

# Compiling Complex Data Structures

## CODAS's ANR Project Final Deliverable

Laure Gonnord, 2023.

## 1 Introduction

The main objective of the CODAS project was to design a general framework for reasoning about covertly regular programs operating on complex data structures, in order to efficiently schedule and parallelize them. For this purpose we have used emerging ideas from three communities: compilation, static analysis, and rewriting; in particular:

- Task 1 was partly achieved via our work on Terms as intermediate representation; which was necessary for our term rewriting activities as well as our Algebraic Data Types (ADT) study.
- Task 2 and Task 3 was the main output of the Phd thesis of Paul lannetta, who proposed a language approach to lay terms (tree-like data structures) into arrays and benefit of further optimisations from the compiler. The compact representation expected in Task 3.1 is not fully defined, but the two Phd of Thais Baudon and Paul lannetta made two steps in this direction. Task 3.2 has been partly addressed in the GPCE '21 paper.
- We developed building blocks of the prototype planned in Task 4.

In conclusion, this project is a real success, and made major steps into bringing more non polyhedral data structures in the HPC world.

## 2 Results and publications

All publications in the context of the project were pushed on HAL, and can be accessed on the project webpage: [http://laure.gonnord.org/pro/research/anr\\_codas.html](http://laure.gonnord.org/pro/research/anr_codas.html). As for the TRS part of the proposal, a journal submission is still under review. New benchmarks have been performed, whose results are available at the page [https://www.dcs.bbk.ac.uk/~carsten/eval/parallel\\_complexity\\_journal/](https://www.dcs.bbk.ac.uk/~carsten/eval/parallel_complexity_journal/).

In addition, the following prototypes have been developed:

- An extension of the Aprove tool to compute lower and upper bounds of Term Rewriting Systems parallel complexity. We plan to publish publicly after the evaluation of the journal paper under submission (June 2023).
- REW, a compilation towards flat representations for trees/ADT <https://github.com/Drup/adtr/>
- RIBBIT, <https://gitlab.inria.fr/ribbit/ribbit>, a DSL to describe low level representation of ADTs and compile a pattern-matching decision procedure.

## 3 Opening toward future work and projects

The CODAS project has progressed into unexpected directions:

- the goal of polyhedral like compilation (task 4) for trees has not been completely reached, but the study of tree layouts has led to significant performance results. It has deeply inspired the Phd proposal of Thais Baudon on low-level representation of ADTs.
- the idea to work on non regular intermediate representations has led to an ANR proposal (SxC) in 2022, on high performance SIMD compilation of streaming languages recognizers. Unfortunately, this proposal has not been accepted, but we definitively believe that the CODAS results might be of great inspiration.