Evolution and Computation

Christos H. Papadimitriou
The Simons Institute
The Algorithm as a Lens

• It started with Alan Turing, 60 years ago
• Algorithmic thinking as a novel and productive point of view for understanding and transforming the Sciences
• Major theme of the Simons Institute
• …and of this symposium
• This talk: Evolution
Evolution before Darwin

• Erasmus Darwin
Before Darwin

• J.-B. Lamarck
Before Darwin

• Charles Babbage

[ca. 1820, paraphrased]

“God created not species, but the Algorithm for creating species”
The Origin of Species

- Natural Selection
- Common Ancestry
- Possibly the world’s most masterfully compelling scientific argument
- The six editions: 1859, 1860, 1861, 1866, 1869, 1872
The Wallace-Darwin papers: Exponential Growth
Brilliant argument, and yet many questions left unasked, e.g.:

• How does novelty arise?

• What is the role of sex?
Cryptography against Lamarck

• A. Weismann

[ca. 1880, paraphrased]

“The mapping from genotype to phenotype is one-way”
Genetics

- Gregor Mendel [1866]
- Number of citations between 1866 and 1901: 3
The crisis in Evolution
1900 - 1920

• Mendelians vs. Darwinians
• Geneticists vs. Biometricists/Gradualists
Big questions remain e.g.:

- How does novelty arise?
- What is the role of sex?
Disbelief

“Our thesis is that Neo-Darwinism cannot explain the basic phaenomena of evolution on the basis of physico-chemistry”

Schützenberger, 1966
Disbelief at the top

“The eye to this day gives me a cold shudder.”
Disbelief, algorithmic version

“What algorithm could have achieved all this in a mere $10^{12}$ steps?”

(surprise: we have an answer…)
Valiant’s Evolvability

“How do you find a 3-billion long string in 3 billion years?”

L. G. Valiant

Computationally-inspired model of Evolution shown to amount to a weak form of learning
Evolution and CS Practice: Genetic Algorithms [ca. 1980s]

- To solve an optimization problem…
- …create a population of solutions/genotypes
- …who evolve through mutations and sex…
- …and procreate with success proportional to their objective function value
- Eventually, some very good solutions are bound to arise in the soup
And in this Corner…
Simulated Annealing

- Inspired by *asexual* reproduction
- Mutations are adopted with probability increasing with fitness/objective differential
The Mystery of Sex Deepens

• Simulated annealing (asexual reproduction) works fine
• Genetic algorithms (sexual reproduction) don’t work
• In Nature, the opposite happens: Sex is successful and ubiquitous
A Radical Thought

• What if sex is a mediocre optimizer of fitness (= expectation of offspring)?
• What if sex optimizes something else?
• And what if this something else is its *raison d’être*?
• [Livnat et al, PNAS 2008]
• Simulations show that natural selection under *asex* optimizes fitness
• But under *sex* it optimizes *mixability*:
• = The ability of alleles (gene variants) to perform well with a broad spectrum of other alleles
Explaining Mixability: The Fisher-Wright model

- Fitness landscape of a 2-gene organism

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Rows: alleles of gene A

Columns: alleles of gene B

Entries: fitness of the combination
Explaining Mixability (cont)

- Asex will select the largest numbers

<table>
<thead>
<tr>
<th>3</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>
Explaining Mixability (cont)

- But sex will select the rows and columns with the largest average

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>
Pointer Dogs
Pointer Dogs

C. H. Waddington
Waddington’s Experiment (1952)

**Generation 1**

**Temp:** $20^\circ C$
Waddington’s Experiment (1952)

Generation 2-4
Temp: $40^\circ$ C
$\sim$15% changed
Select and breed those
Waddington’s Experiment (1952)

Generation 5
Temp: 40°C
~60% changed
Select and breed those
Waddington’s Experiment (1952)

Generation 6
Temp: $40^\circ C$
~63% changed
Select and breed those
Waddington’s Experiment (1952)

(...)

Generation 20

Temp: 40° C

~99% changed
Surprise!

Generation 20
Temp:  20° C
~25% stay changed!!
Genetic Assimilation

- Adaptations to the environment become genetic!
Is There a Genetic Explanation?

Function $f(x, h)$ with these properties:

- Initially, $\text{Prob}_{x \sim p[0]}[f(x, h = 0)] \approx 0\%$
- Then $\text{Prob}_{p[0]}[f(x, 1)] \approx 15\%$
- After breeding $\text{Prob}_{p[1]}[f(x, 1)] \approx 60\%$
- Successive breedings, $\text{Prob}_{p[20]}[f(x, 1)] \approx 99\%$
- Finally, $\text{Prob}_{p[20]}[f(x, 0)] \approx 25\%$
A Genetic Explanation

• Suppose that “red head” is this Boolean function of 10 genes and “high temperature”
  “red head” = “x_1 + x_2 + ... + x_{10} + 3h \geq 10”

• Suppose also that the genes are independent random variables, with p_i initially half, say

• All properties of the Waddington experiment satisfied

• [Stern AN 1958]
Arbitrary Boolean Functions

• What if we have an arbitrary function of genes (no environmental variable $h$)?
• Suppose the satisfying genotypes have a fitness advantage ($1 + \varepsilon$ vs. 1, say)
• Will this trait be fixed eventually?
Arbitrary Functions: Yes!

Theorem: Any Boolean function of genes which confers an evolutionary advantage will be eventually fixed (with high probability)

(2013; with Adi Livnat, Aviad Rubinstein, Greg Valiant, Andrew Won)
Which means that…

• “With sex, all moderate-sized Boolean functions are evolvable.”
• “Look, Ma, no mutations!”
• Novel complex traits can emerge, through sex, in the whole population, without “Fisherian propagation”
Neutral Theory and Weak Selection

• Kimura 1970: Evolution proceeds not by leaps upwards, but mostly “horizontally,” through statistical drift

• Weak selection: the values in the fitness matrix are very close, say in $[1 - \varepsilon, 1 + \varepsilon]$
Changing the subject: The experts problem

- Every day you must choose one of n experts
- The advice of expert $i$ on day $t$ results in a gain $G[i, t]$ in $[-1, 1]$
- Challenge: Do as well as the best expert in retrospect
- Surprise: It can be done!
- [Hannan 1958, Cover 1980, Winnow, Boosting, no-regret learning, MWUA, …]
Multiplicative weights update

- Initially, assign all experts same weight/probability
- At each step, increase the weight of each by $(1 + \varepsilon G[i, t])$ (and then normalize)
- **Theorem**: Does as well as the best expert
- MWUA solves: zero-sum games, linear programming, convex programming, network congestion,…
Disbelief

Computer scientists find it hard to believe that such a crude technique solves all these sophisticated problems

(cf: the other disbelievers)
Theorem: Under weak selection, evolution of a species is a game

- the players are the genes
- the strategies are the alleles
- the common utility is the fitness of the organism (coordination game)
- the probabilities are the allele frequencies
- game is played through multiplicative updates

(2013, with E. Chastain, A. Livnat, U. Vazirani)
Finally...

- **Variance preservation**: MWUA is known to maximize entropy
- The curious successes of Evolution and MWUA: Two mysteries united
- *This* is the role of sex in Evolution
Sooooo...

- The theory of life is deep and fascinating
- And rife with interesting technical problems that are unexpectedly computational
- How robust? (Strong selection? Complex landscapes?)
- Environment? Mutations?
- Test?
Thanks!