

Mathematics of Life

Henry Steyer

E-mail address: henry.steyer@freenet.de

URL: <http://www.mathematicsoflife.net>

Copyright © 2009-2012 Henry Steyer. All rights reserved

typesetting L^AT_EX

Contents

Preface	v
Chapter 1. The Relations Ancestor and Predecessor	1
Chapter 2. The Relation Parent	3
Chapter 3. Trees	4
Chapter 4. Dual Trees	5
Chapter 5. Asymmetric Trees	6
Chapter 6. Multiple Dual Trees	7
Chapter 7. Constructing a Dual Tree with a Double Tree	8
Chapter 8. Multilevel Forest	9
Chapter 9. Environment	10
Appendix A	
Cooperation and Group Fitness	11
Chapter 1. Single Inheritance Cooperation	12
Principle of group fitness and cooperation	12
Cooperation in artificial environments	12
Cooperation by depletion and cooperation with recombination	13
Chapter 2. Simple Models of Cooperation	14
Simple model of cooperation by depletion	14
Simple model of cooperation with recombination	15
Appendix B	
A Numerical Model	17
Chapter 1. Introduction	18
Calculation for a single trees population over one generation	18
Chapter 2. Single Tree Population	20
Chapter 3. Application	22
Revision	25

History

27

Preface

Life records genealogical trees which can be analysed with the mathematics of partially ordered sets or posets; the algebra of transitive asymmetrical relations. For genealogical trees, the relation is discretonal which allows to distinguish a predecessor from a precursor and to identify the former as parent. A multitude of such posets can be brought into relation. For instance, to the simple inheritance poset of the gene genealogy can be superposed the double inheritance one of the sexued reproduction and the multiple ones of some groups of individuals. The result is a multilevel poset. We will then define a function of the time with subsets of this multilevel poset as values. This function is the 'present of the living' for example, things which are already born and not yet died. From a given multilevel poset can be constructed all its possible developements and associate to them a probabilistic distribution, its 'environment'. In correspondence with the function of time, we will then be able to defined a 'reality'.

This book was originaly motivated by the question of cooperation and aimed to provide a theoretical frame where this questions may be placed. But it can also provide a contemplation of descendance for racialist and may also add new mathematical perspective. Aspects of cooperation and group fitness are treated independently in the backmatter. The article "For the Good of the Group", American Scientist, from David Sloan Wilson and Edward Osborn Wilson gives an approach to this questions. On D. S. Wilson's blog: <http://evolution.binghamton.edu/dswilson/publications>.

NOTE. This is a work in process with many uncertainties. The definitions may be adjusted and are valid only in the context of the theory exposed here.

CHAPTER 1

The Relations Ancestor and Predecessor

In a set E , the relation $a \triangleleft b$, or a is *precursor* of b , or b is *follower* of a , is transitive and asymmetric.

$$\begin{aligned} a \triangleleft b \wedge b \triangleleft c &\implies a \triangleleft c \\ a \triangleleft b &\implies b \not\triangleleft a \end{aligned}$$

In a subset A of E , $a \prec_A b$, or a is *predecessor* of b in A , if

$$\begin{aligned} a \triangleleft b \\ c \in A \implies a \not\triangleleft c \vee c \not\triangleleft b. \end{aligned}$$

An element a of A is *created in A* if it has no precursor. This is noted $* \prec_A a$.

$$c \in A \implies c \not\triangleleft a$$

Clearly, the relation \prec_A and the propriety $* \prec_A$ depend on the subset A of E .

An element a in E is *procreated* if it is created or has at least one predecessor in all the subsets of E in which it is element.

$$\forall A \subseteq E, a \in A \implies * \prec_A a \vee \exists c : c \prec_A a$$

REMARK. Even if an element in a set is created or has a predecessor, so not necessarily in all its subsets. For example, in the continuous set of numbers $(-1,1]$, we define \triangleleft by

$$a \triangleleft b \iff a = b - 1 \vee (a > 0 \wedge a < b).$$

In $(-1,1]$, all elements are created or have a predecessor, but none in its subset $(0,1]$. In, say, $[0.3, 0.5]$, only 0.3 is created and none have a predecessor.

A procreated element can have as or be predecessor of an infinite number of elements; for example in $[0,1]$ if \triangleleft is defined by

$$a \triangleleft b \iff a \neq b \wedge (a = 0 \vee b = 1),$$

0 is predecessor and 1 have as predecessor infinite elements.

A subset A of E is *complete* if

$$a \prec_A b \implies a \prec_E b.$$

The relation can then be noted \prec .

A subset P of E is a *path* if

$$a, b \in P \wedge a \neq b \implies a \triangleleft b \vee b \triangleleft a.$$

REMARK. In a path, a procreated element b can have an infinite number of precursors followers of its precursor a . For example in the set $\{1/n \cup 0 : n \in \mathbb{N}\}$ with

$$x \triangleleft y \iff x < y$$

and $a = 0, b \neq 0$.

THEOREM. *Let us considere a set E with a relation precursor.*

- (a) *In a finite subset of E , all elements are procreated.*
- (b) *If b is procreated in E then, in all subsets of E , if a is precursor of b , a is precursor of a predecessor or predecessor of b .*
- (c) *In a path, at most one element is created and every element has at most one predecessor. In a path, an element which is procreated but not created has exactly one predecessor.*
- (d) *If, for any a , in all pathes the elements between a and b are finite, b is procreated.*
- (e) *If the followers of a who are precursor of b are almost countable, there exists a complete path between a and b .*

PROOF. (a) is clear.

(b) Suppose that in a subset of E a is precursor of b and not precursor of a predecessor or predecessor of b . Because a is precursor but not predecessor of b , there is an element a_1 such that a is precursor of a_1 and a_1 is precursor of b . By our assumptions a_1 is not predecessor of b . There is an element a_2 which is follower of a_1 and precursor of b . a_2 can not be predecessor of b because a is precursor of a_2 . In this way we get an infinite sequence $\{a_n\}$. In this set augmented by b , b is not created and has no predecessor; therefore b is not procreated.

(c) In a path, for all a and b , $a \triangleleft b$ or $b \triangleleft a$. Therefore a and b can not be both created. With a similar argument, a and b can not be both predecessor of c . If c is procreated but is not created in this path, (b) shows that it has at least one predecessor.

(d) If b is not procreated, the demonstration of (b) shows that it exists an infinite path between an element a and b .

(e) Because the followers of a who are precursor of b are almost countable, they can be arbitrary arranged in a sequence $\{c_n\}$. A sequence $\{a_n\}$ can be constructed beginning with $a_1 = a, a_0 = b$ and an integer $k = 1$. If $a_j \triangleleft c_i$ and $c_i \triangleleft a_{j-1}$, $1 \leq j \leq k$, c_i is inserted before a_j and k is then incremented by one, else c_i does not belong to the partial sequence $\{a_k\}$. $\{a_n\}$ is clearly a path. Suppose $\{a_n\}$ is not complete. Then there is at least one element c_i of $\{c_n\}$ which is precursor of an a_{j-1} and a_j precursor of c_i , which is impossible by construction. \square

If in E all elements are procreated, \triangleleft is a relation *ancestor*. If \triangleleft is a relation *ancestor* and $a \triangleleft b$, we say a is *ancestor* of b and b is an *descendant* of a .

CHAPTER 2

The Relation Parent

A *forest* is a set with a relation *ancestor* and a relation *parent*. A parent is an ancestor and a predecessor is a parent. The relation parent is noted \triangleleft .

$$a < b \implies a \triangleleft b$$

$$a \prec b \implies a < b$$

A parent is *natural* if it is not ancestor of any other parent of its child (so far, no assumption has been made on the number of parent a child may have).

$$a < b$$

$$a \triangleleft c \implies c \not\triangleleft b$$

THEOREM. *A parent is natural if and only if it is predecessor of its child.*

PROOF. If the parent is not natural, it is not a predecessor.

$$a \triangleleft c \wedge c < b \implies a \not\triangleleft b$$

If the parent is not predecessor of its child, it is ancestor of one of the child's predecessor; a predecessor is necessarily parent. \square

COROLLARY. *It exists only one relation parent where all the parents are natural.*

PROOF. The relation predecessor (without index) is determined by the relation ancestor. \square

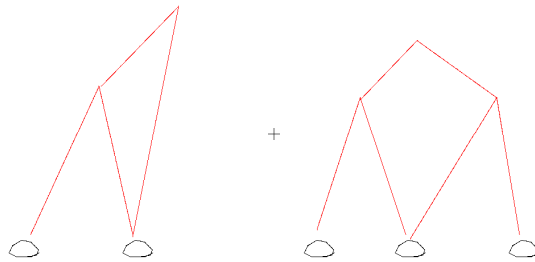


FIGURE 2.1. The relation parent is natural on the right figure

CHAPTER 3

Trees

\square denotes a partition of the poset A . $[a]$ is the class of the element a for this partition. The condition

$$\forall a, b; \quad [a] \neq [b] \implies a \not\leq b$$

or equivalently

$$\forall a, b; \quad a \triangleleft b \implies [a] = [b]$$

defines a set of partitions P . The partition *relative* or \square_{rel} is such that

$$\forall \square \in P, \forall a; \quad [a]_{rel} \subset [a].$$

The partition relative exists, is element of P and is unique.

A *phylogenetic tree* E is a class of relatives.

$$E \in A /_{rel}$$

N -uples refers to the number of parents a child may have. An *n-uple phylogenetic tree* is noted *nPT*. For $n=1$ it is a *single* and for $n=2$ a *double phylogenetic tree*.

$$card\{x < a \mid x \in E\} \leq n; \quad n \in N$$

CHAPTER 4

Dual Trees

The subset of all the elements of one or more single tree from a forest, if its size is even or infinite, can be partitioned in unordered pairs. The corresponding equivalence relation is noted \parallel .

$$A = \uplus_i S_i \quad S_i \text{ is a } 1PT$$

$$p \in A/\parallel \implies \text{card } p = 2$$

To these pairs is associated the relation $<\parallel$ defined by the condition

$$\{a, b\} <\parallel \{c, d\} \iff a < c \vee a < d \vee b < c \vee b < d,$$

and the relation $<\parallel$ which satisfies

$$p <\parallel q \iff p <\parallel q \vee (\exists r : p <\parallel r \wedge r <\parallel q),$$

and is therefore transitive. If $<\parallel$ is asymmetric, it is a relation ancestor, the relation $<\parallel$ the *associated* relation parent and the set a *dual phylogenetic tree*.

$$p <\parallel q \implies q \not<\parallel p$$

THEOREM. *In a dual phylogenetic tree, if $p <\parallel q$, an element of q can not be an ancestor of an element of p .*

$$a \in p \wedge b \in q \wedge p <\parallel q \implies b \not< a$$

PROOF. If b is ancestor of a , q is ancestor of p which contradicts the fact that p is ancestor of q . □

If the two elements of a pair p does not have the same parent (for $<$), the relation $q <\parallel p$ is *integre*.

$$a \parallel b \wedge a \neq b \wedge c < a \implies c \not< b$$

If the relation is integre, the parents of a pair form a *couple*. This relation is noted \circ .

$$a \parallel b \wedge a \neq b \wedge c < a \wedge d < b \implies [c]_{\parallel} \circ [d]_{\parallel}$$

If the relation couple is reflexive, its child is a *clone*.

$$\{a, b\} <\parallel \{c, d\} \wedge ((a < c \wedge b < d) \vee (a < d \wedge b < c))$$

If the pair have no or only one parent, it is *orphan* or *half-orphan*.

PROPOSITION. *A dual tree is one ore more double tree for $<\parallel$.*

CHAPTER 5

Asymmetric Trees

A single phylogenetic tree S is possibly *monal* or *virtually monal* if it has at most one element in each pair of the dual tree.

$$a \in S \wedge a \in [b]_{\parallel} \wedge a \neq b \implies b \notin S$$

If each pair include at most and each couple whose child is not a clone exactly one element of a monal tree (the monal tree needs not to be same for all couple), the dual tree is an *asymmetric phylogenetic tree*.

$$\begin{aligned} S_i \in [S] \\ \{a, b\} \neq \{c, d\} \implies (a \in S_1 \iff b \notin S_2 \wedge c \notin S_3 \wedge d \notin S_4) \\ \{a, b\} = \{c, d\} \implies (a \in S_1 \implies b \notin S_2) \end{aligned}$$

The pair with the monal tree is the m pair and the one without the w pair.

THEOREM. *There exists only one path from a m child to his m ancestor where all the link are m pair and the same apply for the w pairs.*

CHAPTER 6

Multiple Dual Trees

Dual trees are a *n-uple dual phylogenetic tree* when their pairs are in a relation of equivalence so that

$$\begin{aligned} \text{card}[q]_{\sim} &= n \\ p \sim q &\implies \neg(p \circ q) \\ p \sim q \wedge q \circ r \wedge r \sim s \wedge [p]_{rel} &\implies p \circ s \end{aligned}$$

A multiple dual tree is asymmetric if one of its dual tree is asymmetric.

CHAPTER 7

Constructing a Dual Tree with a Double Tree

Each element of a double phylogenetic tree is divided in a pair. At least one element of a parent pair is arbitrary chosen and connected to only one element that is not yet in a relation child of the child pair. The result is a dual phylogenetic tree. The pairs without precursor in the double tree are orphan. Those which only one parent are half-orphan or clone, depending on that one or two elements of the parent pair are connected.

$$\begin{aligned} & A \text{ is a } 2PT \\ & f : A \xrightarrow{\text{bijection}} E/\parallel \\ & \{a, b\} = f(p), \{c, d\} = f(q) \\ & p < q \iff (a < c \vee a < d) \vee (b < c \vee b < d) \end{aligned}$$

CHAPTER 8

Multilevel Forest

A *life function* of the forest E is a function defined on a subset of the ordered set of *time* which values are collections of subsets of E .

Life of the forest E is a set L of life functions with at least the life functions *born*, \mathcal{B} , and *deceided*, \mathcal{D} .

$$\begin{aligned} a \in \mathcal{B}_t \wedge b \in E \setminus \mathcal{B}_t &\implies b \not\prec a \\ &\mathcal{D}_t \subseteq \mathcal{B}_t \\ s < t &\implies \mathcal{B}_s \subseteq \mathcal{B}_t \wedge \mathcal{D}_s \subseteq \mathcal{D}_t \end{aligned}$$

Life is *continuously* if

$$\forall a \in E \quad \exists t : \forall b \in E \quad b < a \implies a, b \in \mathcal{B}_t \setminus \mathcal{D}_t.$$

Life is *condensed* if

$$\forall \mathcal{L} \in L \quad t, u \in \text{dom}(\mathcal{L}) \wedge t < u \implies \mathcal{L}_t \neq \mathcal{L}_u.$$

A life function of E is *immanent* if

$$\mathcal{L}_t \subseteq \mathcal{B}_t \setminus \mathcal{D}_t.$$

If for all t , \mathcal{L}_t is a partition of $\mathcal{B}_t \setminus \mathcal{D}_t$, \mathcal{L} it is a *group* function.

If a life function has single sets as values, it is a *propriety*.

CHAPTER 9

Environment

A forest a is a *germen* of the forest b , or $a \dot{\subset} b$, if all elements of a are elements of b and if for any two elements e_1 and e_2 of a the relation parent is the same in both sets.

$$a \subseteq b$$

$$e_1, e_2 \in a \implies (e_1 <_a e_2 \iff e_1 <_b e_2)$$

For a set of forests L , a *reality* is a n -uple of L^n in which each forest at a position in the suite is the germen of the forest at a later one. A position in the suite is referred as *generation* beginning with the generation zero. The pair (a, p) , which associates a forest with its generation, is a *conscience*.

An *environment* is a distribution on a set of realities. In an environment, $\Pr \{(a, p), \dots\}$ represents the probability of all the realities which include the consciences $(a, p) \dots$. It follows that

$$\sum_{a \in L} \Pr \{(a, p)\} = 1.$$

An *instant* is a distribution on a set of forest. Hence to an environment corresponds a suite of instants $m_0, m_1, \dots, m_p, \dots$ with

$$m_p(a) = \Pr \{(a, p)\}.$$

THEOREM. *For a suite of instants in an environment, it holds*

$$p < q \implies \sum_{f \dot{\subset} a} m_p(f) \geq m_q(a).$$

PROOF.

$$\Pr \{(a, q)\} = \sum_{f \in L} \Pr \{(f, p), (a, q)\} = \sum_{f \dot{\subset} a} \Pr \{(f, p), (a, q)\} \leq \sum_{f \dot{\subset} a} \Pr \{(f, p)\}$$

□

An environment is *deterministic* if

$$(r < p \wedge r < q) \vee (r > p \wedge r > q) \implies$$

$$\Pr \{(a, r), (x, p), (y, q)\} * \Pr \{(b, r), (x, p)\} = \Pr \{(b, r), (x, p), (y, q)\} * \Pr \{(a, r), (x, p)\}.$$

An environment which is not deterministic is *historical*.

Appendix A
Cooperation and Group Fitness

CHAPTER 1

Single Inheritance Cooperation

In this chapter, the conceptualization of cooperation is limited to single inheritance posets.

Principle of group fitness and cooperation

Group fitness is made of an ordered sample with replacement of three phases: the elements births or deaths inside the group (I), the groups division (T), the groups selected as unity (U). To each position in the suite, called a generation, corresponds a partition of the elements' entire population in groups. The elements are naturally connected by descent thus, given a grounding generation, to any choice of elements (which does not need to coincide with a group of the partition) corresponds the set of their descents at a later one. In particular, the complete population corresponds to itself.

Given a generation and a later one, the general fitness g of a set of elements is the relative change of its size which is the product of the relative change at each generation. g can be decomposed in a term p for the phases I, and a term q for U while T has always the value 1 and does not contribute.

$$g = \prod_i g_i = \prod_i p_i * \prod_j q_j = p * q$$

DEFINITION. For two sets a and b , a is *more cooperative* than b if $q_a > q_b$.

$$q_a > q_b \iff p_a/g_a < p_b/g_b.$$

The set of all elements is e . The *cooperation factor* c for a set a is

$$c_a \equiv q_a/q_e.$$

A set a is *cooperative* if

$$c_a > 1.$$

Cooperation in artificial environments

The fitness of a set depends on the environment. In two environments o_1 and o_2 , its general fitness for example is noted g_1 and g_2 . The index e indicates the set of all elements. If the environmental change does not affect the selection inside the groups, so that $p_1 = p_2$ and $p_{e1} = p_{e2}$, then

$$g_2/g_{e2} > g_1/g_{e1} \iff c_2 > c_1.$$

If, further, in the environment o_1 there is no selection or a random selection between groups, so that $q_1 = q_{e1}$, it holds

$$g_2/g_{e2} > g_1/g_{e1} \iff c_2 > 1.$$

Thus:

THEOREM. *A cooperative set of elements is less succesful when there is no longer selection pressure on the groups and this change does not affect the selection inside the group.*

NOTE. For the math of the two formulae consider

$$g_2/g_{e2} = g_1/g_{e1} \iff p_2q_2/p_{e2}q_{e2} = p_1q_1/p_{e1}q_{e1}.$$

Cooperation by depletion and cooperation with recombination

If the elements of a child group can come from different parent groups it is cooperation with recombination (and depletion) and if not it is cooperation by depletion (only).

CHAPTER 2

Simple Models of Cooperation

Usually evolutionary models involve many generations and assume that the number of all the elements grow till a limit or at least is not zero. Here, the two models go only for one generation and could evolve to extinction. This do not contradict the assumptions made for cooperation to be. As minimalistic but nevertheless working models, they expose the fundamental concepts. They also show the parallel between cooperation by depletion and cooperation with recombination.

Simple model of cooperation by depletion

A set of two elements generates two sets from which one is the empty set and the other a set with one element or equivalently one of the element is removed inside the group and then the group divided in itself and the empty set (phases I and T). The empty set or the set with one element is then removed (phase U). This describes a group selection process.

$\{e_1, e_2\}$ generates $\{e_1\}$ and \emptyset or $\{e_2\}$ and \emptyset . The probabilities for one generation of groups are

$$\begin{aligned} p_1 &= \Pr(\{e_1\}, \emptyset) \\ p_2 &= \Pr(\{e_2\}, \emptyset) \end{aligned}$$

After the groups are selected

$$\begin{aligned} q_1 &= \Pr(\{e_1\} \mid \{e_1\}, \emptyset) \\ q_2 &= \Pr(\{e_2\} \mid \{e_2\}, \emptyset) \end{aligned}$$

The fitness for e_1 and e_2

$$\begin{aligned} g_1 &= p_1 * q_1 \\ g_2 &= p_2 * q_2 \end{aligned}$$

Simple model of cooperation with recombination

The sets $\{e1, e2\}$ and $\{e\}$ generate the sets $\{e1, e\}$ and \emptyset or $\{e2, e\}$ and \emptyset . One set is then removed.

The probabilities for one generation of groups are

$$p_1 = \Pr(\{e1, e\}, \emptyset)$$

$$p_2 = \Pr(\{e2, e\}, \emptyset)$$

After the groups are selected

$$q_1 = \Pr(\{e1, e\} \mid \{e1, e\}, \emptyset)$$

$$q_2 = \Pr(\{e2, e\} \mid \{e2, e\}, \emptyset)$$

The fitness for $e1$ and $e2$

$$g_1 = p_1 * q_1$$

$$g_2 = p_2 * q_2$$

Appendix B

A Numerical Model

CHAPTER 1

Introduction

Two groups have the same population if to each individual in a group corresponds a clone by descent in the other group. A population based on single philogenetic trees is used for illustration. In this kind of population an individual is identified by an element of the philogenetic tree it belongs. From a group $\{e1, e2\}$ there is among others the following possible phylogenies:

$$\begin{aligned} a : \{e1, e2\} &\rightarrow 1a\{e1\} \rightarrow 2a\{ \\ b : \{e1, e2\} &\rightarrow 1b\{e1, e1\} \rightarrow 2b\{e1\} \\ c : \{e1, e2\} &\rightarrow 1c\{e1, e2\} \rightarrow 2c\{e1\} \\ d : \{e1, e2\} &\rightarrow 1d\{e1, e2, e2\} \rightarrow 2d\{e2\} \end{aligned}$$

The groups $1a\{e1\}$ $2b\{e1\}$ and $2c\{e1\}$ have the same population $P = \{e1\}$. The groups $2b\{e1\}$ and $2c\{e1\}$ are the possible realisation of the population P in the second generation.

Studied is the probability that after a given number of generations from an initial group originates a particular population.

Calculation for a single trees population over one generation

DEFINITION. • In a set an individual is represented with e . e' is the underset with all individual having the same descent than e . In the founding generation this undersets are determined arbitraly.

- $1(y, p_1, \dots, p_x)$ is a function equal to 1 if $p_1 + \dots + p_x = y$ else equal to 0. If $x = 0$ then if $y = 0$ the function is equal to 1 else it is equal to 0.

$$\sum_{p_1, \dots, p_x} 1(y, p_1, \dots, p_x) * dp_1 * \dots * dp_x$$

is the sum of all the term $1(y, p_1, \dots, p_x) * dp_1 * \dots * dp_x$ where the index p_1, \dots, p_x are the permutations with replacement of x integers ranging from 0 to y .

- To each individual is associated his birth distribution which is the distribution of the number of his children: $d0, d1, \dots$. $d0$ is the probability that he has no children, $d1$ one and so one. Like every distribution $d0 + d1 + d2 + \dots = 1$. The sum $0 * d0 + 1 * d1 + 2 * d2 + \dots$ is his birth-rate.

CONDITION. The birth distribution of an individual does not depend on the actual number of children of any other individual. This can also be said this way: if the actual number of children of an individual is known that does not change in the current generation the birth distribution for any other individual.

If an individual has a birth-rate over 1 the number of his descendants statistically rises and under 1 that it sinks. The maximal size of the next group is the add up of the maximal number of possible descendants of each individual. The minimal possibly number of the next group is the sum of the minimal number of possible descendants. If for every individual $d0$ is not null it is the empty set and the group can become extinct. The maximal possible number of children for any individual is supposed to be s . From a first group with a population of n individual from m different descent (which is the number of the different undersets e') one forms all the possible set from 0 to $n * s$ individual from a choice of m individual from different descent while each individual can occurs several times. A subset e' with y individual has the probability to come from a subset e' with x individual of:

$$\sum_{p_1, \dots, p_x} 1(y, p_1, \dots, p_x) * dp_1 * \dots * dp_x$$

The values for all subsets e' making up the next group are then multiplied. It is the probability that this particular group originates. j identifies all the groups e' in the generated group. $k1$ identifies the originating and $k2$ the originated group.

$$\prod_j \sum_{p_1, \dots, p_x}^{x=N(e'jk1)} 1(y = N(e'jk2), p_1, \dots, p_x) * dp_1jk1 * \dots * dp_xjk1$$

It is possible to give each individual from a group e' a different birth distribution. The calculation proceeding is the same but the factor of $dp_1jk1 * \dots * dp_xjk1$ is given an index: $dp_1i_1jk1 * \dots * dp_xi_xjk1$ with $a \neq b \iff i_a \neq i_b$. Because the order of the factor in the expression $dp_1jk1 * \dots * dp_xjk1$ does not play any role it is possible to use the same index x : $dp_11jk1 * \dots * dp_xxjk1$.

$$\prod_j \sum_{p_1, \dots, p_x}^{x=N(e'jk1)} 1(y = N(e'jk2), p_1, \dots, p_x) * dp_11jk1 * \dots * dp_xxjk1$$

According to the condition it remains that the probability that a given population comes from an other only depends on the birth distribution every individual. By replacing the factors with the product symbol and if $1(y, p_1, \dots, p_x) = 1(y)$. The formula is then written:

$$\prod_j \sum_{p_1, \dots, p_x}^{x=N(e'jk1)} \prod_{1 \leq i \leq x} 1(N(e'jk2)) * dp_i ijk1 = \prod_j \sum_{p_1, \dots, p_x}^{x=N(e'jk1)} 1(N(e'jk2)) * \prod_{1 \leq i \leq x} dp_i ijk1$$

CHAPTER 2

Single Tree Population

In this chapter the probability that a population originate from a group over more than one generation is studied. The generation index is t . Each individual is identified with $eijk_t$ and his birth distribution is:

$$d(eijk_t) = d0ijk_t, d1ijk_t, \dots$$

The probability that from a population k_1 in a generation t_1 comes the population k_2 in a generation t_2 is expressed as:

$$w((Pk_1t_1, Pk_2t_2)) = wk_1t_1k_2t_2$$

The probability that from a subset $e'j$ from a population k_1 in the generation t comes $e'j$ in the next generation symbolised with t^{+1} in the population k_2 is:

$$v(e'jk_1t, e'jk_2t^{+1}) = vjk_1k_2t$$

$$= \sum_{p_1, \dots, p_x}^{x=N(e'jk_1)} 1(y = N(e'jk_2), p_1, \dots, p_x) * dp_1 1jk_1t * \dots * dp_x xjk_1t$$

For the whole population k_1 and k_2 it is:

$$\begin{aligned} wk_1t_1k_2t_2^{+1} &= \prod_j vjk_1k_2t \\ &= \prod_j \sum_{p_1, \dots, p_x}^{x=N(e'jk_1)} 1(y = N(e'jk_2), p_1, \dots, p_x) * dp_1 1jk_1t * \dots * dp_x xjk_1t \\ &= \prod_j \sum_{p_1, \dots, p_x}^{x=N(e'jk_1)} \prod_{1 \leq i \leq x} 1(N(e'jk_2)) * dp_i ijk_1t \end{aligned}$$

To determine the probability of a population over several generations the path method is applied. All the elements on each path are multiplied together and then all this values are summed up.

$$\begin{aligned}
wk_1t_1k_2t_2 &= \sum_{k_1=k_0, k_1, \dots, k_n=k_2}^{t_2-t_1=n} \prod_t^{0 \leq t < n} \prod_j v_j k_t k_{t+1} t_1^{+t} \\
&= \sum_{k_1=k_0, k_1, \dots, k_n=k_2}^{t_2-t_1=n} \prod_t^{0 \leq t < n} \prod_j \\
&\quad \sum_{x=N(e'jk_t)}^{p_1, \dots, p_x} 1(y = N(e'jk_{t+1}), p_1, \dots, p_x) * dp_1 j k_t * \dots * dp_x x j k_t \\
&= \sum_{k_1=k_0, k_1, \dots, k_n=k_2}^{t_2-t_1=n} \prod_t^{0 \leq t < n} \prod_j \sum_{p_1, \dots, p_x}^{x=N(e'jk_t)} \prod_{1 \leq i \leq x} I(N(e'jk_{t+1})) * dp_i i j k_t
\end{aligned}$$

$wk_1t_1k_2t_2$ is the probability that from the population k_1 at the generation t_1 comes the population k_2 at the generation t_2 .

CHAPTER 3

Application

The group $G_0 = \{e_1, e_2, e_3\}$ is the founding generation with e_2 having the same genotype as e_3 . The maximal number of children is 2. The distribution for e_1 in all groups is given with $d_0 = 1/4$ $d_1 = 3/4$ $d_2 = 0$. The distribution for e_2 and e_3 is for all groups but $\{e_1, e_2\}$ $1/2$ $1/3$ $1/6$. For the group $\{e_1, e_2\}$ it is given with $1/7$ $4/7$ $2/7$.

What is the probability of the population $\{e_2\}$ in the second generation?

Because e_3 and e_2 have the same genotype e_2 is equated with e_3 .

This 15 groups can come from $G_0 = \{e_1, e_2, e_2\}$:

$\{\}, \{e_2\}, \{e_2, e_2\}, \{e_2, e_2, e_2\}, \{e_2, e_2, e_2, e_2\},$
 $\{e_1\}, \{e_1, e_2\}, \{e_1, e_2, e_2\}, \{e_1, e_2, e_2, e_2\}, \{e_1, e_2, e_2, e_2, e_2\}$
 $\{e_1, e_1\}, \{e_1, e_1, e_2\}, \{e_1, e_1, e_2, e_2\}, \{e_1, e_1, e_2, e_2, e_2\}, \{e_1, e_1, e_2, e_2, e_2, e_2\}$

Because the maximal number of children is 2 the other sets like $\{e_1, e_1, e_1, e_2\}$ formed with e_1 or e_2 cannot come from $\{e_1, e_2\}$ and do not need to be counted for.

$$\begin{aligned}
&= \sum_{k_1=k_0, k_1, \dots, k_n=k_2}^{t_2-t_1=n} \prod_t^{0 \leq t < n} \prod_j \sum_{p_1, \dots, p_x}^{x=N(e'jk_t)} \prod_{1 \leq i \leq x} 1(N(e'jk_{t+1})) * dp_i j k_t \\
&= ((d * \dots + p + \dots) * j * \dots) * t * \dots + k + \dots \quad d(e_1) = 1/4, 3/4, 0 \quad d(e_2) = 1/2, 1/3, 1/6 \text{ or } 1/7, 4/7, 2/7 \text{ for } \{e_1, e_2\} \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1\} \rightarrow \{e_2\} \quad 0 \\
((1/2) * (1/2 * 1/2)) * ((1) * (0)) & \quad 1/36 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_2\} \rightarrow \{e_2\} \\
((1/4) * (1/2 * 1/3 + 1/3 * 1/2)) * ((1) * (1/3)) & \quad 5/216 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_2, e_2\} \rightarrow \{e_2\} \\
((1/4) * (1/2 * 1/6 + 1/3 * 1/3 + 1/6 * 1/2)) * ((1) * (1/2 * 1/3 + 1/3 * 1/2)) & \quad 1/144 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_2, e_2, e_2\} \rightarrow \{e_2\} \\
((1/4) * (1/3 * 1/6 + 1/6 * 1/3)) * ((1) * (1/2 * 1/2 * 1/3 + 1/2 * 1/3 * 1/2 + 1/3 * 1/2 * 1/2)) & \quad 1/864 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_2, e_2, e_2, e_2\} \rightarrow \{e_2\} \\
((1/4) * (1/6 * 1/6)) * ((1) * (1/2 * 1/2 * 1/2 * 1/3 + 1/2 * 1/2 * 1/3 * 1/2 + 1/2 * 1/3 * 1/2 * 1/2 + 1/3 * 1/2 * 1/2 * 1/2)) & \quad 0 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1\} \rightarrow \{e_2\} \\
((3/4) * (1/2 * 1/2)) * ((1/4) * (0)) & \quad 1/28 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1, e_2\} \rightarrow \{e_2\} \\
((3/4) * (1/2 * 1/3 + 1/3 * 1/2)) * ((1/4) * (4/7)) & \quad 5/288 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1, e_2, e_2\} \rightarrow \{e_2\} \\
((3/4) * (1/2 * 1/6 + 1/3 * 1/3 + 1/6 * 1/2)) * ((1/4) * (1/2 * 1/3 + 1/3 * 1/2)) & \quad 1/192 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1, e_2, e_2, e_2\} \rightarrow \{e_2\} \\
((3/4) * (1/3 * 1/6 + 1/6 * 1/3)) * ((1/4) * (1/2 * 1/2 * 1/3 + 1/2 * 1/3 * 1/2 + 1/3 * 1/2 * 1/2)) & \quad 1/1152 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1, e_2, e_2, e_2\} \rightarrow \{e_2\} \\
((3/4) * (1/6 * 1/6)) * ((1/4) * (1/2 * 1/2 * 1/2 * 1/3 + 1/2 * 1/2 * 1/3 * 1/2 + 1/2 * 1/3 * 1/2 * 1/2 + 1/3 * 1/2 * 1/2 * 1/2)) & \quad 0 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1, e_1\} \rightarrow \{e_2\} \\
((0) * (1/2 * 1/2)) * ((1/4 * 1/4) * (0)) & \quad 0 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1, e_1, e_2\} \rightarrow \{e_2\} \\
((0) * (1/2 * 1/3 + 1/3 * 1/2)) * ((1/4 * 1/4) * (1/3)) & \quad 0 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1, e_1, e_2, e_2\} \rightarrow \{e_2\} \\
((0) * (1/2 * 1/6 + 1/3 * 1/3 + 1/6 * 1/3)) * ((1/4 * 1/4) * (1/2 * 1/3 + 1/3 * 1/2)) & \quad 0 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1, e_1, e_2, e_2, e_2\} \rightarrow \{e_2\} \\
((0) * (1/3 * 1/6 + 1/6 * 1/3)) * ((1/4 * 1/4) * (1/2 * 1/2 * 1/3 + 1/2 * 1/3 * 1/2 + 1/3 * 1/2 * 1/2)) & \quad 0 \\
&\{e_1, e_2, e_2\} \rightarrow \{e_1, e_1, e_2, e_2, e_2, e_2\} \rightarrow \{e_2\} \\
((0) * (1/6 * 1/6)) * ((1/4 * 1/4) * (1/2 * 1/2 * 1/2 * 1/3 + 1/2 * 1/2 * 1/3 * 1/2 + 1/2 * 1/3 * 1/2 * 1/2 + 1/3 * 1/2 * 1/2 * 1/2)) & \quad 4/1573
\end{aligned}$$

Revision

History

2012-10-13. Chapter 1 revised.

2012-05-20. Chapter 3 revised.

2011-08-29. Part 1 revised

2011-05-21. Part 2, Chapter 1

$$(r < p \wedge r < q) \vee (r > p \wedge r > q) \implies$$

$$\Pr \{(a, r), (x, p), (y, q)\} * \Pr \{(b, r), (x, p)\} = \Pr \{(b, r), (x, p), (y, q)\} * \Pr \{(a, r), (x, q)\}.$$

becomes

$$(r < p \wedge r < q) \vee (r > p \wedge r > q) \implies$$

$$\Pr \{(a, r), (x, p), (y, q)\} * \Pr \{(b, r), (x, p)\} = \Pr \{(b, r), (x, p), (y, q)\} * \Pr \{(a, r), (x, p)\}.$$

2011-05-20. Part 2, Chapter 1

$$\Pr \{(a, p)\} = \sum_{f \in L} \Pr \{(f, p), (a, q)\} = \sum_{f \dot{c} a} \Pr \{(f, p), (a, q)\} \leq \sum_{f \dot{c} a} \Pr \{(f, q)\}$$

becomes

$$\Pr \{(a, q)\} = \sum_{f \in L} \Pr \{(f, p), (a, q)\} = \sum_{f \dot{c} a} \Pr \{(f, p), (a, q)\} \leq \sum_{f \dot{c} a} \Pr \{(f, p)\}$$

2011-05-15. Appendix A, Chapter 1, Subsection 2

$$g_2 > g_{e2} \iff c_2 > 1$$

becomes

$$g_2/g_{e2} > g_1/g_{e1} \iff c_2 > 1.$$

Part 2 added.