**Genetic Improvement Workshop 2022 Proposal**

Bobby R. Bruce  
University of California, Davis  
bbruce@ucdavis.edu

Aymeric Blot  
University College London  
a.blot@cs.ucl.ac.uk

Vessna Nowack  
Lancaster University  
v.nowack@lancaster.ac.uk

Emily Winter  
Lancaster University  
e.winter@lancaster.ac.uk

W. B. Langdon  
University College London  
w.langdon@cs.ucl.ac.uk

Justyna Petke  
University College London  
j.petke@ucl.ac.uk

**ABSTRACT**

Genetic improvement is the process of using automated search to improve existing software. It has successfully been used to fix bugs, transplant functionality from one system to another, improve predictions, and reduce software’s runtime, energy and memory consumption; all without the necessity of costly human labour. Research within this field has already won three “Humies”. Despite impressive findings, genetic improvement is a relatively new field of research, with many opportunities to improve the state-of-the-art.

The genetic improvement workshop series has been successfully run since 2015 when it was co-located in Madrid with the top conference on evolutionary computation, GECCO. Each year it has attracted between 5 to 14 accepted papers with participant numbers, reaching almost 50 in its 2017 edition. In 2018 the workshop was held, for the first time, both at GECCO and at ICSE, attracting 33 registrations at ICSE that year. Building upon this success, the workshop was held at GECCO and ICSE, attracting up to 50 registrations at ICSE. (Note workshop registrations at ICSE are not included in the conference registration but must be paid for separately.)

Having observed keen interest from evolutionary computing and the software engineering communities, we believe a 11th workshop at GECCO will continue to build interest, collaboration, and discussion.

**KEYWORDS**

Genetic Improvement, GI, Automated Program Repair, APR, Search-Based Software Engineering, SBSE, Software Improvement

**ACM Reference Format:**


**1 MOTIVATION AND OBJECTIVES**

Genetic improvement (GI) is a new area of research which builds on the idea that all software exists within a search space and that this space may be traversed towards better products. For example, GI was able to find a code variant of a complex bioinformatics software system 70 times faster than the original [11]. Changes involving another bioinformatics system, which led to 3-fold speed-up, have since been adopted into production code [12]. Work on software transplantation using GI has won multiple awards, including an ACM distinguished paper award at ISSTA 2015 [3], a distinguished artifact award this year at ICSE [7], and as a winner of the 2021 SSBSE Challenge track [6]. Work into GI has also been awarded three prestigious cash prizes, “Humies” [2], for human-competitive results produced by evolutionary computation [3, 4, 19]. Software transplantation work was also widely covered by the media [1, 21] and inspired further research into software transplantation [20]. Another example has been optimisation of energy consumption using a similar evolutionary approach [5], as well as other search techniques [8]. The 2019 SSBSE Challenge track winner demonstrated the possibility of using a freely available GI framework to solve common software engineering problems [17]. The GI survey shows the growth of the area both in the evolutionary and software engineering communities [18] (see also Figure 1).
GI @ GECCO 2020 Workshop Proposal Submission

(1) Title:
11th International Workshop on Genetic Improvement
GI @ GECCO 2022

(2) The scope and content is given in the last page (attached cfp.pdf)
(3) http://geneticimprovementofsoftware.com/events/gecco2022
(4) We anticipate 35–50 will attend.
(5) Wes Weimer, the inventor of automated bug fixing, will give the keynote.
(6) We anticipate four (110 minute) sessions.

(7) GI @ GECCO 2022 will be the eleventh international GI workshop. The first GI workshop was held at GECCO 2015 (Madrid), and has been held annually since. In 2018, 2019, and 2020, workshops were held at both GECCO and at the International Conference on Software Engineering (ICSE). In 2021 one workshop was held at ICSE. We wish to hold the 2022 workshop exclusively at GECCO, alternating between the GECCO and ICSE venues in the future. A Dagstuhl Seminar on Genetic Improvement was held in 2018. There were four CREST open workshops (COWs) on GI at UCL so far. There have also been several tutorials on GI held at venues such as GECCO, and the International Conference on Automated Software Engineering (ASE 2020), as well as special sessions at CEC.

2015 • 1st International Workshop on Genetic Improvement (GI@GECCO 2015, Madrid)
http://www.sigevo.org/gecco-2015/workshops.html#gi
2016 • 45th CREST Open Workshop on Genetic Improvement - 25 and 26th January 2016
http://crest.cs.ucl.ac.uk/cow/45/
• GI@GECCO 2016: 2nd International Workshop (Denver)
http://gecco-2016.sigevo.org/index.html/Workshops.html#id_Genetic%20Improvement%20Workshop
• CEC 2016 special sessions on Genetic Improvement of Software (Vancouver)
http://www.cs.ucl.ac.uk/staff/W.Langdon/cec2016/
2017 • The 50th CREST Open Workshop - Genetic Improvement
http://crest.cs.ucl.ac.uk/cow/50/
• GI@GECCO 2017: 3rd International Workshop (Berlin)
http://gecco-2017.sigevo.org/index.html/Workshops.html#id_Genetic%20Improvement%20Workshop
2018 • Dagstuhl Seminar 18052 on Genetic Improvement of Software
https://www.dagstuhl.de/en/program/calendar/sempn/#semmr=18052
• GI@ICSE 2018: 4th International Workshop (Gothenburg)
http://geneticimprovementofsoftware.com/events/papers/icse2018
• 5th edition GI@GECCO 2018 (Kyoto)
http://www.cs.stir.ac.uk/events/gecco-gi-2018/
2019 • The 61st CREST Open Workshop - Statistics, Learning, Genetic Improvement and Testing for Programs
http://crest.cs.ucl.ac.uk/cow/61/
• GI@ICSE 2019, 6th International Workshop on Genetic Improvement (Montreal)
http://geneticimprovementofsoftware.com/events/icse2019
• GI@GECCO 2019, 7th International Workshop on Genetic Improvement (Prague)
https://workshop07.gi-workshops.org/
2020 • 62nd CREST Open Workshop on Program Repair and Genetic Improvement
http://crest.cs.ucl.ac.uk/cow/62/
• GI@ICSE 2020, 8th International Workshop on Genetic Improvement (Seoul/Internet)
http://geneticimprovementofsoftware.com/events/icse2020
GI@GECCO 2020, 9th International Workshop on Genetic Improvement (Cancun/Internet)
https://gi-gecco-20.gi-workshops.org/
• Special session on Genetic Improvement (WCCI 2020)
http://geneticimprovementofsoftware.com/events/wcci2020
• Also GI tutorials at GECCO 2020, ASE 2020 and PPSN 2020
2021 • GI@ICSE 2021, 10th International Workshop on Genetic Improvement (Madrid/Internet)
http://geneticimprovementofsoftware.com/events/icse2021
• Also GI tutorial at GECCO 2021
1.1 Goals and Outcomes

Through the genetic improvement workshop series at GECCO, research on GI has grown within the evolutionary computation community. We hope to continue its expansion both within the evolutionary computation and software engineering community. Since it is a new and exciting research area, there continues to be many unanswered questions. Many of these, such as the best type of test suites for automated bug fixing tools, have already been asked by those in the program repair field [16, 24]. It is still unknown, for instance, when it is beneficial to use GI [14]; what other optimisation objectives one might target (there’s little work on memory optimisation [23], for instance); what ‘good’ mutation operators are; whether mutation operator effectiveness differs depending on the optimisation objective or programming language. Research possibilities and directions are endless. Therefore, by bringing together researchers doing work in GI as well as GI enthusiasts we hope to move the field forward by sharing experiences and exchanging ideas.

2 FORMAT AND REQUIRED SERVICES

The main goal of the workshop is to facilitate discussion, hence we will retain the successful elements from earlier GI workshops. Although we hope it will be possible to hold GI @ GECCO 2022 in person in Boston, it is wise to prepare for the possibility that the COVID-19 Pandemic will require it to be held (in part) online. All of the proposed organisers have either hosted or assisted in the virtualisation of the 2021 ICSE workshop, or have considerable experience with online events, should this be necessary again.

2.1 Plans for Generating Discussion

Keynote. In previous years we have been able to secure the services of fantastic keynote speakers such as Prof. Wolfgang Banzhaf, who gave the keynote on “Genetic Programming between Artificial Intelligence and Software Engineering”. In 2018, Claire Le Goues gave a keynote (author of the ICSE 2009 paper “Automatically finding patches using genetic programming” [22], which was awarded the Most Influential Paper award at ICSE 2019). At GI 2019 Nadia Alshahwan gave a keynote on “Industrial experience of Genetic Improvement in Facebook”. For GI 2020, despite the COVID-19 pandemic we secured Prof. Mark Harman whose keynote “WES: Agent-based User Interaction Simulation on Real Infrastructure” not only described possibilities of implementing GI at scale in Facebook, but also served as an invitation for research proposals. At the last GI workshop the keynote was given by Prof. Stephanie Forrest. Prof. Westley Weimer has agreed to give the GI @ GECCO 2022 keynote. Wes is a full professor at the University of Michigan in Ann Arbor. His research covers a wide range including medical imaging, automated program repair and automatically improving software. As in other years, we expect the keynote talk will trigger many interesting questions and discussion points, which will be followed up during the rest of the workshop.

Research Papers. We invite submissions that discuss recent developments in all areas of research, theory and applications of Genetic Improvement. The workshop will provide an opportunity for researchers interested in GI to exchange ideas and find out about current research directions in the field. Topics of interest include, but are not limited to, using genetic improvement to automatically:

- improve runtime efficiency
- decrease memory consumption
- decrease energy consumption
- transplant new functionality
- specialise software
- improve predictions
- port software
- fix bugs

We are also interested in research into approaches to carry out, as well as understand, genetic improvement. Such research avenues may include studies into loop perforation, genetic programming, constraint-based synthesis, search based software engineering, program analysis, testing (including flaky tests), data mining, machine learning, and entropy and information theory. Genetic improvement combines various fields of research, hence exchanging ideas of how insights from one area fit within GI will lead to fruitful discussions.

Position Papers. From our experience in organising Genetic Improvement workshops, we have noticed that position papers facilitate participation. This gives researchers opportunity to be bold in their ideas and ask questions about the future of GI. We find this very useful for PhD students who then gain useful feedback for their initial ideas for research that are presented at the workshop. We also very much welcome research and position papers that consider GI techniques in industrial settings.

2.2 Workshop Duration

We plan to hold the workshop over one day (4 sessions). The first session will be devoted primarily to Prof. Westley Weimer’s keynote and subsequent discussion. The following three sessions will contain research and position paper presentations, with a dedicated discussion session at the end. The workshop will also have a prize giving ceremony. The workshop will be followed by a social event in Boston.

2.3 Workshop Format

We will start with a keynote and then intersperse research and position papers throughout the day, based on topics of the accepted
papers. We anticipate about 12 papers accepted, based on past experience. Each paper session will have a mix of research and position papers. A preliminary schedule is outlined below in Table 1. As organisers, we will chair all the sessions and facilitate discussion throughout the workshop.

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00am - 9:10am</td>
<td>Opening Remarks</td>
</tr>
<tr>
<td>9:10am - 10:10am</td>
<td>Wes Weimer Keynote</td>
</tr>
<tr>
<td>10:10am - 10:25am</td>
<td>Q &amp; A</td>
</tr>
<tr>
<td>10:25 - 10:50</td>
<td>Research paper presentation</td>
</tr>
<tr>
<td>10:50am - 11:10am</td>
<td>Break</td>
</tr>
<tr>
<td>11:10am - 13:00pm</td>
<td>Paper Session (4 papers)</td>
</tr>
<tr>
<td>13:00pm - 2:00pm</td>
<td>Lunch</td>
</tr>
<tr>
<td>2:00pm - 3:50pm</td>
<td>Paper Session (3 papers)</td>
</tr>
<tr>
<td>3:50pm - 4:10pm</td>
<td>Break</td>
</tr>
<tr>
<td>4:10pm - 4:55pm</td>
<td>Paper Session (4 papers)</td>
</tr>
<tr>
<td>4:00pm - 5:45pm</td>
<td>Discussion, including student feedback</td>
</tr>
<tr>
<td>5:45pm - 5:55pm</td>
<td>Awards</td>
</tr>
<tr>
<td>5:55pm - 6:00pm</td>
<td>Closing</td>
</tr>
<tr>
<td>6:00pm -</td>
<td>Social Event</td>
</tr>
</tbody>
</table>

### 3 TARGET AUDIENCE
We expect little previous knowledge of genetic improvement from the workshop attendees. Given that it is a new area of research we want to attract both senior as well as junior researchers. It should be interesting to researches using evolutionary techniques to improve software. It should also be attractive to novices looking for inspiring new research directions.

#### 3.1 Industrial Relevance
GI techniques have only recently been adapted into industry. The key examples being adoption of GI-evolved changes to a bioinformatics software system called BarraCUDA, giving up to a 3 times speed-up [12], RNAfold [13], incorporation of an automated GI-based software repair process in the Janus Manager management system for medical rehabilitation centres in Iceland [9], Facebook’s Sapfix [15] and at Bloomberg [10]. We hope that these success stories lead to further interest from industry and wider adoption in general.

### 4 PARTICIPANTS
We plan to hold an open workshop. In the previous editions of the Genetic Improvement workshop series there were up to 14 accepted papers. Participant numbers started with over 30 in the first year, with almost 50 in the 2017 edition of the workshop at GECCO. At ICSE-2018 (where registrations are charged) there were 33 registrations, at ICSE-2019 we had 40 and last year (ICSE-2020) it was 50, with about 40 attending online via YouTube. 34 paid for registration at ICSE 2021. We expect a similar number (35–50) of registrations at GECCO 2022.

As in previous years, we will advertise the workshop on relevant mailing lists, both evolutionary computation and software engineering fora, such as the SSBSE mailing list, EvolutionaryComputation, Metaheuristics. We will also send individual emails to potential participants and submitters. We have a list of such email addresses from previous years.

### 5 PROCEEDINGS
We expect to accept about 10–13 papers. We anticipate half will be full papers, while the rest will be position papers. Research papers will be limited to 8 pages, with position papers limited to 2 pages. As with 2021, in the event of the pandemic, we expect the keynote and presentation videos to be available, e.g. via YouTube. Similarly, there will be a post workshop write up, e.g. in SIGEvolution.

#### 5.1 Review Process
As with last year, we intend to use HotCRP for the review process. We had a large PC in 2021, consisting of 12 members, whose expertise ranged from software engineering to evolutionary computation, both from academia and industry. The PC members from last year were:

- Ayaz Akram, UC Davis
- Afnan AlSubaieh, King Saud University based in Riyadh, Saudi Arabia
- Gabin An, KAIST, Korea
- Erik Fredericks, Grand Valley State University
- Saemundur Haraldsson, University of Stirling
- Mohamoud Bakari, University of Adelaide
- Eric Schulte, GrammaTech
- Christopher Timperley, Carnegie Mellon University
- Leonardo Trujillo, Instituto Tecnologico de Tijuana
- Emily Winter, Lancaster University
- Jifeng Xuan, Wuhan University
- Yuan Yuan, Michigan State University
- Mohamoud Bakari, University of Adelaide
- Eric Schulte, GrammaTech
- Christopher Timperley, Carnegie Mellon University
- Leonardo Trujillo, Instituto Tecnologico de Tijuana
- Emily Winter, Lancaster University
- Jifeng Xuan, Wuhan University
- Yuan Yuan, Michigan State University

We will make sure to balance the expertise of the PC in the 2022 edition as well. Each paper will get at least 3 reviews. We will follow the double-blind reviewing model. Acceptance will only depend on the quality of the contribution, based on reviews. Ability to create discussion will be the main criterion for position papers, as it was in previous editions of the workshop.

### 6 WEBSITE: GENETICIMPROVEMENTOFSOFTWARE.COM
All information about the workshop will be provided via the website. It contains links to past workshop editions, a general FAQ, and resources on genetic improvement. See http://geneticimprovementofsoftware.com/events/gecco2022 etc. The web pages are managed by Aymeric Blot.

### 7 WORKSHOP HISTORY
The Genetic Improvement workshop series has been co-located with the top conference in evolutionary computation, namely: the International Conference on Genetic and Evolutionary Computation (GECCO). Since 2015, the workshop has attracted around 30 to 50 participants. In the three years that it has been held at ICSE the participation numbers were: 33, 40, 50 and 34, respectively. Information on previous editions of this workshop can be found on the dedicated website (Section 6).
8 ORGANISERS’ BIOS

**Bobby R. Bruce** Bobby R. Bruce is a Project Scientist at UC Davis where he primarily works on the Gem5 computer architecture simulator. Prior to UC Davis, Bobby carried out research into the automatic optimization of Java bytecode at UCLA. In 2018 Dr. Bruce gained his PhD with Justyna Petke on Genetic Improvement, particularly automated improvement of software non-functional properties such as energy efficiency. His research interests are centred around Search-based Software Engineering, and its application to improving software performance.

**Vesna Nowack** Since she gained her PhD in Software Engineering in 2016 from the Universitat Politècnica de Catalunya in Barcelona, she has conducted research in supercomputing (Spain) and taught robotics in Germany. Her recent research has been on APR in the UK, including 12 months with Bloomberg (London) published this summer as “On the Introduction of Automatic Program Repair in Bloomberg” by IEEE Software [10]. She is now a Senior Research Assistant at Lancaster University where she continues her work on using GI to automatically fix bugs.

**Aymeric Blot** is a Research Associate conducting research in genetic improvement at the CREST and SOLAR groups in University College London. He received in 2018 a doctorate from the University of Lille following work on automated algorithm design for multi-objective combinatorial optimisation. His research focuses on strengthening GI techniques using knowledge from automated machine learning, algorithm configuration, and evolutionary computation. He maintains and evolves the community website on genetic improvement.

**Emily Winter** is a Senior Research Associate at Lancaster University, specialising in the human and socio-technical aspects of software engineering. She works on the Fixie Project: Exploiting Defect Prediction for Automatic Software Repair, investigating developer needs and preferences for how they interact with an Automatic Software Repair tool. As part of her research, she is seconded as a contractor to Bloomberg LP [10].

**William B. Langdon** His PhD was the first book to be published in John Koza and Dave Goldberg’s book series. He has previously run the GP track for GECCO 2001 and 2014 and was programme chair for GECCO 2002 having previously chaired EuroGP for 3 years. More recently he edited FOGA 2011 and run the computational intelligence on GPUs (CIGPU 2008–2012) and EvoPAR workshops. His books include A Field Guide to Genetic Programming, Foundations of Genetic Programming and Advances in Genetic Programming III. He also maintains the genetic programming bibliography. His current research uses GP to genetically improve existing software, CUDA, search based software engineering and Bioinformatics. He co-organised GI 2015, GI 2017, GI 2018, GI 2019, GI 2021 and the 2018 Genetic Improvement of Software Seminar at Dagstuhl.

**Justyna Petke** is a Principal Research Fellow and Proleptic Associate Professor, conducting research in genetic improvement. She has a doctorate in Computer Science from University of Oxford and is now at the Centre for Research on Evolution, Search and Testing (CREST) at University College London. Her work on genetic improvement was awarded a Silver and a Gold ‘Humie’ at GECCO 2014 and GECCO 2016 as well as an ACM SIGSOFT Distinguished Paper Award at ISSTA 2015. She was the PC co-Chair for the International Symposium on Search-Based Software Engineering in 2017. She also organised seven Genetic Improvement Workshops. She currently serves on the editorial board of the Genetic Programming and Evolvable Machines (GPEM) and Empirical Software Engineering (EMSE) journals.
REFERENCES


