

doi:10.1631/FITEE.1601783

题目：一种用于程序约束挖掘的过采样方法

概要：自动协议挖掘是获取精确而完备的 API 使用协议的有效方法。然而，与其它数据挖掘应用类似，自动协议挖掘方法需要足够多训练数据（即对象使用场景）作为输入。虽然通过增加程序的规模可提取更多数量的对象使用场景，但这会导致程序分析运行时开销较大。本文针对面向对象程序提出一种基于继承关系的对象使用场景过采样方法。给定一个面向对象程序 p ，一般情况下，执行 p 所能获得的对象使用场景数不超过运行时实例化的对象数。而本文方法可获得多达上述 n 倍的对象使用场景，其中 n 为程序 p 中一般对象使用场景的平均父类数。为了验证效果，在前期 API 使用协议动态挖掘原型工具 ISpecMiner 中集成上述方法并开展实验研究。实验采用扩展后的 ISpecMiner 从多个实际的程序中挖掘 API 使用协议。结果显示，采用本文方法获得的对象使用场景数是一般化方法的 1.95 倍。不仅如此，对比实验结果表明本文方法有利于挖掘更加精确而完备的 API 使用协议。特别值得关注的是，本文方法适用于无法实例化的类并挖掘出其 API 使用协议。这类 API 使用协议对于验证软件架构、程序说明和理解具有重要意义。虽然本文方法会增加一定的运行开销，但其仍在可接受范围内。

关键词：对象使用场景；API 协议挖掘；程序时序约束挖掘；过采样