L3 internship: Formalizing dictionaries in Coq

Guillaume Melquiond

Inria, LMF, Université Paris-Saclay

1 Context

The ERC Fresco project[1] aims at turning the Coq proof assistant into a competitive tool for doing verified computer algebra. One key component is the design of a dedicated programming language as well as some high-level data structures. This internship looks at the issue of associative data structures, *i.e.*, dictionaries.

People have formalized dictionaries in Coq for a long time, but they have mostly focused on implementations based on binaries trees, be they radix trees (*e.g.*, in CompCert) or AVL-balanced search trees (*e.g.*, in Coq's standard library). Indeed, trees are an inductive data structure that is especially well-suited for formalization in Coq.

The recent addition of primitive arrays to Coq paves the way of a new approach. Indeed, they make it possible to implement and formalize dictionaries implemented by hashtables. But, by virtue of all data structures (including arrays) being immutable in a formal system such as Coq, implementations that are sensible in the real world might exhibit different complexity properties once formalized in Coq.

2 Objectives

The goal of this internship is to investigate array-based implementations of dictionaries in Coq. Concretely, the objectives are as follows:

- survey the literature on (semi-) persistent arrays and hashtables,
- devise and formally verify some implementations of hashtables in Coq,
- evaluate these implementations and compare them to tree-based dictionaries.

3 Location

The internship will take place at the Formal Methods Laboratory[2] of Université Paris-Saclay, in the Inria Team Toccata, which is dedicated to writing tools for deductive program verification.[3]

4 Prerequisites

Knowledge of the Coq proof assistant, or of a similar formal system (*e.g.*, Lean), is highly recommended.

[1] `https://fresco.gitlabpages.inria.fr/`
[2] `https://lmf.cnrs.fr/`
[3] `https://toccata.gitlabpages.inria.fr/toccata/`