gem5/Z3/gcc/Clang/Redis Heap Fitness Landscapes

Evo* 2025 Late-Breaking Abstracts
23 April 2025

Humies $10000 prizes
Submit by **Friday 30 May**

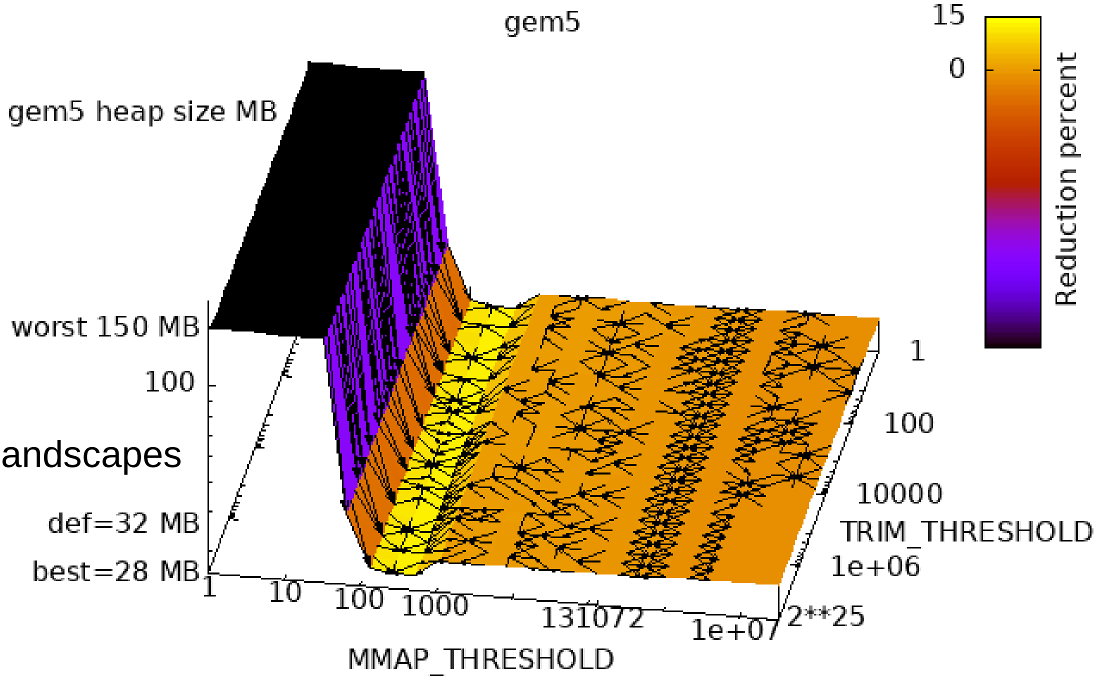
W. B. Langdon, UCL

GP bibliography needs your papers
http://gpbib.cs.ucl.ac.uk/

Can sometimes tune C++ dynamic memory.
Heap memory landscapes are smooth.

The gem5 C++ glibc Heap Fitness Landscape,
GI @ ICSE 2025, Ottawa
W.B. Langdon and B.R. Bruce.

gem5/Z3/gcc/Clang/Redis glibc Heap Fitness Landscapes
Evo* 2025 Late-Breaking Abstracts.
W.B. Langdon, J.Petke, D. Clark



23.4.2025

The gem5 C++ glibc Heap Fitness Landscape

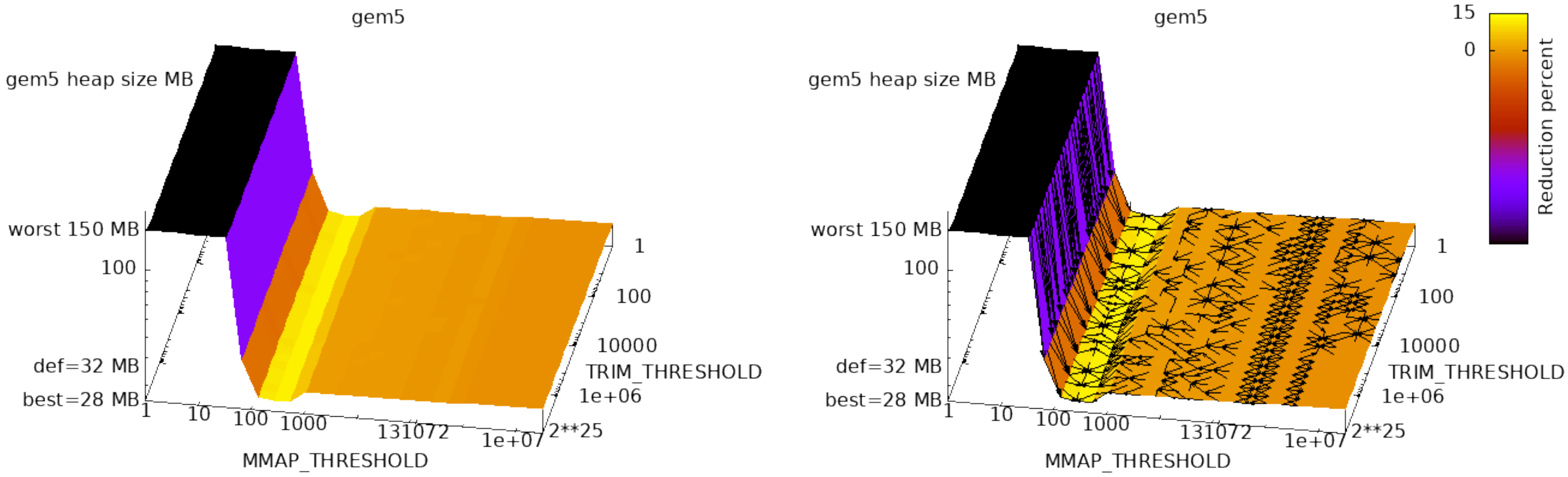
- GI 2025 gem5 landscape smooth saving 12%
  Evo* 2025 LBA try 4 more: smooth but little improvement
- What is gem5?
- What is GNU C malloc() Heap?
- What are Fitness Landscapes?
- Why glibc heap is a “nice” landscape
- Measuring memory: malloc_info(), massif, top
- Genetic Improvement: Magpie, CMA-ES
- New
  - Z3 Microsoft theorem prover
  - GNU g++ compiler
  - Clang LLVM compiler
  - Redis key-value store
- Heap differences: smooth but little gain

What is the GNU C malloc Heap?

- In C dynamic memory is created by malloc() etc.
- In g++ malloc is also used
- 37 /lib64/ld-linux-x86-64.so.2 --list-tunables
- g++ malloc 7 tuning parameters:
  - M_ARENA_TEST, M_ARENA_MAX, M_PERTURB, M_TOP_PAD
  - **M_MMAP_MAX, M_TRIM_THRESHOLD M_MMAP_THRESHOLD**
- Only last three relevant to single thread programs
- Total space 2**(32+64+64) = 2^{160} = 1.5 10^{48}

Why glibc heap is a “nice” landscape

- Although representation is 96 bits (2^{96}) only three continuous dimensions
- One dimension (M_MMAP_MAX) makes almost no difference, any reasonable value will do.
- Can navigate landscape near defaults (2^{16},2^{17},131072) assuming it is continuous
- Sample two dimensional grid at powers of two and half powers
- Add arrows to emphasis gradient to better (lower) values



Genetic Improvement of Parameters with Magpie

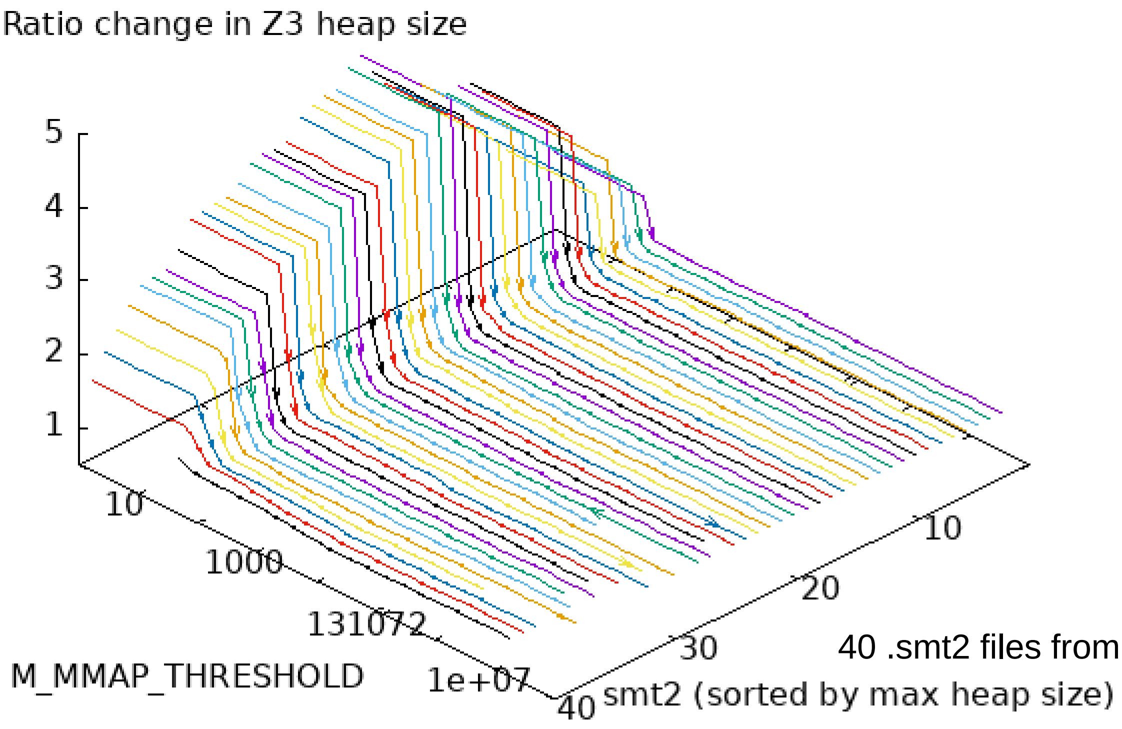
- Genetic Improvement typically applied to source code
- Improvements can be found by tuning parameters [EuroGP 2018]. Cf. deep parameter tuning
- Magpie can do both code and parameter tuning
- Magpie parameter file specify types of mutation
  - Geometric
    - set range 0 to 256x default,
    - set mutation default to be malloc default
    - mean: opps should have been default, was 2.56x default
- setenv environment variable to change malloc parameters
- Run with and without parameter changes
- Fitness = reduction in peak heap size
- 1000 fitness trials
- 12% reduction

New C++ Examples

- Use Valgrind’s Massif
- Z3 Microsoft SMT theorem prover
  - 600,000 lines of code
  - Bench mark Certora Prover example from SMT competition
- GNU g++ compiler
  - Compile largest MySQL source code
- Clang LLVM compiler
  - Fitness function as g++
- Redis key-value store (make MALLOC=libc)
  - 150,000 lines of code
  - Use redis-benchmark
- Heap landscape again smooth but little gain
- Is small improvement because they do not stress heap?

Microsoft Z3 on Certora Prover benchmarks

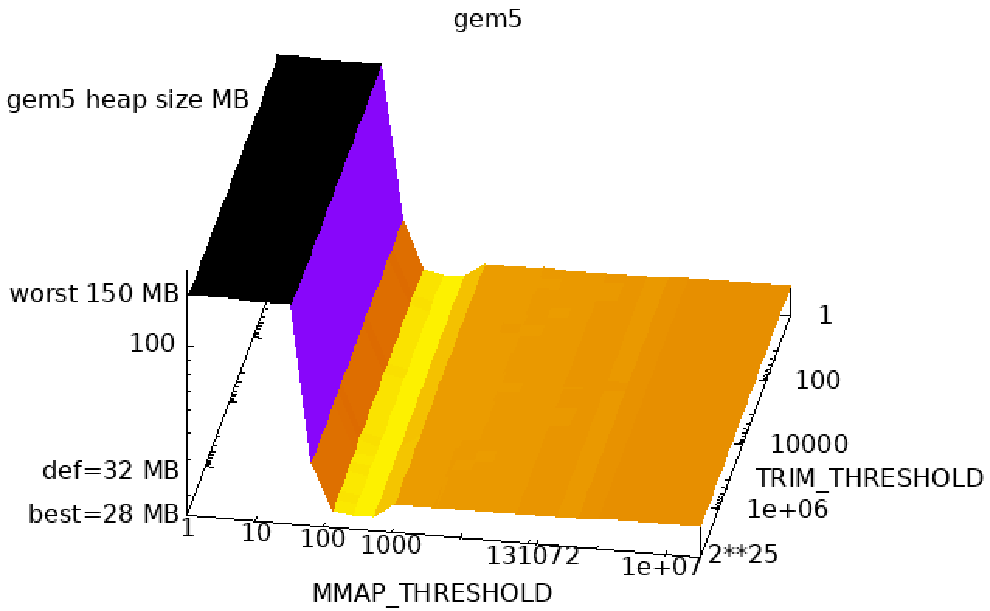
- Z3 Massif peak heap
- Smooth fitness landscape (plot one dimension mmap)
- Best 1.5% improvement



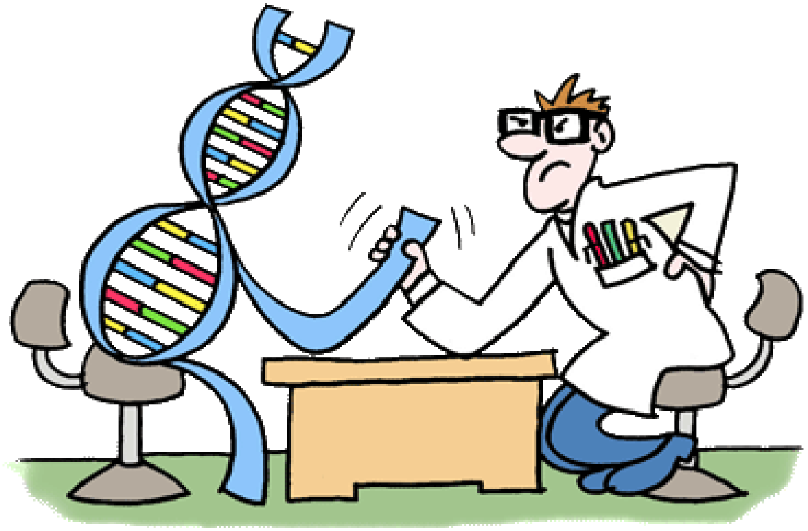
40 .smt2 files from SMT-COMP 2024

Conclusions

- Not all search problems are hard.
- Genetic Improvement can be applied to software parameters as well as code. Eg glibc 37 run time parameters. Use any optimiser
- gem5 is a million+ lines but C++ 7 dimensional new/delete landscape is smooth, collapses to essentially one dimension broad good fitness valley (4 10^{17} solutions) large basin of attraction. Gives 11% heap reduction without loss of speed
- Other non-trivial C++ programs have similar smooth landscapes but tuning GNU glibc malloc gives only marginal improvement
- Magpie can tune parameters as well as multi-language code



4e17 = 400 Quadrillion



Human-Competitive results $10,000 prizes
https://www.human-competitive.org

Email your entry to goodman@msu.edu
by **Friday 30 May**

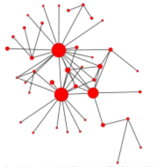
The Genetic Programming Bibliography

http://gpbib.cs.ucl.ac.uk/

17586 references, 17000 authors

Make sure it has all of your papers!

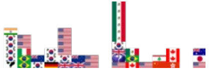
E.g. email W.Langdon@cs.ucl.ac.uk or use | Add to It | web link



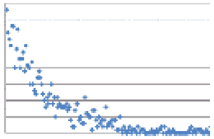
Part of gp-bibliography 84 40 Revision:1.1794 29 May 2011

Co-authorship community.
Downloads

Downloads by day



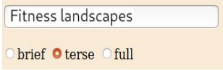
A personalised list of every author’s GP publications.



Your papers

blog

Googling GP bibliography, eg:
Development and learning site:gpbib.cs.ucl.ac.uk



Fitness landscapes

○ brief ● terse ○ full

Text search