Synthesizing Benchmarks for Predictive Modeling

CGO'17 Code Generation and Optimization
Austin, Texas, USA

BEST PAPER AWARD

We mine thousands of repositories on GitHub for program fragments and apply deep learning techniques to automatically construct models for how humans write programs. We then sample the models to generate an unbounded number of runnable training programs, covering the program feature space ever more finely. The quality of generated code is such that even human developers struggle to distinguish our generated programs from hand-written code.

We use our generator for OpenCL programs, CLgen, to automatically synthesize thousands of programs and show that they improve the performance of state of the art predictive models by 1.27x. In addition, the fine covering of the feature space automatically exposes weaknesses in the feature design which are invisible with the sparse training examples from existing benchmark suites. Correcting these weaknesses further increases performance by 4.30x.

Deep Learning in Compilers

Chris Cummins

End-to-end Deep Learning of Optimisation Heuristics

PACT'17 Parallel Architectures and Compilation Techniques
Portland, Oregon, USA

We introduce a better way for building compiler heuristics. Using deep neural networks over raw program code, we learn heuristics entirely without using code features. The networks simultaneously construct appropriate representations of the code and learn how best to optimize, removing the need for manual feature creation. Further, we show that our neural nets can transfer learning from one optimization problem to another, improving the accuracy of new models, even if the optimization tasks are entirely dissimilar.

We compare the effectiveness of our automatically generated heuristics against ones with features hand-picked by experts. In two challenging domains the quality of our fully automatic heuristics surpass that of state of the art predictive models, providing 14% and 12% more performance respectively, with no human effort expended on designing features.

THE UNIVERSITY of EDINBURGH

https://chriscummins.cc

EPSRC Centre for Doctoral Training in Pervasive Parallelism