Apache 1,1,1,1.....)。以供 reduce 函数调用处理。Reduce 函数使用系统已经封装好的 longSumReducer，主要就是实现 values 的累加，指定 map 的输入格式。

5、 最终执行 MapReduce 任务，获得分析结果。在 MapReduce 中，由 Job 对象负责管理和运行一个计算任务，并通过 Job 的一些方法对任务的参数进行相关的设置。此处设置了使用 TokenizerMapper 完成 Map 过程中的处理和使用 IntSumReducer 完成 Combine 和 Reduce 过程中的处理。还设置了 Map 过程和 Reduce 过程的输出类型：key 的类型为 Text，value 的类型为 IntWritable。任务的输出和输入路径则由命令行参数指定，并由 FileInputFormat 和 FileOutputFormat 分别设定。完成相应任务的参数设定后，即可调用 job.waitForCompletion()方法执行任务。

【专业实践总结与收获】

在本次专业实践中，我指导 2 名大数据学生进行相关项目工作，学生已经顺利完成所有相关工作任务，并按要求制作完成了考生系统开发的相关模块。与企业相关工作人员就企业情况、企业需求、学生情况和人才培养等各方面进行了深入的讨论，并将企业对人才的要求融入课堂教学中。在今后的教学过程中，我将更加注重激发学生学习的主动性和创造性，

重视培养学生的“工匠精神”，有效提升学生的就业能力，是教育的核心。在日常教学中，使学生明确目标，树立正确的职业观，知己之所能，有所为，有所不为，实现个人价值。