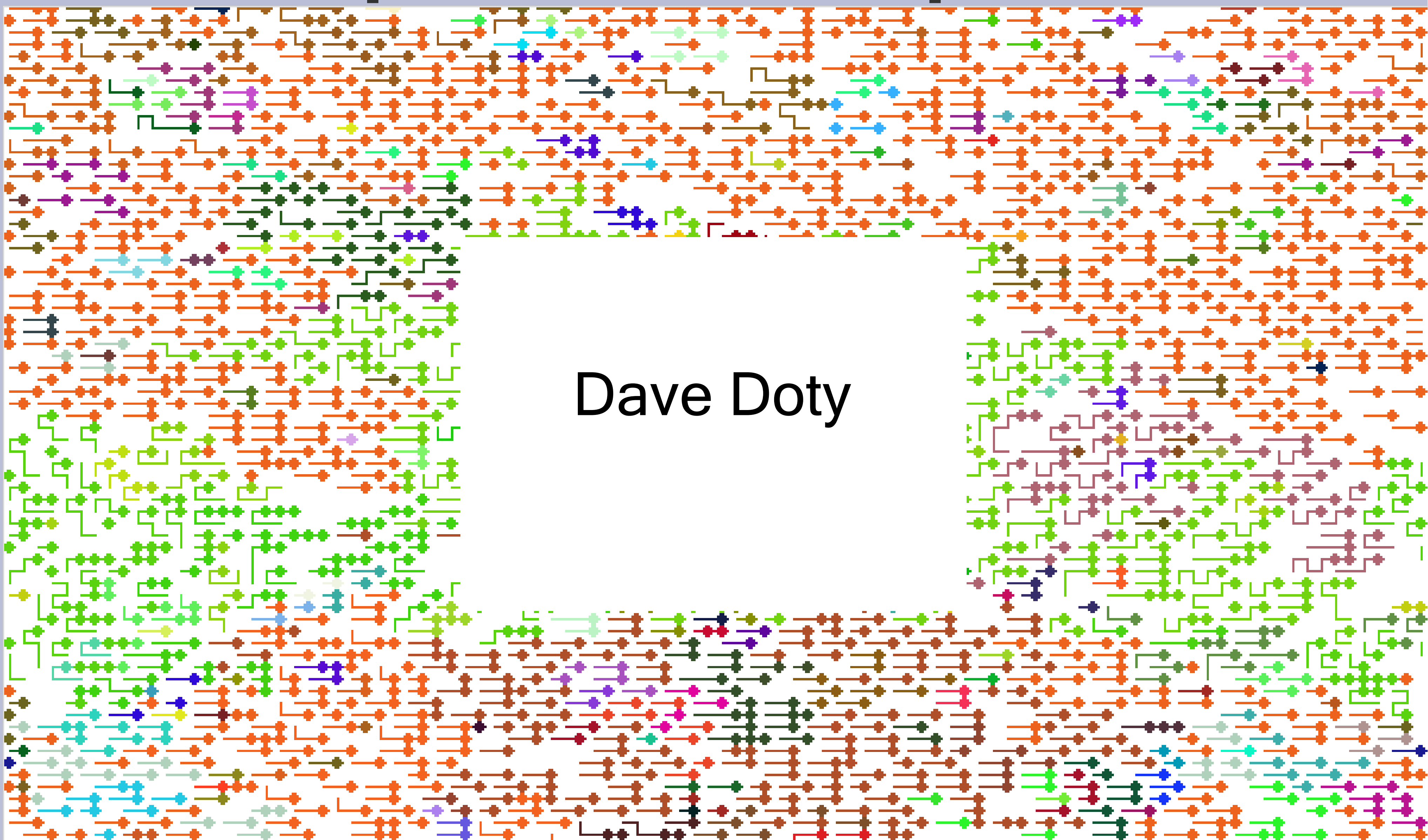
Coevolution and Non-local Adaptation in Gridplants



Dave Doty

Coevolution in biology

- Coevolution is
  - “a type of evolution in which two species become increasingly adapted to each other, resulting in a highly specific interaction.”
  - “changes in the genotypes of two or more species that are a direct consequence of the species' interaction with one another.”
  - “the process whereby genes or gene fragments are changing together and not diverging.”
  - “the process whereby two or more organisms interact and affect each others' evolution.”

Coevolution in CAS

- Definition of evolution by natural selection
  - An evolutionary system is a collection of reproducing agents wherein agents differ in their reproductive ability (termed *fitness*), and in which traits affecting fitness are heritable.
- Generalized definition of coevolution
  - A coevolving system is an evolutionary system in which the fitness of any agent is dependent on interactions with other agents.

Adaptation in coevolutionary systems: two views

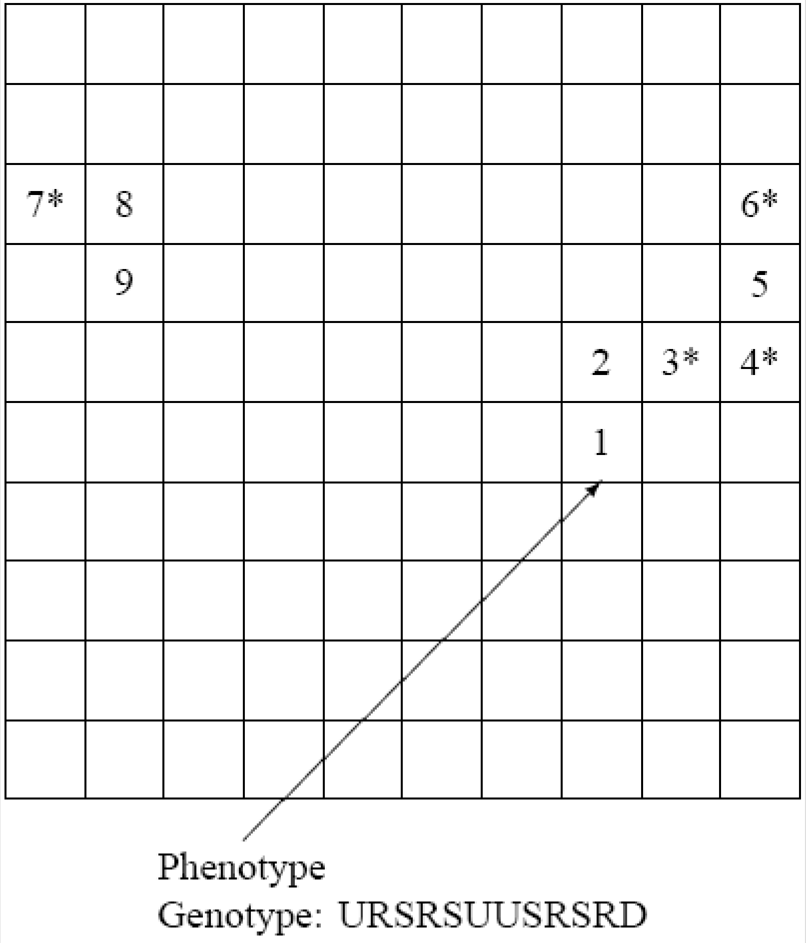
- Adapt only to local environment?
  - “It is widely believed that habitat specialization will arise because a jack of all trades is a master of none.”

    - Andrew Cockburn
- Develop general adaptive abilities?
  - “... if, by the medium of a time machine, predators from one era could meet prey from another era, the later, more 'modern' animals, whether predators or prey, would run rings around the earlier ones. This is not an experiment that can ever be done...”

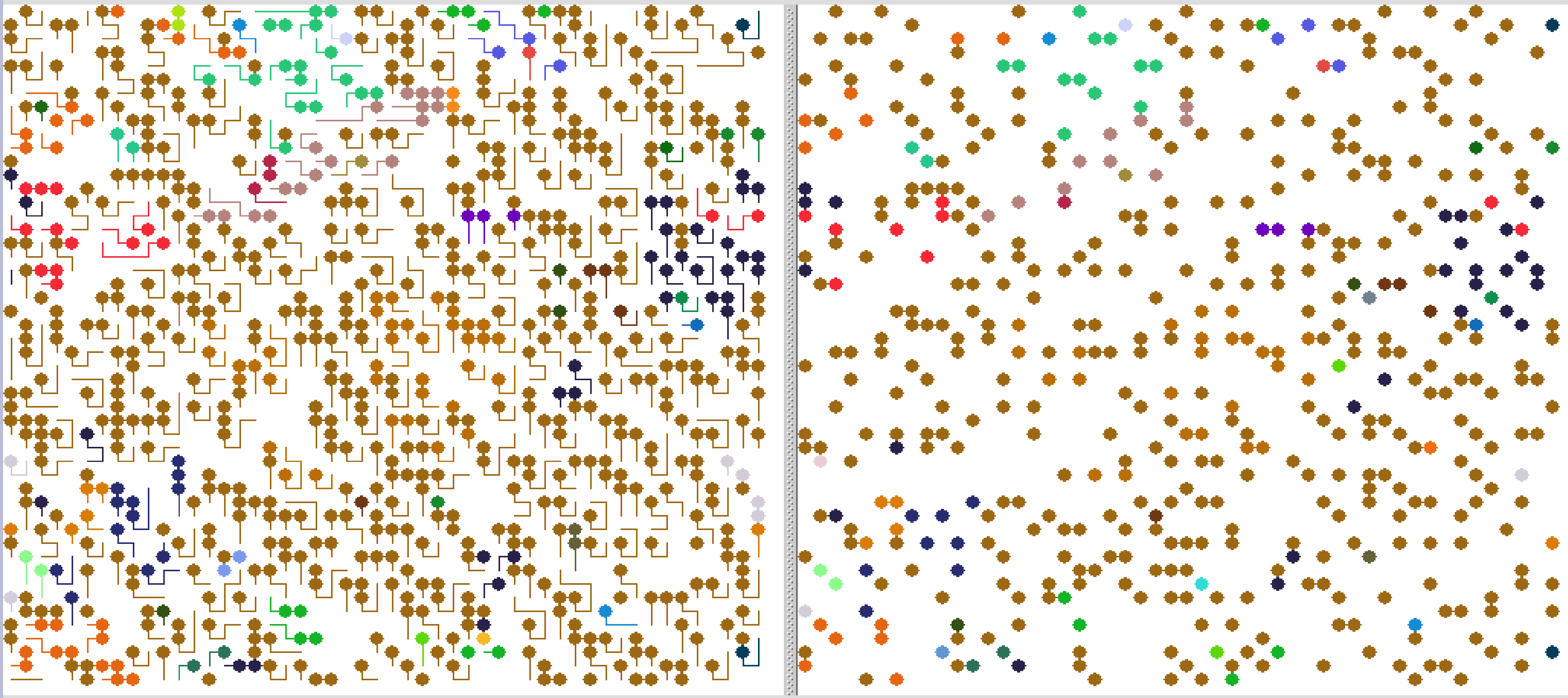
    - Richard Dawkins

Experiment: Gridplants

- Gridplants are simulated plants that live on a 2-dimensional toroidal grid
- Each plant starts as a single seed, and it grows and reproduces by following instructions in its genome (Up, Down, Left, Right, Seed).



Cell layer and seed layer



Growth simulation

- Actions cost energy; occupying space gains energy
- Impossible actions cost no energy
- If not enough energy to execute an action, plant must wait until it has enough
- Mutation and seed mortality happen at random

```
EVOLVE-GRIDPLANTS
seed ← random seed genome
plant seed at random position on grid
for generation ← 1, 2, ...
    do SPROUT
       for i ← 1 to season length
           do GROW
```

```
SPROUT
for each seed seed
    do if seed survives
          then plant ← seed
               stochastically mutate genome of plant
               energy[plant] ← starting energy
```

```
GROW
▷ plants is list of all plants
randomly permute order of plants
for each plant plant ∈ plants
    do energy[plant] ← energy[plant] + INCREMENTAL-ENERGY(plant)
       cost ← cost of action
       if energy[plant] ≥ cost
          then if EXECUTE-ACTION(plant)
                  then energy[plant] ← energy[plant] − cost
```

Growth Parameters

- Costs:
  - Read genome: 1
  - Grow: 3
  - Plant seed: 10
- Initial energy: 4
- Gain 1 energy per cell that plant occupies - 1
- Mutation rate: 1%
- Generation length: 40 iterations
- Save population (genomes) at generations 10,100,1000,10000,100000
- Do 30 runs (lineages) of this (5 * 30 = 150 saved populations total)

Test Setup

- Recall that we wish to test whether more evolved populations can outperform less evolved populations
- Compare all sampled generations to each other
  - (10 vs. 100, 10 vs. 1000, ... , 10000 vs. 100000)
- Each of the two populations can be picked from 30 lineages
- So each pair of generations (g_i,g_j) gives us 30*30 = 900 pairwise comparisons between lineages

Test Parameters

- For each pairwise comparison
  - Turn off mutation
  - Randomly seed 10% of board with seeds drawn at random from two saved populations
  - Run for 50 generations and score by how many cells are occupied by plants from each population
- This gives two numbers.
- All 900 pairs of lineages give us two 30 x 30 matrices for each pair of generations.

Example of test between two generations- 10 and 100

- Test gives us two matrices: each tells score of runs from one generation

	Gen 10, run 1	Gen 10, run2	...	Gen 10, run 30
Gen 100, run 1	Score of gen 10	Score of gen 10	...	Score of gen 10
Gen 100, run 2	Score of gen 10	Score of gen 10	...	Score of gen 10
...	...	...	...	...
Gen 100, run 30	Score of gen 10	Score of gen 10	...	Score of gen 10

	Gen 10, run 1	Gen 10, run2	...	Gen 10, run 30
Gen 100, run 1	Score of gen 100	Score of gen 100	...	Score of gen 100
Gen 100, run 2	Score of gen 100	Score of gen 100	...	Score of gen 100
...	...	...	...	...
Gen 100, run 30	Score of gen 100	Score of gen 100	...	Score of gen 100

Statistical test

- First try: for each pair of generations:
  - Treat the two matrices as 900-entry vectors of paired data
  - Perform Wilcoxon signed rank test between the two
- This didn't work well: rows and columns not independent
- Second try:
  - Sum rows of each matrix to get two 30-entry vectors of paired data
  - Perform Wilcoxon signed rank test between them

Population Test

- All possible pairs of generations tested
- Table gives p-value indicating the probability that generation *top-row* had the same average score as generation *left-column*
- In all cases, if the p-value is low (< 0.01), generation *top-row* outscored *left-column*

	10	100	1000	10000	100000
10	0.975387	0.010444	0.000002	0.000002	0.000002
100		0.909931	0.000002	0.000002	0.000002
1000			0.877403	0.000028	0.000005
10000				0.643517	0.001287
100000					0.544006

Individual Test

- Is adaptation due to better individual strategies, or to interactions with other members of one's own population?
- To test this, instead of
  - ~~Randomly seed 10% of board with seeds drawn at random from two saved populations~~,
  - Randomly seed 10% of board with *copies of one seed* drawn at random from each saved population

	10	100	1000	10000	100000
10	0.909931	0.025637	0.000002	0.000002	0.000002
100		0.643517	0.000002	0.000002	0.000002
1000			0.926255	0.000020	0.000004
10000				0.614315	0.001036
100000					0.958990

Conclusions

- Non-local adaptation definitely occurs in gridplants
- In this system, adaptive abilities are stored in single genomes (recall each genome is only a 20 character string)
- Effect appears to be logarithmic
  - i.e. there exists a constant c such that populations evolved to g generations are statistically worse than populations evolved to less than $g \cdot c$ generations
  - This is why the generations were sampled at 10,100,1000,10000,100000

Related work

- Mayfield, Ashlock (1998)
  - Iterated prisoner's dilemma
- Gandrud, Ashlock (2003)
  - Paintbots
- Jordan (in progress)
  - Iterated prisoner's dilemma 2