Evolution, the First Programmer

Kyle Harrington
DEMO Lab, Computer Science
Brandeis University
Primate Programmers

• The Infinite Monkey Theorem:
  – In the limit of infinite time, monkeys can write Shakespeare.

6 monkeys, a computer, and 1 month later 5 pages of mostly the letters S, A, J, L and M.
(From U Plymouth, reported in Wired)
Micrographs to 3D structures

Rendering of experimentally-derived ribosome structure

3D reconstruction

Images courtesy of Axel Brilot
Wikimedia
Recap: Axel’s Fancy Tools

(see references in Axel’s slides)
BEADS Evolutionary Automated Discovery of Structures
Example: Structure Search
Local Optima Problem
Evolutionary Algorithm

- Evolution acts in parallel
- Proteins evolved in parallel

Graph showing Goodness of Fit vs. # of models evaluated.
Representation

• How does one represent a candidate model?

• Computational evolution uses bit strings

• What about representing this?

Images courtesy of UCSC, RNA Center
Our Representation

• Volumes of density as spheres
  – Variable number of spheres
  – Variable radius
Evolving Structural Models

1. Initialize Population
   - Initialize population
     * Generate initial models (random or biased)

2. Evaluate fitness of population
   - Evaluating a candidate model
     * Computing fitness of a model
       - Discover original particle orientation
       - Discover particle structure

3. Reproduce selections with variation

4. Select more fit individuals

5. Continue Evolving?
   - Yes
   - No

Example: Populations of Structures
Initialize Population

- Generate initial models (random or biased)
Evaluating a Candidate Model

• Computing fitness of a model
  – Simulate microscope imaging
  – Ray-trace from imaging plane
Particle Orientation

- Coevolve orientation and structural model

Common lines approach (see Axel’s slides)
Selection

- A: 23%
- B: 21%
- C: 19%
- D: 19%
- E: 18%
Mutation

• Mutation randomly perturbs candidate model
Populations of Models

- Individual models may favor different features
Fitness Landscapes

- Fitness landscape describes the distribution of phenotypes by fitness value
- Local optima
- Multiple peaks
- Deceptive optima

Images courtesy of M. Frame
Example: Evolutionary Branching
Recombination Optimizes Mixability

• Selection acts on the mixes of individuals (children)
Concluding Particle Reconstruction

- Evolutionary particle reconstruction is a natural next step for the field

- Future directions:
  - Unifying existing algorithms with evolutionary algorithms
  - Recombination and environmental variation
Evolving Mathematical Models

• Goal: discover best model from experimental data with minimal human effort.

Evolving Mathematical Models

Target expression: \( g(x) = x^2 + 2x + 1 \)

Example individual:

\[
\begin{align*}
&+ \\
&1 \quad * \\
&x \\
&+ \\
&x \quad 1
\end{align*}
\]

Mathematical representation:

\( f(x) = 1 + x \cdot (x + 1) \)

- \( X = 0, f(X) = 1, g(X) = 1 \)
- \( X = 1, f(X) = 3, g(X) = 4 \)
- \( X = 2, f(X) = 7, g(X) = 9 \)

Individual’s error: 3
Mutating Expressions
Crossing Over Expressions

Parent 1

1 +
   / 
  * 1
 /   
X  +
   / 
  X  1

Parent 2

+  *
/  /
X 1
 /
X  X

Child 1

+  *
/  /
1 1
 /   
X  X

Child 2

+  *
/  /
+ 1
 /  /
X  X
   / 
  X  1
Predicting Epitope Binding

• Evolving mathematical models of molecular docking
Evolving Other Things

- Swarm ecosystems
- Animats
- Robots
Brevis

- [http://brevis.golemics.org](http://brevis.golemics.org)
- Open-source, interfaces with existing science software, and stable release coming-soon
Thank you

- Jordan Pollack
- James Chin
- Jessica Lowell
- Jeff Gelles and Jane Kondev
- QB program
- Audience