Proposition de stage Master 2 recherche

Expressivité du modèle mémoire du greffon Value

Mots-clés : méthodes formelles, analyse statique, vérification de programmes, programmes C, interprétation abstraite

Cadre

Le CEA LIST est un centre de recherche technologique sur les systèmes à logiciel prépondérant qui mène ses recherches en partenariat avec les grands acteurs industriels du nucléaire, de l'automobile, de l'aéronautique, de la défense et du médical pour étudier et développer des solutions innovantes adaptées à leurs besoins. Au sein du CEA LIST, le Laboratoire de Sûreté des Logiciels (LSL), localisé à Saclay (Essonne, 91), développe des outils d'aide à la validation et à la vérification de logiciels et de systèmes matériels/logiciels, tout particulièrement dans le domaine des systèmes embarqués critiques.

L'un des nos outils, nommé Frama-C (`http://frama-c.com`), est une plateforme logicielle facilitant le développement d'outils d'analyses statiques de programmes C. Le stage se déroulera au sein de l'équipe de R&D développant Frama-C, qui est programmé en *OCaml*.

Objectifs

L'une des analyses les plus avancées de Frama-C est le greffon Value [KKP+15], qui calcule par interprétation abstraite une sur-approximation des comportements possibles d'un programme. Pour cela, le greffon (sur-)approxime l'ensemble des valeurs atteignables pour chacune des variables du programme : scalaires, agrégats, tableaux.

Frama-C/Value utilise deux approches pour représenter les tableaux. Lorsque le contenu du tableau est connu précisément, chaque case est représentée indépendamment. En revanche, si les valeurs possibles pour chaque case ne sont connues qu'imprécisément, la structure de données *résume* le contenu en une seule case. Cette technique d'analyse est bien connue dans la littérature, mais Value se distingue sur deux points :

- le passage d'une représentation *in extenso* à une représentation résumée a lieu plus dynamiquement que dans la plupart des analyseurs ;
- la structure de données utilisée [BC11] est particulièrement riche, et permet de traiter de façon transparente d'autres fonctionnalités telles que les unions et les structs C.

Les résumés actuellement utilisés ne fonctionnent en revanche pas bien pour les tableaux de structs. L'utilisateur doit choisir entre une description complète du tableau, sur laquelle les opérations de lecture ou de mise à jour peuvent être lentes, ou une version très approchée, dans laquelle les contenus des différents champs de la structure sont mélangés. D'autres limitations existent pour les accès à des tableaux de tableaux, cette fois dûes à la façon dont sont représentées les *locations mémoires*.

L'objectif du stage sera de lever une de ces deux restrictions. Pour améliorer le traitement des tableaux de struct, le candidat définira une notion de *périodicité* dans la structure de données OCaml représentant actuellement les tableaux.

Ensuite, toutes les fonctions d'accès à la mémoire (lecture, écriture, union, ...) seront complétées pour prendre en compte ces plages de valeurs périodiques.

Pour les accès aux tableaux de tableaux, le stage consistera à définir des locations mémoires plus symboliques que celles existant actuellement, par exemple `&t[4..8][1..3].c`. Le stagiaire écrira ensuite les opérations abstraites correspondant au décalage et à l'union entre deux locations symboliques, ainsi qu'à la lecture et à l'écriture dans une location.

La structure de données utilisée pour représenter les tableaux dans Frama-C/Value étant très complexe, un bon esprit d'abstraction est attendu du candidat. Un minimum d'agilité avec les modulo et les congruences est également nécessaire.

Candidatures

Profil des candidats

- Bonnes connaissances en OCaml
- Intérêt pour l'analyse automatique de programmes réels
- Capacité de travail en équipe
- Une connaissance du C serait un plus pour écrire les exemples

Conditions : stage indemnisé, aide au logement possible, transports CEA en Île-de-France.

Contact : Boris Yakobowski (*prénom.nom@cea.fr*)

Les délais administratifs au CEA étant de 2 à 3 mois minimum, merci de prendre contact le plus tôt possible.

Références

[BC11] Richard Bonichon and Pascal Cuoq. A mergeable interval map. *Stud. Inform. Univ.*, 9(1) :5–37, 2011.

[KKP+15] Florent Kirchner, Nikolai Kosmatov, Virgile Prevosto, Julien Signoles, and Boris Yakobowski. Frama-C : A software analysis perspective. *Formal Asp. Comput.*, 27(3) :573–609, 2015.