本科毕设开题答辩稿

各位老师好，我是10级本科生赵继伟，我的本科毕设题目是B\*语言解释器的实现，指导老师是马殿富马老师。

目前机载软件复杂度日趋提高，美国F35战斗机的机载软件计划约860万行代码，而空中客车公司的在板软件大小也随着型号的更新成倍提升。与此同时，机载软件的安全标准也日趋严格，2011年DO-178C标准取代了旧的DO-178B标准，对机载软件提出了更高的要求，而其中更是着重强调了形式化验证部分。ISO/IEC 15408标准也是将完全形式化验证作为最高的软件标准。

而我的毕设内容，便是要开发一种形式化验证工具，将验证过程中的机械验证的部分自动化完成，来加快形式化验证的速度。

近十年来，国外已经有很多研究人员在软件的形式化验证领域取得了一定的成果，而国内相关的研究还处于相对较少的阶段。我的导师马老师吸取了前人的经验和教训，提出了一套自己的形式化验证模型。虽然这不是我提出的，但是我的任务便是要实现这个模型，所以我还是要对这个模型以我的理解做一个简短的介绍。

描述一个程序，简单来说可以分成高层规格描述——程序要做什么，低层规格描述——程序怎么做两个层级，这个模型主要针对的是低层规格描述。

对于绝大多数程序的行为，我们可以将其归纳为针对不同的输入完成不同的输出。根据输入，我们可以将程序的行为分为功能性行为与鲁棒性行为，也就是对于正确还是错误的输入而产生输出的行为。

对于每一段程序而言，它的行为不仅受到其自身的影响，还受到其语境——程序上下文的影响。

所以，我们可以将程序的行为，归纳为以下的一个表达式

欧米伽X撇K表示某一段程序的语境，伽马表示程序，即是毕设题目中的B\*语言，R是鲁棒性的逻辑语句，Q是功能性的逻辑语句。

我们通过把程序转换为这样的表达式，就可以把由程序所建立的输入与输出之间的关系是否和需求确立的输入与输出之间的关系做出比较，来验证程序是否是正确的。

此模型的具体推理系统在此由于时间关系我就不一一展开了，而是将模型的这些细节随着我研究内容一起说了，总体而言，由这个模型出发，我们需要区别对待的是分支语句，循环语句，以及数据结构的模型。这三点也构成了我研究的核心内容，也是研究的难点。

一般解释器是一个线性的过程，但是这个模型我需要的是遍历分支语句的每个分支，我在对一个分支语句进行处理时，true和false两个结果所引导的语句块，需要分别对前文形成的语境进行作用，这样就形成了两个平行的分支，在每个分支中继续向后演算，虽然这些最后输出的逻辑语句中统一，但是在实现过程中是需要分开做的。

另外，对于循环语句，目前只能由人工，通过数学归纳法来做，这也是我的毕设不能完全自动化的原因，所以针对这部分我需要实现一个便捷的交互模式，让验证人员可以方便的完成验证工作，并且还能确认是否验证正确。

最后一个研究的难点就是数据结构部分，在这个形式化验证模型中，所有的常用数据结构都需要用集合来表示，这给我带来的难点在于就是如何在常用数据结构与集合之间实现一个映射关系。

我毕设的研究基本流程是一个迭代的过程，先把B\*语法的BNF范式完成，然后确定一个典型的程序，再以这个程序往后推进。

由于我是保研的，上学期就开始做了，不过由于过程中走过一些弯路，目前的进度做到了第三步，对于非线性变化的语境已经给出了一个包含数据结构和算法伪代码的解决方案。循环语句目前可以给出一个流程图，很快就能完成数据结构和算法伪代码。之所以先着手分支语句与循环语句主要是因为这两部分完成就可以为整个解释器搭建起一个完整的运行框架，数据结构部分是在这个框架中的一个部分，只是其表示形式与以往的不很一样。

总体的时间安排是希望在中期检查前完成软件的主体编写，尽快编写毕设论文。