Journal of Theoretical Biology 340 (2014) 285–293

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Reproductive value in graph-structured populations

Wes Maciejewski

Department of Mathematics, The University of British Columbia, Vancouver, British Columbia, Canada

HIGHLIGHTS

- Introduces Fisher's reproductive value into the study of evolutionary graph theory.
- Proves basic properties of reproductive value on graphs.
- Calculates neutral fixation probability for any graph.

ARTICLE INFO

Article history:
Received 13 November 2012
Received in revised form
19 September 2013
Accepted 24 September 2013
Available online 1 October 2013

Keywords:
Fixation probability
Evolutionary graph theory
Evolutionary game theory
Reproductive value

ABSTRACT

Evolutionary graph theory has grown to be an area of intense study. Despite the amount of interest in the field, it seems to have grown separate from other subfields of population genetics and evolution. In the current work I introduce the concept of Fisher's (1930) reproductive value into the study of evolution on graphs. Reproductive value is a measure of the expected genetic contribution of an individual to a distant future generation. In a heterogeneous graph-structured population, differences in the number of connections among individuals translate into differences in the expected number of offspring, even if all individuals have the same fecundity. These differences are accounted for by reproductive value. The introduction of reproductive value permits the calculation of the fixation probability of a mutant in a neutral evolutionary process in any graph-structured population for either the moran birth–death or death–birth process.

1. Introduction

Population structure has, for some time, been recognized as an important factor in determining the outcome of an evolutionary process. Structure can act to arrange individuals and produce evolutionary outcomes not seen in well-mixed populations (Nowak and May, 1992). Early models considered an infinite number of islands of individuals, each linked by global dispersal (Wright, 1931). Subsequent work, like the stepping-stone model of Kimura and Weiss (1964) and Weiss and Kimura (1965), considered the spatial arrangement of these islands. These models were refined to the finite population case by considering a finite number of breeding demes linked by dispersal patterns (Levins, 1969, 1970). Drawing on these earlier models, evolutionary graph theory has emerged as a convenient framework for modelling population structure (Lieberman et al., 2005).

An evolutionary graph G is a collection of vertices V and edges E between them. The vertices are occupied by haploid individuals and the edges indicate who interacts with whom and where offspring disperse. Throughout this paper I denote vertices by v_i and the individual residing on v_i by i. It is possible that the vertices are linked by two sets of edges, one indicating interactions and the other, replacements (Ohtsuki et al., 2007), but these two sets are often assumed to coincide, as they do in this paper.

Since their introduction in Lieberman et al. (2005), evolutionary graphs have become a well-studied representation of structured populations. The exact features of graphs that promote, or work against, cooperation are, however, still elusive. For highly symmetric (vertex-transitive) graphs exact results for any additive game undergoing a weak-selection evolutionary process have been obtained (Ohtsuki et al., 2006; Taylor et al., 2007). This is the largest class of graphs for which results are known, encompassing many other results (Ohtsuki and Nowak, 2006; Grafen, 2007). Actual interaction graphs are often highly non-symmetric (Santos et al., 2008) and it is of great interest to study evolution in these environments.

Very few results have been obtained for non-symmetric graphs. There has been some interest in the role of vertex degree. Some work (Santos et al., 2008) has focused on the distribution of the degrees of vertices. Certain distributions (scale-free) have been shown to promote altruistic and cooperative behaviours more than others (e.g., regular graphs). These approaches have uncovered global features of graphs and a description of the process at the

---

E-mail address: wes@math.ubc.ca

0022-5193/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jtbi.2013.09.032

level of the individual is desirable. One of the challenges faced in the study of heterogeneous populations is dealing with individuals of differing quality. *Reproductive value* (Fisher, 1930) is a way of accounting for such differences.

Antal et al. (2006) are perhaps the first to consider heterogeneous graphs at the individual scale. They have found that it is advantageous for the fitter mutant to occupy high-degree nodes in a Moran death–birth model (their 'biased voter model') and lower-degree vertices in the birth–death process (their 'biased invasion process'). This has been confirmed by subsequent research (Broom et al., 2011). In the current paper I show that these results, when phrased in terms of reproductive value (Fisher, 1930; Grafen, 2006), are two sides of the same coin.

The work of Antal et al. (2006) and Broom et al. (2011) focuses on the case of *constant selection*, where the resident population has fecundity 1 and a mutant with fecundity $r>1$ arises. The probability of this mutant taking over the entire population is calculated and compared against the neutral case of $r=1$. If this mutant fixation probability is greater, the mutant is advantageous. An extension of the results of Antal et al. (2006) and Broom et al. (2011) to the case of a public-goods game, as in Santos et al. (2008), is highly desirable. I attempt to make headway by presenting an example that illustrates that a mutant individual can have greater evolutionary success depending on where it first emerges.

The main thrust of this paper is a complete description of the fixation probability of an allele in any graph-structured population undergoing neutral drift. For a structured population of size N with the property that all sites are equivalent—for example, degree-regular graphs—then this fixation probability is $1/N$, irrespective on which vertex the allele is first found. This is not the case for degree-heterogeneous graphs. In general, the fixation probability depends on the degree of the vertex on which the allele initially appears. In the current paper I calculate these fixation probabilities for both the birth–death and death–birth Moran processes on any graph. A general rule is derived: fixation probability is positively associated with relative reproductive value. An allele will have a higher fixation probability if it first emerges on a vertex with a higher reproductive value in both the birth–death and death–birth processes.

2. Reproductive value

Reproductive value has been defined in various ways by different authors. The core of the definitions is the notion of long-term genetic share of a population. Fisher (1930) first introduced reproductive value as a means of accounting for the differences in the reproductive output of different ages of females. Since that time reproductive value has been applied to age (Charlesworth, 1980), sex (Taylor, 1990), and spatially structured (Rogers and Willekens, 1978) populations and has been placed on a rigorous mathematical footing (Grafen, 2006). At an intuitive level, the relative reproductive value of an individual i is the probability that i is the ancestor of a randomly chosen individual in a distant future generation (Taylor and Frank, 1996).

To define reproductive value, I suppose that the individuals in the population under consideration are neutral with respect to selection. That is, the genotype of an individual does not affect their fitness. Births and deaths occur at random in the population. Throughout this paper I work with two Moran processes, which will be made explicit, that ensure a fixed population size. In the birth–death process, a birth occurs randomly in the population and the new offspring displaces a neighbouring individual, who dies. In the death–birth process an individual is chosen to die and a neighbouring individual is chosen at random to place an offspring on the newly vacated site. These birth and death probabilities are captured by a transition matrix M. Specifically, I define the i,j entry of M to be the probability p_{ij} that the current individual i is the offspring of individual j produced during a birth/death event. This entry will differ depending on whether births precede deaths or vice versa, and examples throughout the paper will illustrate this. An individual may be unaffected by the birth/death event in which case we say that such an individual is "from itself".

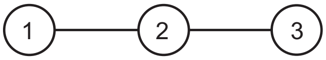

Fig. 1. The line graph on three vertices is the simplest example of a degree-heterogenous graph. Label the vertices v_1, v_2, and v_3 from left to right. As is shown in the penultimate section, spiteful behaviours can evolve in such a population structure and these depend on where the spiteful individual first emerges.

As a first example of such an M matrix, consider a birth–death process on the 3-line graph in Fig. 1. In the neutral process all individuals have the same fecundity and are therefore chosen to reproduce with equal probability, which in the 3-line case is 1/3. If the centre, or hub, individual is chosen, then it places an offspring on either leaf vertex with probability 1/2. If a leaf is chosen, its offspring disperses to the hub with probability 1. Given the current state of the population, we can ask where the individual on a leaf vertex was before a birth–death event. With probability 1/6, the individual is the offspring of the hub vertex and with probability 5/6 the individual was unaffected by the birth–death event and was already resident on the leaf vertex. For the hub individual, with probability 1/3 it came from one of the leaf vertices and with probability 1/3 it was unaffected by the birth–death event and already resident on the hub. In all, with the vertex numbering in Fig. 1,

$$M=\begin{bmatrix}\frac{5}{6} & \frac{1}{6} & 0\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ 0 & \frac{1}{6} & \frac{5}{6}\end{bmatrix}. \tag{1}$$

This matrix M can be used to find the vector of probabilities of the origin of the left-most leaf individual. Represent this individual with the vector $[1,0,0]$. This yields

$$[1,0,0]\begin{bmatrix}\frac{5}{6} & \frac{1}{6} & 0\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ 0 & \frac{1}{6} & \frac{5}{6}\end{bmatrix}=\left[\frac{5}{6},\frac{1}{6},0\right], \tag{2}$$

which captures the argument above: with probability 5/6 the leaf individual was unaffected by the birth–death event and with probability 1/6 it is an offspring of the hub individual. Another right-multiplication by M yields the probability vector for the generation two previous, and so on.

To find the probability that a randomly chosen individual in the population at a time t, measured in the number of birth/death events, in the future is from the lineage originating from individual i at the present time $t_0=0$, we perform a calculation similar to the above on the vector $[1,1,1]$:

$$[1,1,1]\begin{bmatrix}\frac{5}{6} & \frac{1}{6} & 0\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ 0 & \frac{1}{6} & \frac{5}{6}\end{bmatrix}^t. \tag{3}$$

This expression converges rapidly as t increases (Barton and Etheridge, 2011). Hence, the vector resulting from the calculation in Expression (3) above is stable to additional right-multiplications by M for sufficiently large t. This vector is the vector of reproductive values and when normalized, yields the probability distribution of the origin of a randomly chosen individual. This is captured