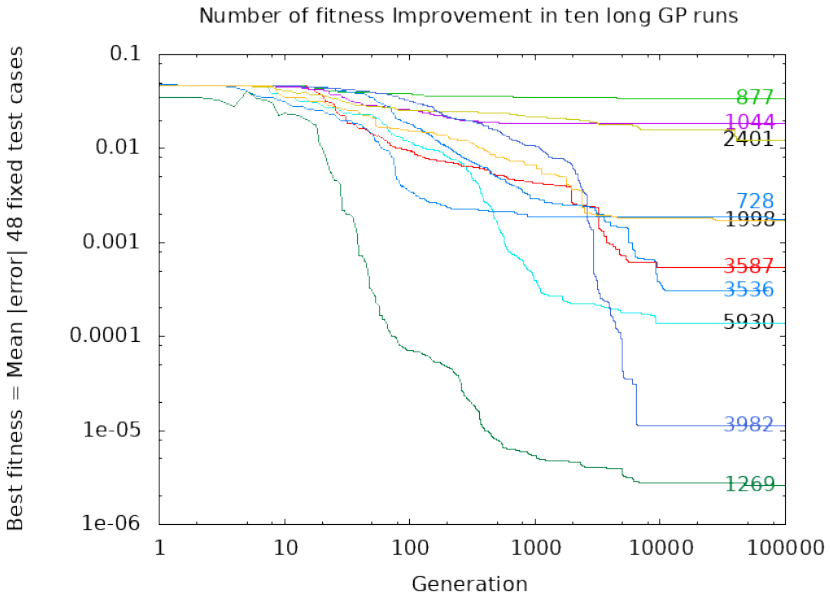
Long-Term Evolution Experiment with Genetic Programming

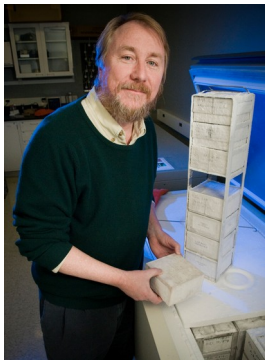
W. B. Langdon and W. Banzhaf. 2022. ALJ 28(2) 173–204
https://doi.org/10.1162/artl_a_00360

W. B. Langdon



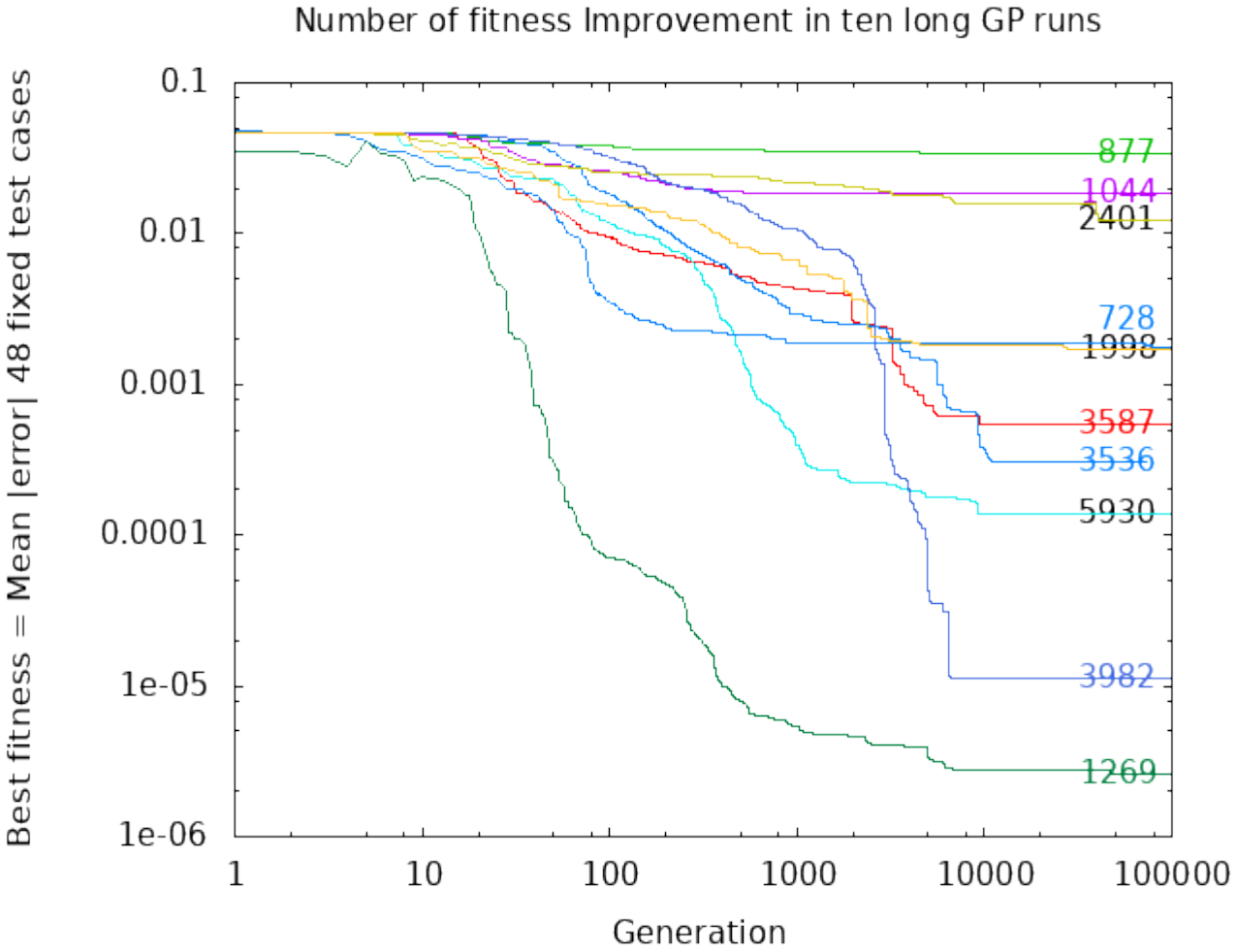
Summer 2022

Long-Term Evolution Experiment with Genetic Programming



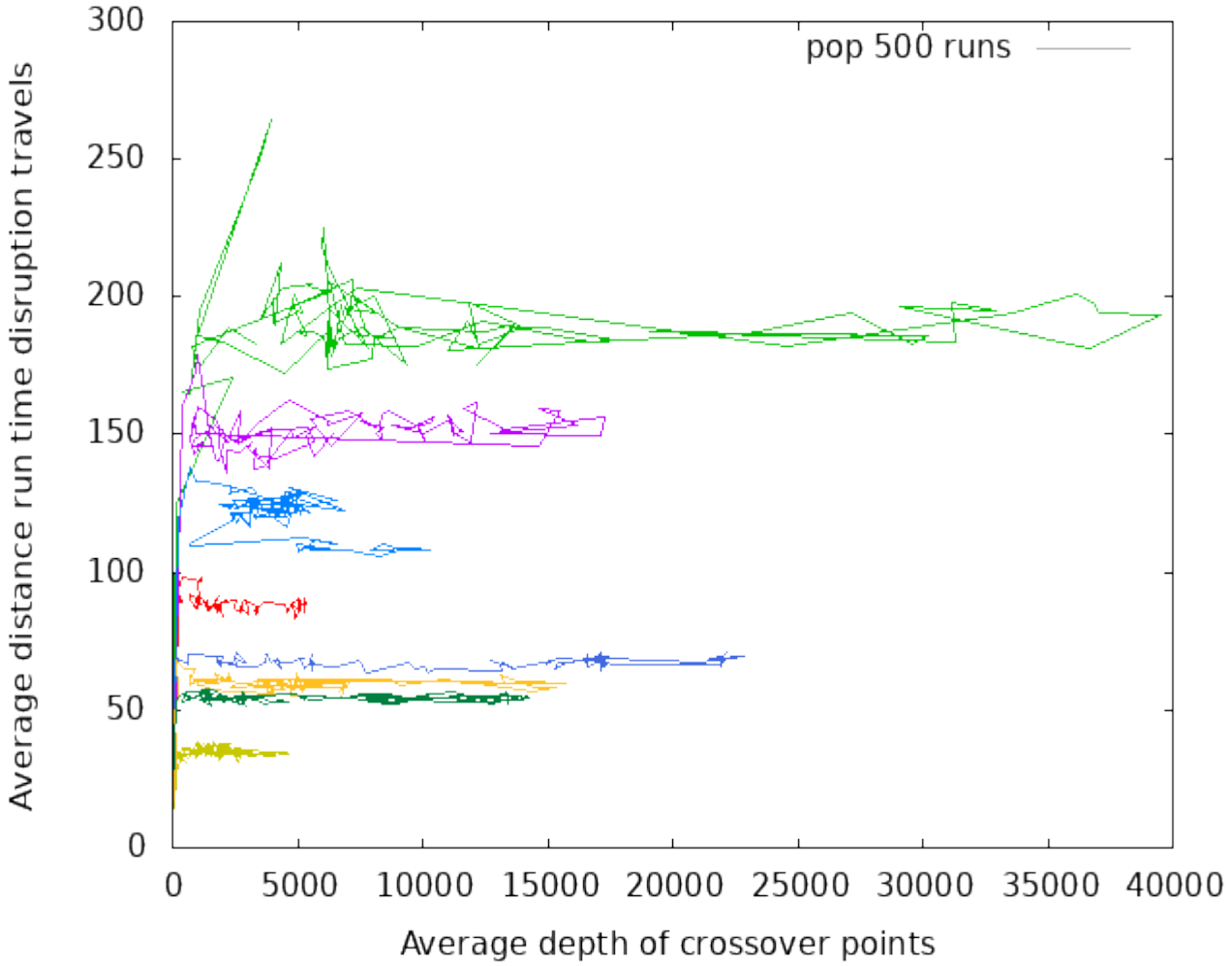
- Paper looks at floating point symbolic regression tree over up to a million generations.
- Rich Lenski's Long Term Evolution Experiment LTEE 75000 generations of bacteria E.coli
- In GP fitness improvement continues but slows
- In big trees crossover genetic disruption gives limited run time disruption due to tree depth
- Rate of disruption inversely proportional to size
- Only a fraction of fitness changes are improvements
- But fitness improvement continues

Fitness improvement continues but slows



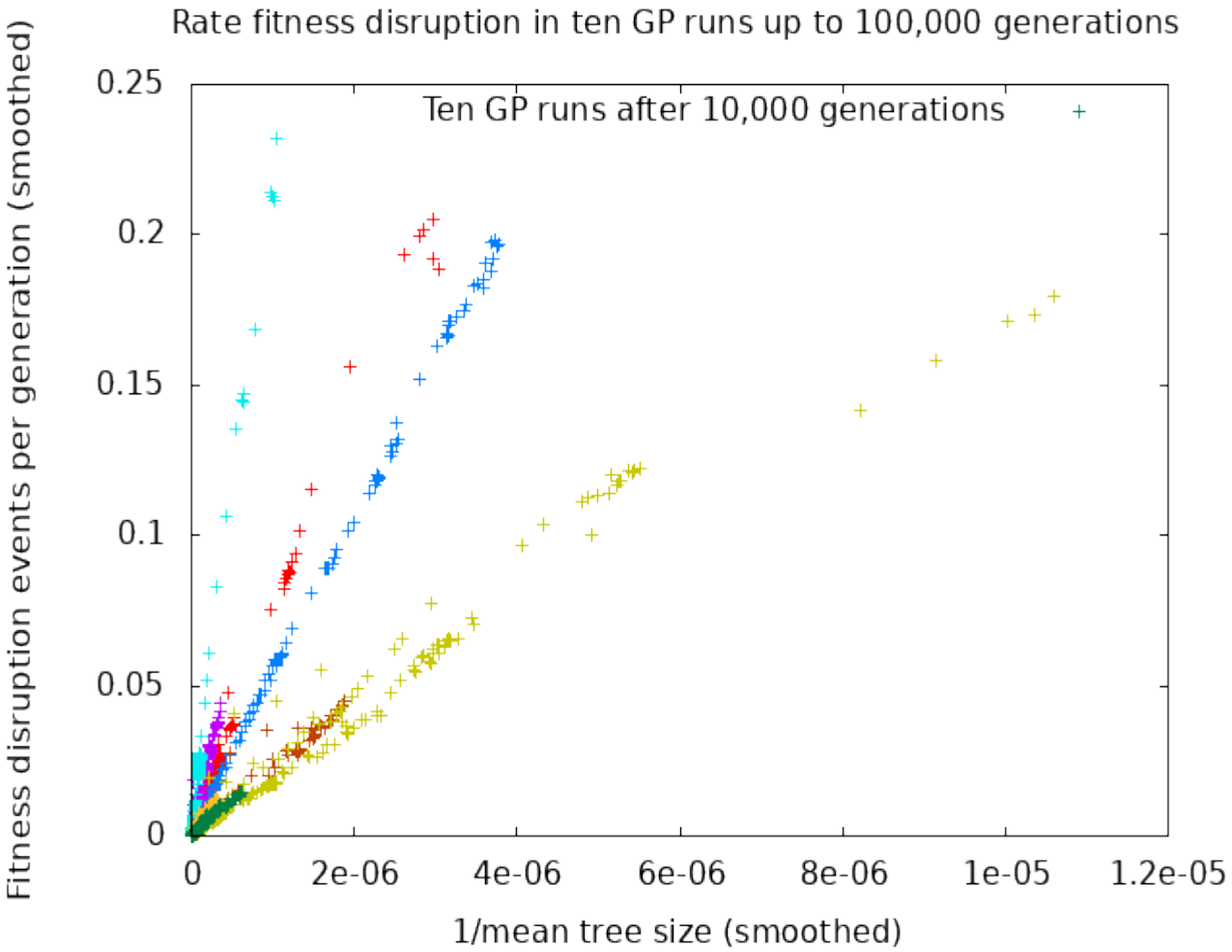
Evolution of mean absolute error in ten runs of Sextic polynomial with population of 500. Runs to 100 000 generations (2 stopped early). Thousands of fitness improvements found. Note log scales.

Crossover syntax changes have limited run time impact



Impact of syntax change propagates only a limited numbers of nested arithmetic operations. (Evolution gives different behaviours between runs but in all cases failed disruption propagation stops the impact of almost all crossovers.)

Crossover fitness disruption 1/size



When trees are large chance crossover hitting sensitive area near root falls in proportion to tree size. (Evolution gives different ratios between runs.) Smoothed by taking running averages over a 100 generations.

Fraction Improvements

- Rate at which crossover produces children with different fitness to their parents is chance of hitting sensitive region by root node, and so falls in proportion to tree size
- Fraction of changes which are beneficial varies between runs and as the population evolves.
- Evolution continues to find small improvements even after many thousands of generations

Total fitness changes generations 10,000 to 100,000

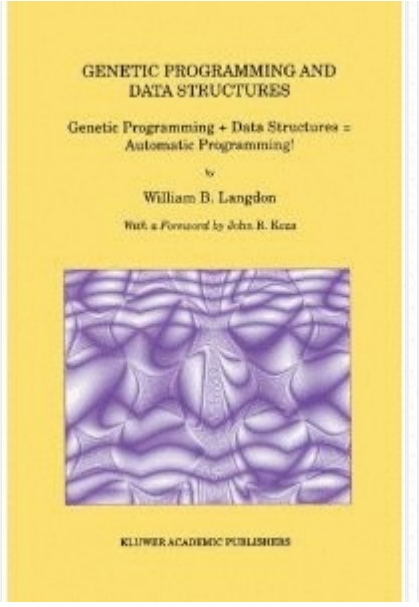
Run	different	better	ratio
1	359843	222	0.00062
2	127422	106	0.00083
3	452932	317	0.00070
4	232026	102	0.00044
5	182222	35	0.00019
6	132879	209	0.00157
7	386257	497	0.00129
8	57661	9	0.00016
9	85591	121	0.00141
10	97316	44	0.00045

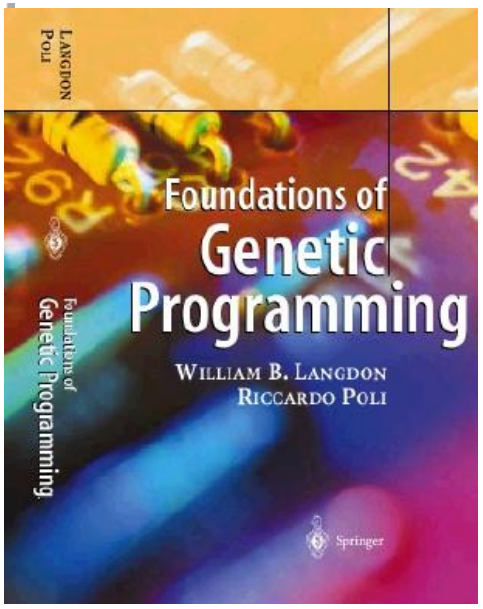
Conclusion Evolution continues

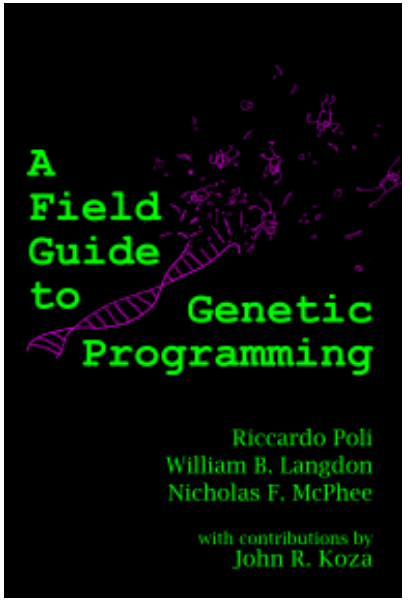
1) Huge GP runs (up to a million generations, trees of up to two billion nodes) are feasible and have lead to improved implementation (*equivalent* to a trillion GP operations per second) and insights.
2) Information theory shows even in pure floating point arithmetic, failure of disruption caused by deep crossover and mutation to impact fitness is inevitable, and leads with current GP to exploration falling in proportion to tree size.
3) With higher entropy loss function sets (eg Boolean and conditionals) we expect a more sever brake on evolution.
4) Nevertheless evolution continues and fitness improvements continue to be found even after many thousand of generations.

Genetic Programming

W. B. Langdon







etc

In software engineering failed disruption propagation will be more severe due to high entropy loss but offset by side effects. Meaning the optimal test oracle needs to be close to the source or error.

This work was inspired by conversations at Dagstuhl Seminar 18052 on Genetic Improvement of Software.

Parallel GPquick code is available via
http://www.cs.ucl.ac.uk/staff/W.Langdon/ftp/gp-code/GPinc.tar.gz

References

[1] William B. Langdon and Wolfgang Banzhaf. 2022. Long-Term Evolution Experiment with Genetic Programming. Artificial Life 28, 2 (2022).

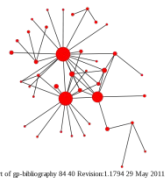
[2] Richard E. Lenski et al. 2015. Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with Escherichia coli. Proc. R. Soc. B: Biol. Sci. 282, 1821 (22 Dec 2015). http://dx.doi.org/10.1098/rspb.2015.2292

The Genetic Programming Bibliography

15586 references, 15000 authors

Make sure it has all of your papers!

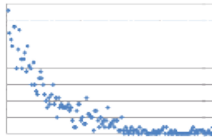
E.g. email W.Langdon@cs.ucl.ac.uk or use | Add to It | web link



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