Swarm Algorithms

- Solve problems using cooperative behavior of a group of agents
- Santa Fe - “We use the term ‘swarm’ in a general sense to refer to any loosely structured collection of interacting agents.”
- Each agent is only part of the overall solution
  - Flocking behavior - Reynold’s (1987) Boids
    - Avoid colliding with neighbors
    - try to go about the same speed as neighbors
    - try to move toward the center of the flock
  - Ant algorithms - food finding, path construction, nest building, routing of information, schedule planning (Bonabeau. Dorigo, & Theraulaz’s 1999 book “Swarm Intelligence”)
- Cooperative interactions where each agent offers a complete solution to a problem
  - Axelrod’s (1997, “The complexity of cooperation) Culture Model
  - Particle Swarm Optimization (Kennedy & Eberhart’s, 2001 book “Swarm Intelligence”)
Culture Model (Axelrod, 1997)

- Model of the spread of culture with social interactions
- Individuals represented by symbolic strings (as in GAs), and arranged on a 2-D grid
- Probability of agent interaction increases with agents’ similarity

Algorithm
- Randomly select an agent
- Randomly select a neighbor of the agent
- Probability of interaction = (# shared string values)/(string length)
- If interact, then copy value on one feature of neighbor’s string to the agent’s string

Behavior
- Preservation of different cultures in same world (compare to Nowak)
- Increased local harmony over time
- Similarity-based interaction creates distinct boundaries
### Figure 6.1
Initial random start for a simulation of a $10 \times 10$ population of individuals made up of strings of five features represented by numerals ranging from zero to nine.

<table>
<thead>
<tr>
<th>27217</th>
<th>74924</th>
<th>31157</th>
<th>53671</th>
<th>22660</th>
<th>37316</th>
<th>07959</th>
<th>57666</th>
<th>33206</th>
<th>92725</th>
</tr>
</thead>
<tbody>
<tr>
<td>66219</td>
<td>08226</td>
<td>26707</td>
<td>45600</td>
<td>48767</td>
<td>39481</td>
<td>62784</td>
<td>89859</td>
<td>27792</td>
<td>35492</td>
</tr>
<tr>
<td>37262</td>
<td>66163</td>
<td>89178</td>
<td>60968</td>
<td>91098</td>
<td>19937</td>
<td>62103</td>
<td>07562</td>
<td>03500</td>
<td>13864</td>
</tr>
<tr>
<td>87746</td>
<td>66209</td>
<td>94122</td>
<td>72784</td>
<td>03593</td>
<td>16647</td>
<td>19776</td>
<td>87819</td>
<td>22160</td>
<td>48185</td>
</tr>
<tr>
<td>16880</td>
<td>09713</td>
<td>76057</td>
<td>30843</td>
<td>92125</td>
<td>41152</td>
<td>74156</td>
<td>98801</td>
<td>64760</td>
<td>00144</td>
</tr>
<tr>
<td>86287</td>
<td>66161</td>
<td>23271</td>
<td>46773</td>
<td>53014</td>
<td>44442</td>
<td>25424</td>
<td>98309</td>
<td>32553</td>
<td>16678</td>
</tr>
<tr>
<td>90624</td>
<td>65685</td>
<td>68785</td>
<td>32385</td>
<td>90770</td>
<td>24676</td>
<td>68806</td>
<td>25347</td>
<td>16640</td>
<td>30602</td>
</tr>
<tr>
<td>98681</td>
<td>11402</td>
<td>67304</td>
<td>68003</td>
<td>16943</td>
<td>01041</td>
<td>44693</td>
<td>63237</td>
<td>76040</td>
<td>61075</td>
</tr>
<tr>
<td>52249</td>
<td>30617</td>
<td>91425</td>
<td>92780</td>
<td>82342</td>
<td>30467</td>
<td>19721</td>
<td>84117</td>
<td>96595</td>
<td>55215</td>
</tr>
<tr>
<td>79949</td>
<td>70851</td>
<td>29089</td>
<td>89311</td>
<td>19176</td>
<td>67653</td>
<td>95954</td>
<td>64805</td>
<td>51332</td>
<td>74301</td>
</tr>
</tbody>
</table>

### Figure 6.2
Result of simulation where interaction is a stochastic function of similarity (Axelrod’s paradigm).
Phase Transitions in the Culture Model

$q < q_c$: Monocultural
Global culture

$q > q_c$: Multicultural
Cultural diversity
Global polarization

Transition well defined as $N \to \infty$

$q = $ Number of traits (values) per feature, $F = $ Number of features
$N = $ Row and Column size of population
$S_{\text{max}} = $ size of the largest homogenous domain
Adaptive Culture Model (Kennedy & Eberhart, 1997)

- Agents interact with each other based not on similarity, but on performance
  - Behavior is not only based on fads, but on results
  - “If neighbor’s sum is larger than my sum, then interact” -> “99999”
  - “If neighbor’s first 3 numbers are close to last 2 numbers (compared to my numbers), then interact” -> pockets of “17769” and “83193”

- Can solve the Traveling salesman problem

- Behavior
  - Preservation of different cultures in same world (compare to Nowak)
  - Increased local harmony over time
Solving $X_1 + X_2 + X_3 = X_4 + X_5$

![Table of numbers]

**Figure 6.4** Result of a simulation where interaction occurred when the difference between the sums of the first three numbers and the last two was smaller for the neighbor than for the individual. Note that in all cases the sum of the first three numbers equals the sum of the last two.
ACM can solve the TSP

All solutions are variants of ABCDEFGH
Particle Swarm (Kennedy & Eberhart, 1997)

- Flocking behavior to solve search problems
- Each particle in a system tends to move toward its own historically best solution and the best of its neighbors
- Individual particles have
  - a location in a ring (that does not change - flying in “description space”)
  - a location, \( X_{i,d,t} \) for each particle \( i \) along each dimension \( d \) at time \( t \)
  - a velocity, \( V_{i,d,t} \) that gives the particle’s inertia
- Execution cycle
  - Initialize population with random positions and velocities
  - For each particle, determine its value on the fitness function
    - Compare particle’s current fitness to its historically best fitness (Pbest). If it is greater, set Pbest to the current particle’s fitness.
    - Compare the particle’s fitness to the neighbors’ historically best fitness (Gbest). If it is greater, set Gbest to the current particle’s fitness
  - \( V_{i,d,t+1} = V_{i,d,t} + C_1 \times \text{rand} \times (P\text{best}_{i,d} - X_{i,d,t}) + C_2 \times \text{rand} \times (G\text{best}_d - X_{i,d,t}) \)
  - \( X_{i,d,t+1} = X_{i,d,t} + V_{i,d,t+1} \)
Other Considerations

- Can replace Gbest with Local best within neighborhood of particle
  - Good solutions if Lbest is about 15% of population
- $V_{\text{max}} =$ maximal speed
  - If $V_{\text{max}}$ is too high, particles may pass good solutions
  - If $V_{\text{max}}$ is too low, particles may not explore far enough, and may become trapped in local minima
- Inertial weight
  - $V_{i,d,t+1} = W V_{i,d,t} + C_1 \cdot \text{rand} \cdot (P\text{best}_{i,d} - X_{i,d,t}) + C_2 \cdot \text{rand} \cdot (G\text{best}_d - X_{i,d,t})$
  - $W$ often decreases during a run
  - Shift from exploration to exploitation with time
- For discrete, not continuous, decisions:
  - if $C < S(V_{i,d,t+1})$ then $X_{i,d,t+1} = 1$, else $X_{i,d,t+1} = 0$
  - $C$ is a number chosen from a uniform distribution $\{0,1\}$
  - $S(x)$ is a sigmoidal squashing function. $S(X) = 1/(1+\exp(-X))$
  - Continuous velocity is converted to discrete location
Finding a good value for $V_{\text{max}}$

$$v = v + C(\text{best-}x) \quad [C \text{ is a random variable}]$$

$$x = x + v$$

no $V_{\text{max}}$

One particle with one dimension

$V_{\text{max}} = 2$

converges close to optimal $x$, but local minima are possible
Finding good weights for the control parameter C

Low C: Not quickly pulled toward best solution

If there is no decrease in velocity weight, particles do not converge on single solution
Applications of Particle Swarms

• Discrete version
  – De Jong’s suite of functions
    • Sum(x^2) is simple
    • multi-peak functions are harder but often solveable
      – often better than GAs because different classes of solution are not bred together
      – advantage of using a Local best rather than Global best in particle swarm
  – Two pulls on behavior: Individuality and conformity
    • Boyd and Richerson’s (1985) Cultural transmission model: Learning by direct experience + learning by culture
    • Ajzen and Fishbein (1980) Reasoned action model: intentions are determined by personal attitudes and subjective norms

• Continuous version
  – Evolve neural network weights (rather than using backprop)- solves the XOR problem
  – Estimate the charge left in batteries in a vehicle with a neural network