

文章编号: 1672-5913(2013)17-0037-04

中图分类号: G642

编译实习课程的创新教学

郭 耀, 王千祥

(北京大学 信息科学技术学院 高可信软件技术教育部重点实验室, 北京 100871)

摘 要: 编译实习是深入学习和理解编译原理和技术的重要手段, 也是编译课程的重要组成部分。针对当前计算平台和系统的迅速演进和发展, 北京大学在编译实习的教学过程中进行创新尝试, 让学生面向包括智能手机等在内的新型计算平台开发一个完整的小型编译器。文章介绍编译实习课程的创新思路, 对北京大学近 3 年来的实践情况进行总结分析并提出展望。

关键词: 编译原理; 编译实习; 课程创新; 移动平台

0 引 言

编译原理是计算机科学与技术专业本科生的必修课程^[1-3]。国内外高校多年的教学实践表明, 编译实习是深入理解编译原理和技术的重要途径^[4-5], 即所有学生都应该能够自己动手实现一个面向某个简单语言或语言子集的小型编译器。编译实习可以与编译原理课程安排在同一个学期, 也可以分开。由于编译原理课程内容较多并且较难, 目前北京大学采用的方法是在学生上完编译原理课程的下一个学期开设编译实习课程。

我们曾经尝试过不同的编译实习题目。例如, 在 20 世纪 90 年代采用的是面向 Pascal 语言子集 Mini-Pascal 的编译器, 后来转为实现 C 语言子集 Mini-C 的编译器, 随着 Java 语言的流行, 最近几年改成了面向 Java 语言子集 MiniJava 的编译器^[6-7]。虽然面向的语言不同, 但总体目标都是要实现一个相对完整的编译过程, 包括词法分析、语法分析、语义分析和类型检查、中间代码生成、编译优化、寄存器分配等编译阶段, 最后生成面向一种虚拟机的可执行目标代码, 编译

器产生的最终结果可以在虚拟机上执行和验证。

进入 21 世纪以来, 各种新型计算平台和环境不断涌现, 特别是在 2008 年之后, 随着 iOS、Android 等移动智能计算平台的流行, 许多大学生都拥有了智能手机并热衷于开发智能手机上的应用。如果能够让学生面向包括智能手机在内的新型计算平台开发自己的编译器, 使生成的目标代码可以在各种实际的计算机硬件和操作系统上运行, 那么会激发学生对计算机技术和编译技术的更大兴趣。

基于此出发点, 作为对编译实习课程的创新探索, 我们从 2010 年起在北京大学设立编译实习实验班(以下简称实验班), 提出把要实现的编译器运行平台从课程提供的虚拟机平台, 转移到实际的计算机硬件和操作系统平台上, 如 Java 虚拟机(JVM)、Windows 等传统平台以及较新的 Android、iPhone 等智能手机平台。通过选修实验班课程, 学生可以利用编译基本原理实现一个在实际平台上运行的编译器; 也可以选择面向新的语言实现编译器或解释器, 如 Ruby、Scheme 等比较新的语言。另外, 在编译器实现过程中, 我们通过数据流分析等技术, 实现更深

基金项目: 国家自然科学基金项目(61103026)。

第一作者简介: 郭耀, 男, 副教授, 研究方向为操作系统、编译技术、低功耗技术、移动计算等, yaoguo@pku.edu.cn。

入的编译优化功能,提高编译器生成程序的执行效率。

1 设计目标与理念

1.1 设计目标

在课程的设计上,实验班专为编程能力较强、有结合最新计算平台开发真实编译器意愿的高年级学生开设。在编译实习普通班上,我们仅仅要求编译结果能在模拟器上运行,但这与实际平台的运行还有一定距离。实验班要求编译结果能在最新出现的平台上运行,如 JVM、Android、Iphone/Ipad 等,挑战性更大,收获也更大。编译实习课的总体目标是培养学生具有良好的实践能力,充分发挥学生的主动性和积极性,给学生上讲台的机会,以充分展示自己。

具体来说,实验班希望学生达到以下主要目标:(1)深入掌握编译技术并将其应用到编译器的实现中;(2)深入掌握运行平台的细节,在使用中得到更好的认识;(3)开阔视野,提高学习计算机科学的兴趣爱好;(4)培养合作开发大型软件能力以及工程实践能力;(5)培养学生的演讲和沟通能力。

1.2 设计理念与特色

实验班的主要目的是把要实现的编译器的运行平台从课程提供的虚拟机平台,转移到实际的计算机硬件和操作系统平台上。实践证明,虽然面向实际应用平台开发编译器原型系统挑战很大,但是这可以在很大程度上提高学生对编译器开发的兴趣,在实际教学中取得很好的效果。通过近 3 年的实践,我们打造了拥有以下特色的实验班课程。

(1)面对新计算平台不断出现的特点,结合实际的计算平台实现可用的编译器。与普通班采用统一虚拟机作为编译的目标平台不同,实验班要求学生选择一个实际的计算平台作为编译的目标平台,同时鼓励学生选择包括 Android、iOS

等移动平台在内的新型计算设备作为编译的目标平台。在实践中,学生对新型平台非常感兴趣,如在 2011 年的编译实习实验板上,4 组学生都选择面向 Android 平台开发相应的编译器。

(2)鼓励学生自主探索相关领域的最新技术,如新计算平台技术、新程序分析技术等。由于使用新的平台需要学习许多新的知识,包括该平台上的体系结构、目标代码格式、相应的运行环境等内容,而这些内容无法通过课程讲授完全获得,因此我们鼓励学生通过阅读相关技术资料,特别是通过上网查找相应的资料以了解新的计算平台相关技术。另外,我们也鼓励学生结合新的计算平台,采用新的程序分析技术对目标程序进行优化。

(3)鼓励学生开展项目小组内和项目小组之间的合作。掌握新平台的知识需要不断学习,因此我们鼓励小组内和小组间展开合作,学生在课程内和课程外通过组间交流互通有无,共同进步。例如,Android 采用的 Dalvik 虚拟机上的 DEX 程序格式比较复杂,各小组之间可以通过分工学习和深入交流提高掌握的效率。学生反映组间的合作让他们受益颇多。

2 教学实践情况

从 2010 年开始的每年秋季学期,我们都在编译实习课程中鼓励学有余力的学生参与到实验班的探索实践中。3 年来,共有 37 名学生分为 17 个项目小组参与到不同种类自选题目的实习项目中。2010—2012 年实验班的项目及选题人数见表 1。

2.1 第 1 年:初次尝试自选题目

我们对编译实习课程的改革尝试始于 2010 年,共有 5 个小组 10 名学生选择自选题目的编译实习项目。首先有一组学生在教师的鼓励下,尝试面向 JVM 实现 MiniJava 编译器。这个小组由 2 名学生组成,编程能力都很强,在跟着普通

表 1 2010—2012 年编译实习实验班的项目及选题人数

学年	项目	选题人数 / 组数
2010	面向 JVM 的 MiniJava 编译器实现	2/1
	面向 Unicore 平台的 MiniC 编译器实现	6/3
	面向 Java 程序的污点分析 (Taint Analysis) 实现	2/1
2011	面向 Android 平台的 MiniJava 编译器实现	12/4
	面向 Windows 平台的 MiniJava 编译器实现	2/1
	面向 JVM 的 MiniJava 编译器实现	3/1
	面向 Unicore 平台的 MiniC 编译器实现	2/1
2012	Scheme 语言的解释器实现	2/2
	面向 ARM 平台的 MiniJava 编译器实现	2/1
	面向 Android 平台的 MiniJava 编译器实现	2/1
	Ruby 语言子集的解释器实现	2/1
总计		37/17

班上过前面几次课之后,基本上就能够自主开发完成面向 JVM 的 MiniJava 编译器,并且最终实现的编译器效果很好。

我们还与体系结构实习课程的任课教师合作,遴选 3 组学生面向北京大学自主研制的众志处理器 UniCore 实现 MiniC 的编译器,不仅要求他们尝试面向新的平台开发编译器,而且还把编译实习与体系结构实习有机结合在一起,进行有益的探索。

除此之外,我们还推荐学生实现更高级的程序分析技术,一组学生在老师的带领下,实现了 Java 程序分析中一种非常重要的新的程序分析技术:污点分析 (Taint Analysis^[8])。这组学生通过自己查阅论文,分析现有的 Java 编译器 (Soot),自主实现污点分析的技术和相应工具,取得了很好的效果。

2.2 第 2 年:移动平台大受欢迎

从 2011 年开始,我们加大对实验班的宣传力度,共有 19 名学生参加并分为 7 组,尝试开发面向 JVM(1 组)、Windows(1 组)、Android(4 组)的 MiniJava 编译器,以及面向 Unicore 的 MiniC

编译器(1 组)。该年度最大的特点是很多学生对智能手机的应用开发产生浓厚兴趣,北京大学有许多学生开发面向 Android 和 iOS 平台的校园智能应用,有 12 名学生分为 4 组分别实现了面向 Android 平台的 MiniJava 编译器,该编译器主要有以下特色。

(1) Android 是一个相对较新的系统,虽然它采用的主要语言是 Java,但是却专门为移动平台开发了新的 Dalvik 虚拟机,取代了 JVM。JVM 字节码的指令采用的是栈式结构,而 Dalvik 字节码采用的是寄存器风格的指令。由于 Android 系统的资料不是很全,许多信息需要在互联网上查找,甚至要自己动手尝试以分析代码格式,因此对学生来讲仍存在较大挑战。

(2) 一般来说,Android 智能手机上的应用是一个 APK 包,其中除了 Dalvik 字节码的目标码格式的可执行文件之外,还有许多其他的辅助文件。要生成合法的可以在智能手机上安装的应用,还需要在编译完成之后进行许多复杂繁琐的收尾工作。

(3) 如果不仅可以生成在 Android 手机上运行的应用,而且还可以让 MiniJava 编译器在 Android 手机上运行,那就更好了。有一组学生完成了这个看似很具挑战性的任务,即可以直接在手机上运行编译器。

总的来说,这一年面向 Android 开发编译器的尝试,不仅对学生是一个挑战,而且也让任课教师收获颇多。由于参与人数比较多,项目主要由学生自主完成,教师期间共组织了多次交流与讨论,进行必要的引导与建议。最终的编译器尽管差异较大,但总体效果非常好,多数学生在完成这个项目后,大大增强了开发大型项目的自信心。

2.3 第 3 年:理性回归+百花齐放

虽然 2011 年的尝试很成功,但是由于参与的学生较多,难以管理,因此我们在 2012 年对学生的自选题目进行了更为严格的筛选,最终有

8 名学生分别进行了 4 个不同题目的开发。我们每年都会鼓励学生选择新的题目,因此 2013 年只有一组学生选择了面向 Android 平台的编译器实现,而其他学生则分别根据自己的兴趣选择了不同的项目。

其中,2 名学生在选择函数式程序设计课程时,对 Scheme 语言产生浓厚兴趣,在教师的鼓励下,决定尝试编写 Scheme 语言的解释器。作为一种函数式语言,Scheme 的解释器和 C++、Java 等命令式语言的编译过程有较大区别。最开始 2 名学生作为一个小组,但是由于在项目中期出现实现理念上的冲突,因此 2 名学生分别实现了 2 个版本的 Scheme 解释器。一个版本的解释器效率较高,程序执行速度较快;另外一个版本的解释器功能较强,可以支持更复杂的 Scheme 语言结构。2 名学生都体现出了非常好的基础素质以及对编译原理和构造技术的深入理解。

另外 2 名学生也对 Android 等移动平台感兴趣,但是却另辟蹊径:编译器的目标不是生成 Dalvik 虚拟机上执行的 APK,而是生成可以在 ARM 平台上执行的本地(Native)应用。由于之前还没有学生尝试过类似项目,因此还需要调研 ARM 指令的格式与 ARM 平台上 Linux 相关库的实现,以便最终实现的 MiniJava 编译器可以生成在 Android 平板电脑上执行的本地程序。

值得一提的是,还有 2 名学生选择实现一个支持 Ruby 语言子集的解释器。这也是一个很有挑战的项目,因为 Ruby 语言较新,资料不完整。2 名学生虽然在前期表示出很大的信心和决心,也进行了较为详细的调研工作,但是由于在中期遇到来自其他项目的压力,无法投入足够的时间完成最后的实现工作,很遗憾地选择了中途退出。这个例子也提醒我们,在课程初期需要更加深入地考察学生的能力和所选择项目的难度,尽可能提高每个项目的成功率。

经过 3 年的实践,我们积累了较多经验,收到的学生反馈也大多是正面和积极的。我们计划将实验班的课程作为一个长期的课程实践开展下

去。目前,2013 年秋季的课程也再次列入计划之中。

3 经验与挑战

在课程实践过程中,我们发现选择将新出现的计算平台作为编译目标平台,挑战性比较大。由于平台新、发展快,资料更新快,教师提供的信息往往也不全面,因此这就要求学生有更强的自主解决问题能力。新计算平台容易吸引学生,因为许多学生喜欢新技术,他们认为新技术实现空间大,也更能体现自己的能力,还能接受挑战。这对于激发学生的热情,释放学生的创新能力很有意义。然而,在实践过程中,我们也遇到了一些挑战和问题。

(1)如何更好地传承课程经验?学生在结束开发工作后,都会按照要求提交实习报告,这些报告是课程的宝贵财富。如何充分利用好这些内容,值得我们继续思考。

(2)如何与普通班进行更好的协调?尽管实验班的题目与普通班有差异,但有些内容(如词法、语法分析工具、编译器的基本框架等)还是很接近的。这就需要在内容设置、评价细节等方面进行持续优化。

(3)如何设置更合理的选题?虽然我们鼓励学生自选题目,但每年还是会在开课初提供一些选题列表,将前几年的项目题目作为参考发给学生。学生比较容易被新颖的题目吸引,但还是会综合考虑创新性和难度。

(4)如何更好地发挥教师的作用?对于自身编程能力较强的学生,教师除了为其提供好的题目之外,还应如何给予恰当的帮助?教师指导得好既使学生能够快速获取必要的信息而少走弯路,又避免因建议太多而束缚学生的思路。

(5)如何提高评价的客观程度?尽管主观评价是主要手段,但增加可行的客观手段是非常好的补充,尤其对于相同题目,增加在目标码的执

(下转第48页)

园, 在全校学生范围内起到良好的宣传效果。结合竞赛主题及形式布置课程大作业, 便于学生积极参与竞赛。教师通过竞赛掌握了学生的计算机应用情况, 也提高了自己的教学水平及能力。

“北京大学计算机应用设计大赛”已成为校园品牌, 学生参赛累计人数达上千人。2008—2012年期间, 北京大学的学生在“中国大学生计算机设计大赛”中取得了2个一等奖、5个二等奖和4个三等奖。

6 结 语

实践证明, 上述改革措施激发了学生的兴趣

及潜能, 增强了学生的信息素养、实践能力以及创新能力。来自北京大学文科院系的本科一年级学生, 多次参加中国大学生计算机设计大赛, 他们以富有创意的大赛作品、娴熟的综合技术运用以及较全面的解决方案, 获得大赛一等奖。课题组的教师曾多次参加全国性文科计算机基础教学研讨会并介绍课程改革方案, 得到来自其他院校教师的良好反馈。下一步我们将以培养学生计算思维能力为目标, 探究一种具有文科学生特点的教学内容和教学方法的改革方案。

参考文献:

- [1] 张铭, 谢柏青. 北京大学计算机基础课程教学体系调查[J]. 计算机教育, 2005(15): 18-21.
- [2] 钱丽艳. 基于Blackboard平台的北大教学网在不同规模课程中的应用与实践[J]. 计算机教育, 2010(2): 94-97.
- [3] 刘志敏, 黄贝宁. 计算机基础课软件应用测试及考试系统[J]. 计算机教育, 2009(4): 130-132.
- [4] 刘志敏, 马郢. 北京大学大学生计算机应用设计大赛组织经验浅谈[J]. 计算机教育, 2012(22): 53.

(编辑: 彭远红)

(上接第40页)

行时间、内存开销、编译速度等方面的比较, 更能激发学生优化编译器的热情。

4 结 语

为了适应迅速发展的计算机硬件和系统平

台, 我们在北京大学信息科学技术学院编译实习课程中进行创新探索, 并在实践过程中获得一些经验, 同时还解决实践过程中出现的问题, 希望能与各高校承担编译课程教学任务的教师进行广泛交流, 也希望大家多提宝贵意见。

参考文献:

- [1] 蒋宗礼. 编译课程教材建设[J]. 计算机教育, 2007(11): 74-76.
- [2] 蒋宗礼. “编译原理”课程与专业能力的培养[J]. 计算机教育, 2009(21): 3-6.
- [3] 何炎祥, 伍春香. 计算机专业不需要开设编译原理课程吗? [J]. 计算机教育, 2009(4): 61-62, 85.
- [4] 杨海燕, 史晓华, 张莉. “编译技术”实践环节的质量管理体系及实践[J]. 计算机教育, 2009(17): 61-63.
- [5] 王生原, 董渊, 张素琴. “编译原理”课程实验项目介绍[J]. 计算机教育, 2009(23): 12-14.
- [6] Palsberg J. From Java to MIPS in five nifty steps[EB/OL]. [2013-07-15]. <http://www.cs.ucla.edu/~palsberg/course/cs132/project.html>.
- [7] Appel A W, Palsberg J. Modern compiler implementation in Java[M]. 2nd ed. Cambridge: Cambridge University Press, 2002.
- [8] Tripp O, Pistoia M, Fink S J, et al. TAJ: Effective taint analysis of web applications[C]// Proceedings of the 2009 ACM SIGPLAN conference on programming language design and implementation. New York: ACM, 2009: 87-97.

(编辑: 宋文婷)