

科创研究院远程科研实习营——机器学习方向（利用关联算法学习不变式，形式化验证协议的正确性）

December 17, 2017

1 简介

此实习项目专门为计划申请计算机、软件工程、电子及自动化等专业的学生所设计。学生将跟随中科院导师一同工作，学习形式化方法，使用机器学习的关联算法学习推导协议不变式，从而验证验证协议正确性。实习结束后，导师会根据学生表现出具推荐信。

2 内容

本项目名称是基于关联算法的协议验证。

- 掌握协议的形式验证：包括协议形式化描述，可达集，不变式的概念，以及利用不变式描述协议的重要性质。
- 利用协议验证工具CMurphi生成可达集，将可达集合转化为数据表，以供数据分析。
- 熟练机器学习中的关联算法中的一种：Apriori。
- 基于Apriori算法对可达集数据表进行分析，提取,并筛选出正确的不变式。
- 利用不变式对若干协议进行形式化验证，并完成协议的验证报告。

通过以上学习，使学生能够：

- 了解并掌握形式化方法基本思想：利用严格的数学方法，并利用计算机程序辅助验证，来保证系统的正确性；

- 理解机器学习中关联技术，理解平凡模式与关联规则的基本思想，以及熟悉Apriori算法实现。
- 尝试将机器学习的关联技术与协议的形式验证结合起来。

3 师资背景

任职教师现任知名研究所副研究员。毕业于上海交通大学，获计算机软件理论专业博士学位。主要研究方向包括形式化方法，协议验证与分析，机器学习。

4 招生对象及要求

大二以上优秀本科生及部分优秀高中生，计划申请信息科学、运筹学、应用数学等相关专业，为了让学生可以更好的完成科研项目，项目组会以笔试和面试的形式对学生进行筛选。

5 项目形式

远程科研指导项目时长一个月，具体时间可根据学生需求以及合适的时间进行安排。该项目优点在于对有足够长申请时间的学生而言，导师可以帮助学生更加深入、更加全面、更加系统的完成一项或几项专业领域科研任务，让学生亲身参与到整个科研项目开展的过程，体验解决科研难题的成就感，同时可以让学生了解到该领域背景及前沿动态等。除了定期科研项目讨论课程之外，项目周期内学生可以随时向导师请教相关问题，得到导师的专业指导，让学生提前体验到一名研究员真实的工作和生活状态。项目结束后，导师会依据学生表现出具推荐信。若学生表现出色，并且对本项目课题做出具体的贡献，本项目在研论文若用到上述贡献，并被发表，可以将学生列为合作者。具体课程安排如下：

1. 第一周：初步了解两种主要的形式化方法：模型检测技术与定理证明技术；以缓存一致性协议为具体实例，以murphi, Nusmv为工具学习并掌握协议的形式化建模；理解协议验证的关键概念：协议的并发语义，协议的可达集以及不变式性质；能够对重要的互斥性质，数据一致性进行描述，并基于上述两种工具进行验证；初步了解协议带参验证的基本概念。并以协议验证为例，更加深入理解形式化方法。
2. 第二周：熟悉Cmurphi工具，以临界区协议等若干经典协议为实例，利用工具生成协议可达状态集，基于python或者其他编程语言将可达集转化成数据表。

3. 第三周：熟悉并掌握机器学习中的关联技术：Aproi算法，掌握并使用基于python或者JAVA语言有关的Aproi程序包，分析在第二周建立起来的数据表，提取并筛选出合适的不变式集合。
4. 第四周：基于所提取的不变式集合，与其他工具结合，看能否给出带参协议的正确性证明，若能，则利用相应的工具生成正确性证明，并给出分析报告。