L-CMP：基于关联规则学习的自动化带参验证工具

李勇坚 曹嘉伦 庞军

L-CMP: an automatic learning-based parameterized verification tool. In Proceedings of the 33rd ACM/IEEE International Conference on **Automated Software Engineering** 2018 Sep 3 (pp. 892-895). ACM.

联系人：曹嘉伦　15959595756　caojl@ios.ac.cn

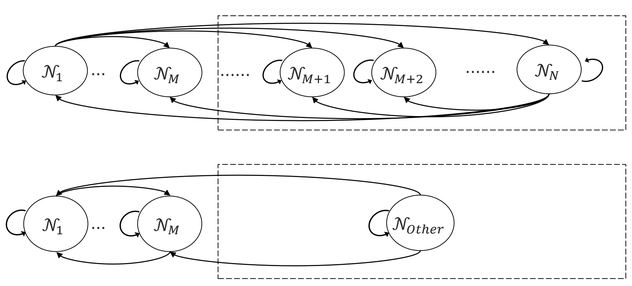
背景介绍

带参系统广泛存在于计算机体系中，提供数据交换、网络传输等系统功能。**带参协议**是对带参系统抽象建模的结果。常见的带参协议包括：缓存一致协议，安全性协议和网络通信协议等。

通常，带参系统$\mathcal{P}(\mathcal{N})$表示包含任意$\mathcal{N}$个具有相同结构的并发执行的主体。而带参系统的验证**目标**是验证安全性质是否在任意规模的系统中始终满足。该问题的**难点**在于：即使安全性质在小的实例上成立，仍然不能推出在任意规模下仍成立。而该问题被证明是不可判定问题。

解决这个问题的方法可以分为手工方法和自动方法。其中，手工方法需要人工提供辅助不变式，这个过程易于出错；而自动方法虽然能够自动地进行验证，但往往不具有严格的理论基础，且无法给协议设计者以启发。

因此，我们设计的L-CMP系统，不仅能够自动地对协议进行验证，还能提供可读性强的辅助不变式，对学者深入理解协议提供启发。并且，L-CMP的理论基础是经典的**``参数抽象及卫式加强”**（也称为CMP方法）。这种方法的核心思想如下图。具体做法是保留少量结点，并将其余结点进行抽象。经过卫式加强之后，使得抽象模型能够模拟原来无限系统的行为。最后，只要证明安全性质在抽象模型中成立，即可推导出安全性质在任意规模的带参系统中同样成立。

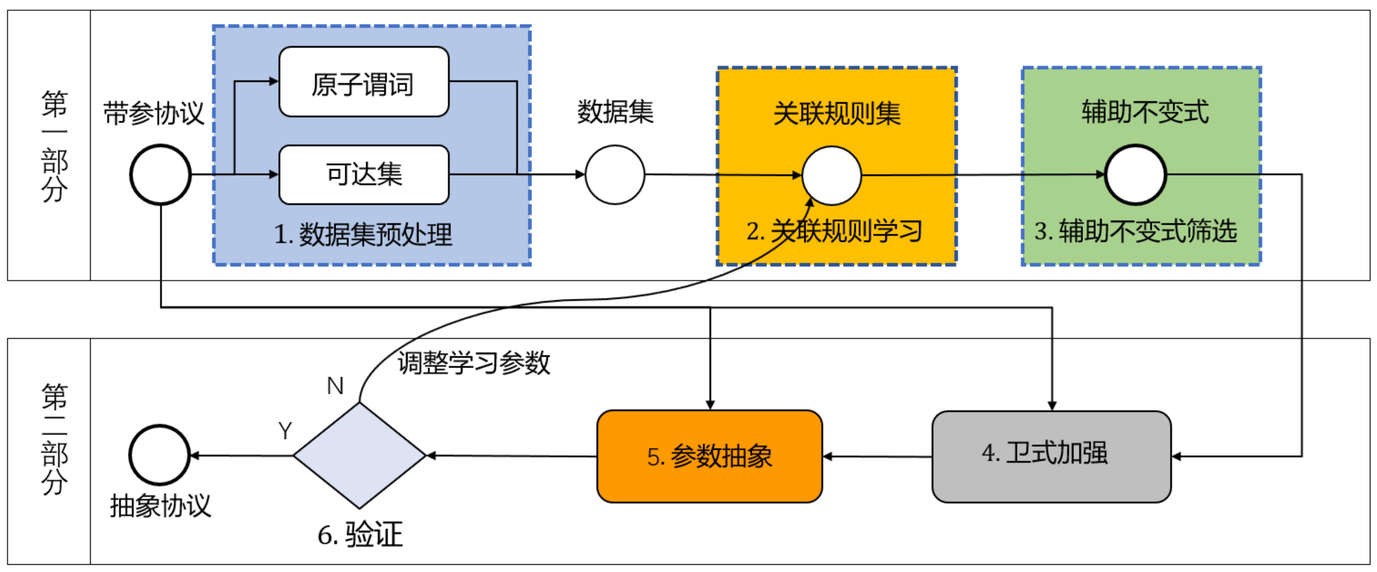


L-CMP 系统框架

L-CMP 的整体架构可以分为两部分：

- **第一部分**.
  - 这个部分主要目的是从带参协议的小实例中学习出关联规则，再经过筛选，筛选出其中的辅助不变式。总共包含三个步骤。给定一个带参协议P，先计算出一个小实例P(M)的可达集 (M通常为2或3)。如果小的实例无法通过模型检测，则需要重新检查带参协议，直到协议在小实例规模上能够通过模型检测；否则，则直接结束。经过适当的数据预处理，可达集将转变为一个可供关联规则学习使用的新数据集。这一步的预处理很关键，因为它能够使得学习出的关联规则足够用于加强与抽象协议。接下去，运用关联规则学习算法，从经过预处理后的数据集中学习出满足条件的关联规则。最后，用模型检测辅助筛选，得到辅助不变式。

- **第二部分**.
  - 这一部分的做法与CMP方法不同，卫式加强是执行在参数抽象之前。换句话说，在这一部分中，本文提出了``不变式指导的"加强策略，使得自动化的加强与抽象成为可能，同时避免了人为分析反例及提供辅助不变式的过程。经过卫式加强与抽象，生成的新协议将由模型检测器进一步进行检验。如果能够通过检测，则完成验证过程，输出抽象协议；否则，则说明学习出的关联规则不足，需要重新调整参数，继续学习。



实验结果

将 L-CMP 应用于经典的带参协议，验证它们的正确性。这些协议包括：MOESI, MESI, Mutual Exclusion (abbrv. MutualEx), Mutual Exclusion with data (abbrv. Mutdata)，German 和 FLASH。 其中，带参协议的实例大小除FLASH协议设置为3外，其他的协议均设置为 2。

协议	# 关联规则	# 辅助不变式	# 用到的辅助不变式	运行时间 (s)	结果
Moesi	736	20	5	24.744	√
Mesi	144	16	5	24.412	√
MutualEx	656	12	3	89.869	√
Mutdata	540	12	6	19.703	√
German	21202	448	8	85.418	√
Flash	358710	1636	327	41371.023	√

创新及贡献

- ✓ 提出了一个基于学习的自动化框架，L-CMP。它能够自动化地从带参系统的小实例中使用**关联规则**学习出辅助不变式，并进行自动化的抽象及验证工作。并且，学习出来的辅助不变式的形式可读性高，能给设计者启发；
- ✓ 提出了一种`“**不变式指导**卫式加强” 的模式。它比原先的`“反例指导‘卫式加强” 模式更有效，因为其避免了循环分析及加强的过程；
- ✓ 将基于统计的学习规则与形式化验证通过学习**辅助不变式**进行结合。这种结合能够给这两个领域的进一步的结合以启发。