

Sujet de projet GIS5 / Polytech'Lille

Comparaison statistique de méthodes informatiques d'analyses de programme

Laure Gonnord et Cristian Preda, Université de Lille, Polytech

Octobre 2012

Contexte Projet GIS, dans deux équipes, Émeraude du LIFL, pour la partie info, encadrée par Laure Gonnord, et Cristian Preda, Laboratoire Painlevé pour la partie Statistiques.

Sujet L'analyse statique de programmes est un ensemble de méthodes informatiques qui permet de caractériser le comportement de programmes séquentiels (C, par exemple). Ces méthodes permettent de calculer des formules numériques vraies (des "invariants") à certains endroits du programme, par exemple "ici, dans la boucle, la variable x a une valeur toujours comprise entre 1 et 10". Comparer deux méthodes est assez difficile, puisqu'en général il est impossible de prouver théoriquement la supériorité d'une méthode sur une autre.

L'objectif à moyen terme est pour nous de publier un benchmark de référence pour les méthodes/outils de génération d'invariants numériques de programmes. Nous avons déjà collecté un certain nombre d'exemples, mais ceux-ci restent à caractériser. Nous désirons proposer une ou des méthodes de comparaison statistique des méthodes de génération d'invariants, à l'aide de ces benchmarks.

Ce projet a donc les deux objectifs suivants :

- collecter et classifier des exemples classiques de la littérature du domaine. La collecte est déjà plus qu'amorcée, il s'agira de :
 - proposer des conventions de nommage appropriées
 - proposer des critères de classification, le nombre de lignes de code ne convenant pas toujours.
 - proposer des méthodes de stockage appropriées (base de données, autre ?)
 - évaluer ces solutions et les implémenter.
 - proposer une/des méthodes statistiques de validation d'outils d'analyse statique de programmes. Cette partie plus exploratoire a pour objectif de répondre aux questions suivantes :
 - Comment comparer deux ou plusieurs méthodes étant donné un benchmarks de programmes ? Comment visualiser et comparer leurs résultats ?
 - Est-il possible d'outiller (scripts ?) une telle comparaison ?
- Selon l'avancement du projet, il sera éventuellement demandé de coder les solutions proposées.

Compétences attendues Double compétence info-stats.

Mots clefs Analyse statique de programmes, comparaisons de méthodes, benchmarks, analyse statistiques de résultats.