

Lecture 13: GAs as Dynamical Systems

Dynamical Systems vs Markov Processes

- ▶ Treating a GA as a Markov Process gave us some powerful analytical tools
 - ▶ Limiting distribution over states
 - ▶ Limiting transition matrix
 - ▶ Expected time to reach an absorbing state
- ▶ With a little work we can adapt our Markov Process treatment of the GA to analyse it as a *dynamical system*
 - ▶ *Definition:* A dynamical system is a mathematical formalisation of a system whose trajectory in a space evolves according to some rule
 - ▶ By treating a GA as a dynamical system we can analyse it in terms of its trajectory in the space of possible populations

Population Dynamics

- ▶ For the Markov Chain analysis of a GA we represented a population as a vector

$$v = (v_0, v_1, \dots, v_{|C|-1})$$

where each v_k is the number of copies of point k in the search space, C and

$$\sum_{k=0}^{|C|-1} v_k = N$$

- ▶ We can remove the dependence on population size if we divide by N giving a population vector

$$p = (p_0, p_1, \dots, p_{|C|-1})$$

- ▶ Now the population vector represents the *proportion* of the population that are copies of each point in the search space
 - ▶ N.B. recall that v and hence p are actually *column* vectors

Population Dynamics

- ▶ The population vector p is a *unit vector*, i.e.

$$\sum_{k=0}^{|\mathcal{C}|-1} p_k = 1$$

- ▶ So all possible populations lie within the *unit simplex*
 - ▶ *Definition:* A simplex is the simplest geometrical shape that can be represented in an n -dimensional space
 - ▶ E.g. In 2-dimensions the simplex is a line
 - ▶ In 3-dimensions a triangle
 - ▶ In 4-dimensions a tetrahedron
 - ▶ Etc...
 - ▶ *Definition:* The unit simplex in an n -dimensional space is the shape having the n vertices $(1 \ 0 \ \dots \ 0)$, $(0 \ 1 \ \dots \ 0)$, ..., $(0 \ 0 \ \dots \ 1)$

Population Dynamics

- ▶ Possible populations are points within the simplex, but clearly such points can only correspond to points with *rational* coordinates
 - ▶ I.e. vectors of the form $p = \left(\frac{v_1}{N} \frac{v_2}{N} \dots \frac{v_{|C|}}{N} \right)$
- ▶ The simplex contains all points with real coordinates, i.e. rational *and* irrational

$$\Lambda = \left\{ x \in \mathbb{R}^n : x_k \geq 0 \text{ for all } k \text{ and } \sum_{k=0}^{|C|-1} x_k = 1 \right\}$$

- ▶ As $N \rightarrow \infty$ the set of points corresponding to possible populations becomes *dense* in the simplex, and the simplex is their *closure*

Population Dynamics

- ▶ To describe the evolution of the dynamical system, we need a *generational operator* mapping points in the simplex back into the simplex

$$\mathcal{G} : \Lambda \rightarrow \Lambda$$

- ▶ Apart from representing population proportions, another interpretation of the population vector p is a probability distribution over all the points in the search space
 - ▶ Then the generational operator gives the probability distribution over all possible populations in the next generation, given the current population
 - ▶ Thus the generational operator is the equivalent of the Markov Process transition matrix

Population Dynamics

- ▶ We can also interpret the generational operator as creating a population distribution
- ▶ *Theorem (Vose)*: If the population vector p is the current population, the expected next population is $\mathcal{G}(p)$
 - ▶ In finite populations stochastic effects will lead to deviations from the expected next population
 - ▶ The variance of this deviation will decrease as N increases
 - ▶ For infinite populations the variance will be zero so the population will behave deterministically
 - ▶ \mathcal{G} gives the infinite population behaviour
 - ▶ We can iterate the application of \mathcal{G} to calculate the (expected) population trajectory from its initial state

Population Dynamics - Selection

- ▶ Construction of the generational operator is done in the same way as construction of the transition matrix for Markov Chain analysis
- ▶ We start with selection only
 - ▶ From the previous lecture, the incidence vector based probability of selection is

$$P(i|v)_1 = \frac{v_i f(i)}{\sum_{j \in \mathcal{C}} v_j f(j)}$$

- ▶ Dividing through by N we obtain

$$P(i|p)_1 = \frac{p_i f(i)}{\sum_{j \in \mathcal{C}} p_j f(j)}$$

Population Dynamics - Selection Operator

- ▶ We can specify the selection operator \mathcal{F} in matrix form
 - ▶ If we consider our fitness function as a vector $f \in \mathbb{R}^{|\mathcal{C}|}$ whose entries $f_k = f(k \in \mathcal{C})$ then the selection operator is

$$\mathcal{F}(p) = \frac{\text{diag}(f)p}{f^T p}$$

where $\text{diag}(f)$ is the diagonal matrix with the entries from vector f on its diagonal, e.g.

$$\text{diag}([1 \quad 2 \quad 3]) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Population Dynamics - Fixed Points of Selection

- ▶ The vertices of the simplex are fixed points of fitness proportional selection
 - ▶ Each corresponds to a different uniform population
 - ▶ They are absorbing states under Markov Chain analysis
- ▶ Other fixed points also exist
- ▶ *Theorem (Convex Hull)*: If $A \subseteq \mathcal{C}$ is a set of individuals all having the same fitness value, then

$$\mathcal{H}\{\mathbf{e}_k : k \in A\} = \left\{ \sum_{k \in A} \alpha_k \mathbf{e}_k : \alpha_k \geq 0 \text{ for all } k \text{ and } \sum \alpha_k = 1 \right\}$$

is the *convex hull* of the vertices, and all members of this set are fixed points of fitness proportional selection

- ▶ where \mathbf{e}_k is the vector containing zeroes everywhere, apart from a 1 at the k -th index

Population Dynamics - Fixed Points of Selection

- ▶ So the fixed points of fitness proportional selection are
 - ▶ The vertices of the simplex
 - ▶ These correspond to all the possible uniform populations
 - ▶ Mixed populations (points inside the simplex) containing individuals with the same fitness
- ▶ The vertices of the simplex are always stable fixed points
- ▶ Mixed populations are only stable fixed points in the infinite population case
 - ▶ In the finite population case, finite population effects will lead the population away from mixed population fixed points
 - ▶ A finite population will eventually reach one of the vertices of the simplex through a combination of selection and *genetic drift*

Population Dynamics - Mutation

- ▶ Mutation can easily be incorporated into the generational operator, using the mixing matrix U as considered in the last lecture
- ▶ The mutation operator is thus

$$\mathcal{U}(p) = Up$$

- ▶ So the combined selection and mutation operator is

$$\mathcal{U} \circ \mathcal{F}(p) = \frac{U \text{diag}(f)p}{f^T p}$$

Population Dynamics - Crossover

- ▶ Crossover is also incorporated by using the mixing matrix $M(k)$ we saw in the last lecture
- ▶ The entries i, j of $M(k)$ are the probabilities that chromosomes i and j combine through crossover to produce chromosome k
 - ▶ Many crossover operators are not symmetric
 - ▶ E.g. UX with Bernoulli parameter $\neq 0.5$
 - ▶ However, we can construct a symmetric matrix M_k by taking mean probabilities, i.e. the entries i, j of M_k are

$$\frac{M_{i,j}(k) + M_{j,i}(k)}{2}$$

- ▶ With such a matrix the crossover operator is given by

$$\mathcal{X}(p)_k = p^T M_k p$$

Population Dynamics

- ▶ We have so far defined three main operators
 - ▶ \mathcal{F} — selection
 - ▶ \mathcal{U} — mutation
 - ▶ \mathcal{X} — crossover
- ▶ The mutation and crossover operators can be composed to give what is known as the *mixing operator*

$$\mathcal{M} = \mathcal{X} \circ \mathcal{U}$$

- ▶ The generational operator can be constructed by composing any of the operators, e.g.

$$\mathcal{G} = \mathcal{M} \circ \mathcal{F}$$

Population Dynamics - Fixed Points

- ▶ We have already analysed the fixed points for proportional selection
- ▶ In general we can find the fixed points of the dynamical system by finding the eigenvectors of the matrix representing the generational operator
 - ▶ E.g. for $\mathcal{G} = \mathcal{U} \circ \mathcal{F}$ the fixed points are the eigenvectors of the matrix

$$U \text{diag}(f)$$

- ▶ N.B. the eigenvectors must be normalised so their entries sum to one, to meet our condition for population vectors
- ▶ The corresponding eigenvalue of the eigenvector is the mean population fitness for that fixed point

Population Dynamics - Fixed Points

- ▶ Not all fixed points will necessarily correspond to real populations
 - ▶ The fixed points may be irrational, or not correspond to rational numbers with denominator N
 - ▶ They may also lie *outside* the simplex
 - ▶ E.g. eigenvectors with negative entries
 - ▶ However, a more general form of the Perron-Frobenius theorem from the last lecture is useful, and tells us that
 - ▶ If the matrix version of the operator \mathcal{G} has only positive entries then it has exactly one eigenvector (fixed point) inside the simplex
 - ▶ This is the *leading* eigenvector, i.e. the eigenvector with largest eigenvalue, or mean population fitness
 - ▶ This fixed point is a *global attractor*

Population Dynamics - Fixed Points

- ▶ While the single interior fixed point tells us where the infinite population GA will end up, the other fixed points are also important in any analysis of finite population behaviour
- ▶ To analyse this, we consider the *force* induced by the generational operator \mathcal{G} at a point in the simplex p

$$\|\mathcal{G}(p) - p\|$$

- ▶ So the force is the distance that a population moves under application of the generational operator
 - ▶ Near a fixed point the force will be very low
 - ▶ At a fixed point the force will be zero

Population Dynamics - Metastability

- ▶ A population at a point *inside* the simplex may pass very close to a fixed point just *outside* the simplex
 - ▶ The force on the population will be low in this vicinity, so a finite population may spend some time there
- ▶ A finite population may not be able to exactly 'hit' a fixed point
 - ▶ In the vicinity of the fixed point the force will also be low, and again the population may spend some time there
- ▶ Such population states as the above can be referred to as *metastable states*

Population Dynamics - Neutrality

- ▶ Connected sets of points within the simplex may have very low force on them
- ▶ Hence a finite population will tend to drift between these points according to stochastic effects
- ▶ Such a set of points can be referred to as a *neutral network*
 - ▶ Neutral networks have attracted some interest, particularly as a means to escape local optima

The Dynamical Systems Model

- ▶ What advantages does the Dynamical Systems formulation offer us?
- ▶ We have seen some elementary analyses
 - ▶ Analysis of trajectories
 - ▶ Analysis of fixed points other than the global attractor
 - ▶ Analysis of finite population behaviour around fixed points
- ▶ More advanced analyses are possible, e.g.
 - ▶ Convergence/nonconvergence properties of trajectories under different operators
 - ▶ Selection with mutation converges, but does crossover? What are the requirements for \mathcal{G} to be *focussed*?
 - ▶ Structure preserving properties of operators with a given representation
 - ▶ E.g. all bit-mask crossover operators provably respect schema membership
 - ▶ There are many open questions...