What GP is and when you should use it

- Evolutionary algorithms that optimise programs
- Useful for discovering novel solutions to problems
- Or solutions to problems you don’t know how to solve

Introduction to tree-based GP:

- Initialisation, crossover, mutation
- Symbolic regression
- Conditional execution
- Handling types
- Bloat and bloat avoidance
Can GP expressions be “pretty printed”?  
- Yes!
- There’s also a gawk script for circular trees: [http://www0.cs.ucl.ac.uk/staff/W.Langdon/gp2lattice/gp2lattice.html](http://www0.cs.ucl.ac.uk/staff/W.Langdon/gp2lattice/gp2lattice.html)
Can GP be used for medical image analysis?

- Yes, there are quite a few examples of this
- Google “genetic programming medical image”
SL Smith, Cartesian Genetic Programming and Its Application to Medical Diagnosis, *IEEE Computational Intelligence Magazine*, November 2011.
Improving 3D Medical Image Registration CUDA Software with Genetic Programming

William B. Langdon, Marc Modat, Justyna Petka, Mark Harman
Dept. of Computer Science, University College London Gower Street, WC1E 6BT, UK
W.Langdon@cs.ucl.ac.uk

ABSTRACT
Genetic Improvement (GI) is shown to optimise, in some cases by more than 35%, a critical component of healthcare industry software across a diverse range of six NVIDIA graphics processing units (GPUs). GP and other search based software engineering techniques can automatically optimise the current rate limiting CUDA parallel function in the Nifty Reg open source C++ project used to align or register high resolution nuclear magnetic resonance (NMRI) and other diagnostic NIfTI images. Future Neurosurgery techniques will require hardware acceleration, such as GPGPU, to enable real time comparison of three dimensional in theatre images with earlier patient images and reference data. With millimetre resolution brain scan measurements comprising more than ten million voxels the modified kernel can process in excess of 3 billion active voxels per second.

Categories and Subject Descriptor: 1.2.8 [search]: heuristic

Figure 1: Performance of modified reg_spline_getDeformationField3D CUDA kernel after optimisation

http://dl.acm.org/citation.cfm?id=2598244
Evolvability

This is the capacity for a program to improve its fitness as a result of an evolutionary process (i.e. mutation and recombination).

For genetic programming, there’s little value in being theoretically able to express a program if it can not be discovered by evolution.
Expressiveness

This is the capacity for a program representation to express different kinds of behaviours.

For genetic programming, you can’t evolve a program if you can’t express it.

In practice, there is often a trade-off between expressiveness and evolvability.
Genetic Programming: Memory, Loops and Modules

Dr. Michael Lones
Room EM.G31
M.Lones@hw.ac.uk
There are various ways of adding memory to GP
  ▶ However, in practice these are not widely used

Consider the approach used by Astro Teller [1994]
  ▶ This introduces a memory, of width M:

```
0 1 2 3 4 5 6 7 ... M
```

  ▶ And 2 new functions:
    • READ(X) – read value from memory location X
    • WRITE(Y, X) – write value Y to memory location X
Example

```
progn3
```

```
WRITE
  2
  0

WRITE
  4
  READ
    2
    0

READ
```

```
0 0 0 0 0 0 0 0 0 0 0 0
```
Write 2 into location 0
Example

Write 4 into the location number read from location 0
Example

WRITE

progn3

WRITE

READ

2 0 4 0 0 0 0 0 0 0 0

Read from location 2, i.e. 4

2 0

4 0

2 0
What if this is mutated?
Into a 3... or a 12?

What happens here?
Using addressable memory is problematic.

An alternative is to use data structures

- These are less expressive than ‘full memory’
- But less susceptible to bad mutations
- e.g. using a stack:
Other Memory Approaches

◊ “Soft memory” [Poli et al 2009]
  ▶ Blends new assignments with old memories
  ▶ Intended to be more evolvable
  ▶ [http://cswww.essex.ac.uk/staff/poli/papers/PoliMcPheeCitiCraneJAEA2009.pdf](http://cswww.essex.ac.uk/staff/poli/papers/PoliMcPheeCitiCraneJAEA2009.pdf)

◊ Cultural memories [Spector & Luke 1996]
  ▶ Memory is shared amongst the population
  ▶ Allows programs to communicate
This is possible, but…

- Requires caution!
- Halting Problem: can’t predict termination
- Need to use constraints to prevent infinite loops
- (e.g. max iterations)
- Or use a time-out during solution evaluation

Alan designed the perfect computer

http://www.coopertoons.com
Adding Loops

- **Koza-style:**
  
  ```
  do {
    move();
    until(IRsensor>5)
  }
  ```

- **Langdon [1996]:**
  
  ```
  for(i=2+3; i<9; i++)
  {
    write(1,i);
    loop index
  }
  ```
Iteration has been found useful for image analysis

- Often processed in segments, e.g. every 10x10 pixels
- Evolved loops can be more efficient and effective
  
It is possible to add syntactic features to GP:
- Types ✔
- Memory ✔
- Loops ✔

But this has consequences:
- More complex initialisation and variation operators
- More constraints during evolution
- Possible biases within the search landscape

So, only use them when necessary
Sub-expressions frequently re-occur in a program

- Ideally, we only want to evolve each one once
- So, not this:
Sub-expressions frequently re-occur in a program

- Ideally, we only want to evolve each one *once*
- This:

```
  define sub:
   \( \times \)  \\
   \( \div \)
   \( \times \)  \(+\)
   \( t \)  \( y \)  \( 3 \)  \( y \)
```

```
( sub \times sub ) \div ( ( t \times y ) + 3 \times y )
```
A program is defined as a forest of trees

- One “result producing branch” (RPB), i.e. main()
- One of more ADFs, each with one or more arguments

```
RPB
  ×
/  \
ADF0  ADF1
3   y   r

ADF0(arg0, arg1)
  ÷
/  \
ADF1(arg0)
  ×
/  \
arg0  y  arg1

ADF1(arg0)
  ×
/  \
arg0  sin
```
ADFs must be defined before a GP run
  ▶ Both the number of ADFs, and the number of arguments for each ADF must be specified in advance
  ▶ This is a prominent limitation of ADFs

Some heuristics for choosing these values:
  ▶ A priori knowledge of problem decomposition
  ▶ Over-specification, i.e. more than will ever be needed
  ▶ Affordable capacity, since ADFs tend to increase the complexity and hence the execution time of programs
Good things about ADFs 😊

▸ They can reduce overall program size
▸ They make it easy to solve modular problems
▸ They have been used to solve hard problems
▸ See Koza’s books!

Bad things about ADFs 😞

▸ They can increase program complexity
▸ Incorrect parameter settings hinder evolution
▸ Modular dependencies may hinder evolution…
This is also seen in software engineering

- Two routines make use of shared code
- The routines’ requirements diverge
- The shared code must be copied and modified
- i.e. it is no longer shared code!
Koza introduced a similar mechanism to GP

- Called ‘case splitting’
- Part of a group of ‘architecture altering operators’
- Basically mutation operators that refactor code
Another group of methods for achieving modularity
- These attempt to identify useful code blocks
- And then turn them into modules
- For potential use elsewhere in evolving programs

Adaptive Representation through Learning (ARL)
- A successful form of analytical modularity
- First, it identifies programs with improved fitness
- Then extracts the most-executed code blocks
- Placing these in a library shared between programs
It is possible to add syntactic features to GP:

- Types ✔
- Memory ✔
- Loops ✔
- Modularity ✔

But this has consequences:

- More complex initialisation and variation operators
- More constraints during evolution
- Possible biases within the search landscape

So, only use them when necessary
Some limitations of GP are due to using trees:
There are advantages to using graphs instead
  ▶ Instant reuse!
There are a number of graph-based GPs
- PADO, PDGP, GNP, CGP, ...

Cartesian GP (CGP) is the best known [Miller]
- Functions are arranged on a Cartesian grid
Other notable properties of CGP

- Constrained grid limits program size (no bloat!)
- Mutation can connect/disconnect nodes
- Disconnected nodes are a form of redundancy
- Redundancy has evolutionary advantages
Other notable properties of CGP

- Constrained grid limits program size (no bloat!)
- Mutation can connect/disconnect nodes
- Disconnected nodes are a form of redundancy
- Redundancy has evolutionary advantages

This sub-graph is still present but is no longer in use

Output
Other notable properties of CGP

- Constrained grid limits program size (no bloat!)
- Mutation can connect/disconnect nodes
- Disconnected nodes are a form of redundancy
- Redundancy has evolutionary advantages

This sub-graph was unexpressed, but is now used
This allows old code to come back into use, and supports ‘neutral evolution’, a good thing!
Things you should know

✧ The meaning of evolvability and expressiveness
  ▶ And that there is often a trade-off between them
  ▶ And be able to give examples of this trade-off

✧ ADFs: what are they, why use them, rules of thumb
✧ Cartesian GP: what is it, why use it

✧ I don’t expect you to know:
  ▶ Details of GP memory and loop implementations
  ▶ About modular dependencies and analytical modularity
This week

- Some ways of making tree GP more expressive
- Some thoughts on how this affects evolvability
- The potential for using different representations

Next week

- Other ways of representing programs
- Turing complete genetic programming
- Conventional languages
- What we can learn from biology
Things to try out

◊ ADFs
  ▶ ECJ has support for these
  ▶ Try them out on a few problems
  ▶ Try using different numbers of ADFs

◊ CGP
  ▶ Add-on available for ECJ
  ▶ Have a look, try it out
