Genetic Programming

CS 5764
Evolutionary Computation
Hod Lipson
Symbolic Regression

What function describes this data?

\[ f(x) = e^x \sin(|x|) \]

John Koza, 1992
Encoding Equations

Building Blocks: + - * / sin cos exp log … etc

\[ f(x) = (x_1 - 3) \cdot \sin(x_2) + x_1 \cdot \sin(x_2) - 7 + x_2 \]

John Koza, 1992
GENETIC PROGRAMMING
ON THE PROGRAMMING OF COMPUTERS BY MEANS OF NATURAL SELECTION
GA with Tree Representations

\[ X + 1 \]
\[ X^2 + 1 \]
\[ 2 \]
\[ X \]

\[ 0.67 \]
\[ 1.00 \]
\[ 1.70 \]
\[ 2.67 \]
Basic Building Blocks

• Trees replace strings
• BBs replace alleles
  – For algebraic problems (+, -, /, *)
  – Trig functions Sin(), Cos()…
  – Exponents, etc.
Variation

• Mutation (small, random)
  – Change coefficient
  – Replace branch with constant

• Crossover (Large, non-random)
  – Swap sub-trees
    • (may or may not align)
Differences from GA

- Open ended representation
- Variable Linkage
- Hierarchical
Challenge

• How to store an open ended equation?

\[ x^2 + 1 + \sin\left(\frac{1}{x}\right) \]
\[ 2x^2 + \frac{1}{x} \]
\[ x^2 + 1 + \text{six} \]
\[ 2x^2 + \frac{1}{n\left(\frac{1}{x}\right)} \]
Challenge

• How to store an open ended equation?

\[ X^2 + 1 \]

“x*x+1”
Heap Data Structure

- Root element at position 1 (not 0)
- Children of element $i$ in position $2i$ and $2i+1$

<table>
<thead>
<tr>
<th>Index</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>*</td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>x</td>
</tr>
<tr>
<td>7</td>
<td>x</td>
</tr>
</tbody>
</table>
Heap Data Structure

- Root element at position 1 (not 0)
- Children of element \(i\) in position \(2i\) and \(2i+1\)

<table>
<thead>
<tr>
<th>Index</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

What is the array size for tree of depth \(h\)?
Graph Representations

\[ f(\theta, \omega) = 4.771 \cdot (3.714 - \omega^2) + \cos(\theta) \]

\[ + (3.714 - \omega^2) \cdot \cos(\theta) \]

\[ < \text{load}[3.714] \]

\[ < \text{load}[\omega] \]

\[ < \text{mul}(1), (1) \]

\[ < \text{mul}(3), (6) \]

\[ < \text{add}(12), (6) \]

\[ < \text{add}(13), (7) \]

\[ A \]

\[ B \]

\[ + \]

\[ x + \]

\[ – \]

\[ 4.771 \]

\[ \omega \]

\[ \omega \]

\[ 3.714 \]

\[ x \]

\[ x \]

\[ \cos \]

\[ \theta \]
Graph Representations

\[ f(\theta, \omega) = 4.771 \cdot (3.714 - \omega^2) + \cos(\theta) 
+ (3.714 - \omega^2) \cdot \cos(\theta) \]

\begin{array}{|c|}
\hline
(0) <- load [3.714] \\
(1) <- load [\omega] \\
(2) <- mul (1), (1) \\
(3) <- sub (0), (2) \\
(4) <- load [\theta] \\
(6) <- \cos (4) \\
(7) <- mul (3), (6) \\
(9) <- load [4.771] \\
(12) <- mul (9), (3) \\
(13) <- add (12), (6) \\
(15) <- add (13), (7) \\
\hline
\end{array}

B

\[ + \]
\[ \times \]
\[ \times \]
\[ \cos \]
\[ \theta \]
\[ 3.714 \]
\[ 4.771 \]
\[ \omega \]
\[ \omega \]
\[ + \]
\[ - \]
\[ \times \]
Additional Variations

• **Snipping**
  – Replace branch with a constant at the average of the branch’s output

• **Pruning**
  – Eliminate branches with relatively low contribution
General Applications

- Think of GP for
  - Circuits
  - TSP
  - Kinematic systems
  - Geometric shapes
Example: Analog Circuits

U. S. patent 1,227,113—George Campbell—AT&T—1917

Post 2000 US patent
LOW-VOLTAGE balun circuit

U. S. patent 1,538,964—Otto Zobel—AT&T—1925

One criterion for innovation:
Patentability
Development

[Diagram showing a combination of series and parallel circuits with components labeled as L, C, and R]
Resilient Circuits
Resilient circuits
How do we simplify?

• **Snipping**
  – Replace a subsystem with a single component

• **Pruning**
  – Eliminate subsystems with relatively low contribution
More building Bloks

\[ f(x) = \sin(x + \text{noise}) \times \sin(x) - 1.2 \times x \times \text{noise} \]
Automatically Defined Functions

b. $\text{ADF}_0$
   - $\div$
     - $\div$
       - $-$
         - $-$
           - $a$
           - $b$
         - $b$
       - $b$
   - $b$

b. $\text{ADF}_1$
   - $\times$
     - $\div$
       - $a$
       - $-$
         - $-$
           - $a$
           - $b$
         - $b$
       - $b$
   - $b$

b. $\text{ADF}_2$
   - $-$
     - $\times$
       - $a$
       - $+$
         - $+$
           - $+$
             - $+$
               - $a$
               - $b$
               - $b$
               - $b$
   - $b$

c.

Cell

- $\times$
  - $\times$
    - $+$
      - $+$
        - $+$
          - $+$
            - $+$
              - $a$
              - $b$
              - $b$
              - $b$
            - $ADF_2$
          - $ADF_0$
        - $ADF_1$
      - $ADF_0$
   - $ADF_0$
Automatic Programming
General Applications

• Think of GP for
  – Circuits
  – Geometric shapes
  – Kinematic systems
  – TSP
The Straight Line Problem

It is easy to think of a mechanism that traces an exact circle without having a circle built in: A compass.

Can you think of a linkage mechanism that will trace a straight line without reference to an existing straight line?
The Straight Line Problem

It is easy to think of a mechanism that traces an exact circle without having a circle built in: A compass.

One solution: The Peaucellier (1873)

The straightness of the links themselves does not matter
The Straight-Line problem

• Needed to guide the piston of the steam engine.
  – *The* breakthrough that made steam engines a success

“Though I am not over anxious after fame, yet I am more proud of the parallel motion than of any other mechanical invention I have ever made”

James Watt, cf. 1810 [15]
More established solutions

Silverster-Kempe’s (1877)

Peaucelier (1873)

Chebyshev (1867)

Robert (1841)

Chebyshev (1867)

Chebyshev-Evans (1907)

Source:
Kempe A. B., (1877), How To Draw A Straight Line, London

See http://kmoddl.library.cornell.edu
Considered fundamental technology

Cornell University acquired in 1882 about 40 straight-line mechanism models and used them in the early engineering curriculum.

See videos at Cornell University Digital Library of Kinematic Models

http://kmoddl.library.cornell.edu
Evolving Straight line mechanisms
Design a straight-line tracing mechanism

Silverster-Kempe’s (1877)
Peaucelier (1873)
Chebyshev (1867)
Robert (1841)
Evolving Photonic Crystals

With Stefan Preble & Michal Lipson, 2004 (PRL)