МИТЧЕЛЛ М. ТОЧНЫЕ МАТЕМАТИЧЕСКИЕ МОДЕЛИ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ[1]

Теория шаблонов позволяет предсказать ожидаемые изменения частот шаблонов от одного поколения к другому. Но она не позволяет напрямую предсказать состав популяции, скорость сходимости или распределение приспособленности в популяции через длительный промежуток времени. В качестве первого шага для обеспечения более детального понимания ГА и более точного прогнозирования его поведения некоторые исследователи разрабатывают «точные» математические модели простых ГА (например, [Goldberg, 1987; Goldberg, Segrest, 1987; Davis, Principe, 1991; Vose, Liepins, 1991; Nix, Vose, 1991; Horn, 1993; Vose, 1993; Whitley, 1993a]). Эти точные модели описывают каждый элемент простого ГА посредством математических операторов. Таким образом, имея разработанную модель, появляется возможность доказывать теоремы о некоторых интересных свойствах этих операторов. В данном разделе я [М.Митчелл – Ю.Ц.] в общем виде опишу модель, предложенную Восом и Липинсом (1991) и обобщу расширения модели, сделанные Никсом и Восом (1991) и Восом (1993).

Формализация ГА

Математики Майкл Вос и Гунар Липинс (Michael Vose, Gunar Liepins, 1991) разработали формальную модель, основанную на следующем простом ГА:

Имеется случайная популяция бинарных строк длины l.
1. Вычислить приспособленность $f(x)$ каждой строки x в популяции.
2. Выбрать (с замещением) две родительские особи из текущей популяции с вероятностью пропорциональной относительной приспособленности каждой строки популяции.
3. Скрестить с вероятностью p_c родительские особи (с выбором одной случайной точки разрыва) для создания двух потомков. Если кроссовер не производится, то потомки являются точными копиями родителей. Выбрать случайно одного из двух потомков и исключить другого.
4. Мутировать каждый бит выбранного потомка с вероятностью p_m и поместить в новую популяцию.
5. Если новая популяция не сформирована, то перейти на шаг 2.
6. Перейти на шаг 1.

[1] Глава из книги Mitchell M. An Introduction to Genetic Algorithms. Fifth printing. Cambridge, MA: The MIT Press, 1999.
Перевел Ю.Р. Цой. Любые замечания, касающиеся перевода, просьба присылать по адресу qai@mail.ru
Данный текст доступен по адресу: http://qai.narod.ru/GA/mitchell_4_3.pdf

Единственная разница между описанным и стандартным простым ГА заключается в том, что выживает только один потомок для каждого использования оператора кроссовера. Тогда для популяции размера n всего имеют место n операций рекомбинации. Данная модификация позволяет упростить формализацию[1].

В формальной модели Воса и Липинса каждой строке в пространстве поиска поставлено в соответствие целое число i[2] в диапазоне от 0 до (2^l-1), закодированное в строке. Например, для l = 8, строке 00000111 будет соответствовать число 7. Популяция в поколении t представлена двумя вещественными векторами, содержащие 2^l компонент каждый. i-я компонента первого вектора (обозначенная p_i (t)) равна доле строк i-го типа, а i-я компонента второго вектора (обозначенная s_i (t)) равна вероятности выбора строки i-го типа в качестве родительской на втором шаге простого ГА, согласно приведенной выше схеме. К примеру, если l = 2 и популяция состоит из двух копий строки «11» и одной копии строк «01» и «10», то

$$\vec{p}(t) = \{0, 0.25, 0.25, 0.5\}.$$

Если приспособленность строки равна количеству содержащихся в ней единичных разрядов, то

$$\vec{s}(t) = \{0, 0.1667, 0.1667, 0.6667\}.$$

Для использования операции умножения на матрицу будем считать эти вектора вектор-столбцами, хотя часто они будут записаны как вектор-строки.

Вектор $\vec{p}(t)$ точно определяет состав популяции в поколении t, а вектор $\vec{s}(t)$ задает вероятности выбора строк для скрещивания для данной функции приспособленности. Эти вектора объединяются через приспособленность. Пусть F является двумерной матрицей, такой, что $F_{i,j} = 0$ для $i \neq j$ и $F_{i,i} = f(i)$. Т.е. все элементы матрицы F за исключением диагональных $((i,\ i))$ равны нулю и равны значениям приспособленности строк i-го типа. Для пропорциональной селекции:

$$\vec{s}(t) = \frac{F\vec{p}(t)}{\sum_{j=0}^{2^l-1} F_{jj}\vec{p}_j(t)}. \tag{4.8}$$

Данная формула есть не что иное, как определение пропорциональной селекции. Таким образом, зная $\vec{p}(t)$ и F можно легко найти $\vec{s}(t)$ и наоборот. Вос и Липинс представляют большинство результатов через $\vec{s}(t)$.

---

[1] Нужно отметить, что изначально в результате скрещивания для выживания случайно выбирался один потомок. Однако такой подход, как отмечено, например, в работе: Booker L.B. Intelligent behavior as an adaptation to the task environment. Unpublished PhD thesis. University of Michigan, 1982 – ведет к потере аллелей (allele loss). Поэтому было предложено оставлять обоих потомков, что к настоящему времени стало практически стандартом (хотя, конечно, есть и довольно многочисленные исключения). О потере аллелей см. работу: De Jong K. An analysis of the behavior of a class of genetic adaptive systems. Unpublished PhD thesis. University of Michigan, 1975 (also university microfilms no. 76-9381) – Прим. перев.

[2] Известное также как тип строки. Далее записи «строка i» и «строка i-го типа» будут обозначать одно и то же. – Прим. перев.

Используя данные определения, Вос и Липинс пытаются определить единственный «оператор» G такой, что применение G к $\vec{s}(t)$ будет точно соответствовать эффекту от запуска ГА в поколении t для формирования популяции в поколении t+1:

$$\vec{s}(t+1) = G\vec{s}(t). \qquad (4.9)$$

Тогда итерационное применение оператора G, начиная с $\vec{s}(0)$, даст точное описание работы ГА (что весьма похоже на модели, разработанные в популяционной генетике Фишером и другими исследователями, см., например, [Ewens, 1979]).

Чтобы прояснить картину, предположим, что ГА использует только селекцию (без кроссовера и мутации). Пусть $E(x)$ обозначает ожидаемую долю строки x. Тогда, поскольку $s_i(t)$ равно вероятности выбора строки i-го типа на каждом шаге селекции, то

$$E(\vec{p}(t+1)) = \vec{s}(t).$$

Пусть запись $\vec{x} \sim \vec{y}$ обозначает пропорциональность. Тогда из уравнения (4.8) имеем

$$\vec{s}(t+1) \sim F\vec{p}(t+1),$$

что ведет к следующему соотношению

$$E(\vec{s}(t+1)) \sim F\vec{s}(t).$$

Получаем искомое отношение (аналогичное виду уравнения (4.9)) из которого следует, что $G = F$ в случае использования только селекции.

Данные результаты дают только ожидаемые значения. Для любой популяции конечного размера будет характерно отклонение полученных величин от ожидаемых значений. В случае бесконечно большой популяции ожидаемые значения будут точными.

Вос и Липинс учли в модели кроссовер и мутацию, определив G как композицию матрицы приспособленности F и «оператора рекомбинации» G[1], который соответствует эффекту от применения кроссовера и мутации. (Вос и Липинс используют термин «рекомбинация», чтобы объединить кроссовер и мутацию. Я буду применять такую трактовку до конца раздела.) Один из способов определить G – найти $r_{i,j}(k)$ – вероятность получить k-ю строку в результате рекомбинации строк i и j. Если $r_{i,j}(k)$ известно, то мы можем вычислить

$$E(p_k(t+1)) = \sum_{i,j} s_i(t) s_j(t) r_{i,j}(k).$$

Другими словами, ожидаемая доля строки k в поколении t+1 равна вероятности ее получения для данной пары родителей, помноженная на вероятность их выбора и просуммированная по всем возможным парам родителей.

---

[1] В оригинале используется другой символ, присутствующий в формате LaTeX и похожий на рукописную «*M*», но поскольку у меня нет уверенности, что: (а) я смогу воспроизвести этот символ, перебрав доступные шрифты; (б) на любом другом компьютере будет установлен этот шрифт для нормального отображения текста, то было решено заменить его на греческую G. – Прим. перев.

Определить $r_{i,j}(k)$ и G довольно непросто. Вос и Липинс сначала определяют более простую матрицу M, элементы которой $M_{i,j}$ равны вероятности $r_{i,j}(0)$, т.е. что рекомбинация строк i и j даст строку 0 (т.е. строку, содержащую одни нули). Я детально опишу данный прием, чтобы читатели, мало знакомые с теорией вероятностей, смогли понять, как все работает. (Другие читатели могут попытаться сделать необходимые выкладки самостоятельно, прежде чем прочитают их ниже). После того как элементы $r_{i,j}(0)$ определены, они могут быть использованы для рассмотрения общего случая.

Выражение для $r_{i,j}(0)$ равно сумме двух произведений: вероятности, что кроссовер между строками i и j не случится, а выбранный потомок (строка i или j) мутирует до нулевой строки (первое произведение) и вероятности, что кроссовер случится и выбранный потомок мутирует до строки из одних нулей.

Если строки i и j выбраны для скрещивания, то вероятность применения кроссовера равна p_c, а вероятность обратного события равна $1 - p_c$. Аналогично, вероятность мутации каждого бита выбранного потомка равна p_m, а вероятность, что мутация не произойдет, равна $1 - p_m$. Если $|i|$ – количество единичных разрядов в строке i длины l, то вероятность, что строка i мутирует до нулевой, есть вероятность, что все $|i|$ единиц в строке мутируют, а ни один из $(l-|i|)$ нулевых разрядов не изменится:

$$p_m^{|i|}(1-p_m)^{l-|i|}.$$

Таким образом, первый множитель в выражении для $r_{i,j}(0)$ равен:

$$\frac{1}{2}(1-p_c)\left[p_m^{|i|}(1-p_m)^{l-|i|} - p_m^{|j|}(1-p_m)^{l-|j|}\right]$$

Т.к. в рассматриваемой модели выживает только один потомок, то множитель ½ отражает равную вероятность выживания каждого из потомков.

Для определения второго множителя, будем считать, что h и k обозначают потомков, полученных в результате применения кроссовера с разрывом в точке c (считая с правого края строки, см. рис. 4.3). Заметим, что всего существует $l - 1$ возможная точка разрыва, поэтому вероятность выбора точки c равна $1/(l - 1)$.

$c = 5$

i_1 i_2

i: 0 0 0 | 1 0 1 0 1 $\rightarrow$ h: 0 0 0 0 1 0 1 0

j_1 j_2

j: 1 1 1 | 0 1 0 1 0 $\qquad$ k: 1 1 1 1 0 1 0 1

Рис. 4.3. Иллюстративный пример для c, i_1, i_2, j_1 и j_2.

Второй множитель может быть записан следующим образом

$$\frac{1}{2}\frac{p_c}{l-1}\sum_{c=1}^{l-1}\left[p_m^{|h|}(1-p_m)^{l-|h|}-p_m^{|k|}(1-p_m)^{l-|k|}\right].$$

Множитель ½ также обозначает равновероятное выживание каждого потомка.

Для завершения нам необходимы выражения для $|h|$ и $|k|$. Пусть i_1 подстрока строки i, содержащая $l-c$ разрядов слева от точки c, и пусть i_2 подстрока строки i, содержащая c разрядов справа от точки c. Аналогично, для строки j определим подстроки j_1 и j_2, как показано на рис. 4.3. Тогда $|h|=|i|-|i_2|+|j_2|$ и $|k|=|j|-|j_2|+|i_2|$. Вос и Липинс упрощают эти выражения, используя удобный прием. Прежде всего, заметим, что

$$|i_2|=|(2^c-1)\wedge i|,$$

где $\wedge$ обозначает логическое «И». Т.к. 2^c-1 представляет строку, в которой за $l-c$ нулями следуют c единиц, то $|(2^c-1)\wedge i|$ дает количество единичных разрядов для c бит в правой части строки i. Точно также:

$$|j_2|=|(2^c-1)\wedge j|.$$

Пусть

$$\Delta_{i,j,c}=|i_2|-|j_2|=|(2^c-1)\wedge i|-|(2^c-1)\wedge j|.$$

Тогда $|h|=|i|-\Delta_{i,j,c}$ и $|k|=|j|+\Delta_{i,j,c}$.

Теперь мы можем записать окончательное выражение для $r_{i,j}(0)$. Для упрощения записи, введем обозначение $h=p_m/(1-p_m)$. Тогда после ряда выкладок получим

$$r_{i,j}(0)=\frac{(1-p_m)^l}{2}\left[h^{|i|}\left(1-p_c+\frac{p_c}{l-1}\sum_{c=1}^{l-1}h^{-\Delta_{i,j,c}}\right)+h^{|j|}\left(1-p_c+\frac{p_c}{l-1}\sum_{c=1}^{l-1}h^{\Delta_{i,j,c}}\right)\right]. \qquad (4.10)$$

Данное выражение отражает суть того, каким образом проведен анализ. Рационально используя логические операторы и перестановки (эта тема выходит за рамки данного обсуждения), Вос и Липинс смогли выразить общий оператор рекомбинации G через M (см. работу [Vose, Liepins, 1991]).

Пусть $G(\vec{x})=F\circ\mathbf{\Gamma}(\vec{x})$ для векторов $\vec{x}$, где $\circ$ – оператор композиции. Тогда в пределе для бесконечно большой популяции

$$G(\vec{s}(t))\sim\vec{s}(t+1).$$

Определим G_p как

$$G_p(\vec{x})=\mathbf{\Gamma}(F\vec{x}/|F\vec{x}|),$$

где $G_p(\vec{p}(t)) = \vec{p}(t+1)$ обозначает сумму компонент вектора $\vec{s}$. Тогда в пределе для популяции $\vec{s}(0)$ бесконечно большого размера операторы G и G_p хотя и влияют на различные характеристики популяции, но могут быть выражены друг через друга путем простых преобразований.

Результаты формализации

Каким образом подобная формализация может помочь нам лучше понять и предсказать поведение ГА? Рассматривая *G* как оператор, задающий динамическую систему, Вос и Липинс сформировали геометрическую картину поведения ГА и затем использовали ее чтобы доказать некоторые свойства этого поведения. Геометрия картины в том, что набор всех возможных векторов $\vec{s}(t)$ образует поверхность *S*, на которой оператор *G* определяет перемещение от точки к точке. Исходной точкой является $\vec{s}(0)$, а итерационное применение *G*, начиная с этой точки, задает траекторию на *S*. Анализируя динамику *G*, прежде всего, необходимо определить точки равновесия для *G* на *S*, т.е. такой набор $\vec{s}(t)$ для которого $G(\vec{s}(t)) = \vec{s}(t)$. Другими словами, мы хотим узнать, какие точки $\vec{s}(t)$ обладают тем свойством, что если ГА окажется в них, то он «не уйдет» оттуда.

Эта задача в ее общей постановке не была решена Восом и Липинсом (1991). Они решили отдельные проблемы нахождения неподвижных точек для F и G и проанализировали их свойства. Несложно показать, что точки равновесия для F (для одной только селекции) соответствуют популяциям, в которых строки сошлись к строкам с равной приспособленностью. Вос и Липинс доказали, что стабильным является только один класс таких точек, включающий точки, соответствующие популяциям строк с максимальной приспособленностью. Другими словами, если популяция сходится к состоянию, в котором не все строки имеют максимальную приспособленность, то небольшое изменение в распределении приспособленности в популяции может привести к «уходу» из текущей фиксированной точки. Но если приспособленность по всей популяции максимальна, то в случае любого достаточно малого изменения в распределении приспособленности ГА снова окажется в этой фиксированной точке.

Вос и Липинс также показали, что в случае применения только оператора G над $\vec{s}(t)$ существует единственная равновесная точка: вектор $\vec{s}(t)$, компоненты которого соответствуют популяции, содержащей с равными вероятностями все строки из пространства поиска. Аналогично, применяя оператор G над $\vec{P}$, получается только одна точка равновесия, в которой все возможные строки присутствуют в популяции с равными долями. Это значит, что в пределе для популяции бесконечного размера использование без селекции только кроссовера и мутации приводит к максимально «перемешанной» популяции, содержащей одинаковое количество всех возможных строк [1].

Вос и Липинс оставили неразрешенной более сложную проблему, касающуюся понимания совокупного эффекта от одновременного применения F и G для прояснения взаимного влияния кроссовера и мутации. Однако, они предположили, что предложенный формализм может пролить свет на феномен «прерывистого равновесия» («punctuated equilibria»), часто наблюдаемого в работе генетических алгоритмов и характеризующегося относительно долгими периодами отсутствия улучшений чередующимися (прерывающимися) с быстрыми повышениями приспособленности. Их идея заключается в

[1] Данное состояние в популяционной генетике известно как равновесие Роббинса (Robbin's equlibrium). См., к примеру, работу: Spears W. The equilibrium and transient behavior of mutation and recombination // Foundations of Genetic Algorithms, № 6, 2001, pp. 241–260. – Прим. перев.

том, что подобное прерывистое равновесие возникает в результате комбинирования «сжимающих» (focusing) свойств F и «рассеивающих» (diffusing) свойств G. В этом случае периоды «застоя» относятся к периодам, когда популяция находится в окрестности точки неустойчивого равновесия, а периоды быстрого улучшения соответствуют периодам движения популяции (в результате действия рассеивающих свойств рекомбинации) от окрестности одной точки равновесия к другой. (отметим, что даже если данная точка равновесия неустойчива, динамическая система может достаточно долго находиться в ее окрестности). Эти эффекты подлежат дальнейшему более подробному количественному обоснованию в рамках предложенной или любой другой модели. (понятие сжимающих и рассеивающих сил, одновременно влияющих на динамику ГА, обсуждалось, хотя и с меньшими подробностями, в 6-й главе книги Холланда 1975 г.)

Формализация Воса и Липинса является превосходным первым шагом к более детальному пониманию и более строгому прогнозированию простого ГА. (Развитие модели Воса и Липинса см. в работе [Whitley, 1993a]). Нужно отметить, что главный недостаток представленной формализации и полученных результатов заключен в допущении бесконечного размера популяции, что выражается в оперировании ожидаемыми величинами. Рассмотрение бесконечной популяции является идеализацией, упрощающей анализ, однако, динамика конечных популяций может значительно отличаться из-за присутствия выбросов и случайных отклонений (sampling error).

Модель для популяции конечного размера

Случай бесконечно большой популяции рассматривает детерминированные переходы от $\vec{p}(t)$ к $\vec{p}(t+1)$ и, следовательно, от $\vec{s}(t)$ к $\vec{s}(t+1)$, т.к. в бесконечной популяции отсутствуют случайности. При моделировании же конечной популяции, напротив, необходимо принимать во внимание появляющиеся стохастические эффекты.

Чтобы рассмотреть случай конечной популяции Никс и Вос [Nix, Vose, 1991] моделировали простой ГА с использованием аппарата марковских цепей. Цепи Маркова описывают стохастические процессы, в которых вероятность, что процесс в момент времени t находится в состоянии j , зависит только от состояния i процесса в момент времени (t-1). Многие природные процессы могут быть описаны посредством марковских цепей. (В качестве введения в математические основы цепей Маркова см. [Feller, 1968][1]).

«Состоянием» ГА с конечной популяцией будет просто произвольная популяция. Набор всех возможных состояний соответствует набору возможных популяций размера n. Их можно пронумеровать в некотором порядке и проиндексировать через i. Никс и Вос представляют i-ю популяцию в виде вектора $\vec{f}_i$ длины 2^l. Элемент с индексом y вектора $\vec{f}_i$ равен числу строк y в популяции P_i. Ясно, что для простого ГА текущая популяция P_j зависит (с некоторой вероятностью) только от популяции предыдущего поколения. Поэтому ГА можно моделировать с использованием цепей Маркова.

Чтобы построить такую модель, необходимо найти вероятность перехода от произвольной данной популяции к любой другой популяции в ходе работы простого ГА. Множество всех возможных популяций размера n можно представить в виде огромной матрицы Z, в которой столбцы соответствуют всем возможным векторам популяций $\vec{f}_i$. Сколько всего существует таких популяций? Ответом является

---

[1] На русском языке см., например, Вентцель Е.С., Овчаров Л.А. Теория случайных процессов и ее инженерные приложения. – Учеб. пособие для втузов. – 2-е изд. стер. – М.: Высш. шк., 2000. – Прим. перев.

$$N = \binom{n+2^l-1}{2^l-1}$$

(Получение данной формулы является упражнением). Элемент матрицы $Z_{y,i}$ равен количеству вхождений строки y в популяцию i.

Рассмотрим простой пример составления матрицы Z для l = 2 и n = 2. В этом случае возможные популяции:

$$p_0 = \begin{Bmatrix}00\\00\end{Bmatrix}, \quad p_1 = \begin{Bmatrix}00\\01\end{Bmatrix}, \quad p_2 = \begin{Bmatrix}00\\10\end{Bmatrix}, \quad p_3 = \begin{Bmatrix}00\\11\end{Bmatrix}, \quad p_4 = \begin{Bmatrix}01\\01\end{Bmatrix},$$
$$p_5 = \begin{Bmatrix}01\\10\end{Bmatrix}, \quad p_6 = \begin{Bmatrix}01\\11\end{Bmatrix}, \quad p_7 = \begin{Bmatrix}10\\10\end{Bmatrix}, \quad p_8 = \begin{Bmatrix}10\\11\end{Bmatrix}, \quad p_9 = \begin{Bmatrix}11\\11\end{Bmatrix}.$$

Матрица Z:

$$Z = \begin{Bmatrix} 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 2 \end{Bmatrix}.$$

Состояние в марковской цепи соответствуют столбцу матрицы Z.

Следующим шагом будет определение матрицы переходов Q. Матрица Q имеет размерность $N \times N$, и каждый ее элемент $Q_{i,j}$ равен вероятности получения популяции P_j из популяции P_i в результате работы простого ГА. Когда матрица задана, она может быть использована для выявления некоторых свойств ГА.

Определение вероятностей перехода $Q_{i,j}$ несколько затруднено, но поучительно. Пусть $p_i(y)$ – вероятность, что строка y будет получена в результате процесса селекции и рекомбинации (т.е. после шагов 3 и 4) простого ГА, работающего с популяцией P_i. Число вхождений строки y в популяцию P_j равно $Z_{y,j}$, поэтому вероятность, что из популяции P_i будет получено точное число[1] строк y, есть вероятность получить $Z_{y,j}$ экземпляров y из популяции P_i. Она равна вероятности, что строка y получена в результате комбинации из $Z_{y,j}$ различных щагов (селекция + рекомбинация), а число таких комбинаций равно количеству способов выбрать $Z_{y,j}$ шагов (селекция + рекомбинация) из n возможных шагов (селекция + рекомбинация). Вслед за Никсом и Восом мы будем рассматривать по порядку все возможные строки.

Количество способов выбрать $Z_{0,j}$ экземпляров строки 0, чтобы заполнить $Z_{0,j}$ позиций в популяции j равно

[1] Т.е., что не проявится влияние конечного размера популяции. – Прим. перев.

$$\binom{n}{Z_{0,j}}.$$

После выбора $Z_{0,j}$ экземпляров строки 0 в новой популяции незаполненными остаются $n-Z_{0,j}$ позиций. Количество способов разместить $Z_{1,j}$ строк 1 по $n-Z_{0,j}$ позициям равно

$$\binom{n-Z_{0,j}}{Z_{1,j}}.$$

Продолжая подобным образом, мы получим выражение для числа всех возможных способов формирования популяции P_j в результате n шагов (селекция + рекомбинация):

$$\binom{n}{Z_{0,j}}\binom{n-Z_{0,j}}{Z_{1,j}}\ldots\binom{n-Z_{0,j}-Z_{1,j}-\ldots-Z_{2^l-2,j}}{Z_{2^l-1,j}}=\frac{n!}{Z_{0,j}!\,Z_{1,j}!\ldots Z_{2^l-1,j}!}$$

Чтобы записать это выражение, мы пронумеровали все строки числами от 0 до 2^l-1. Несложно показать, что вычисления с использованием другого порядка нумерации строк даст такой же результат.

Вероятность получить (из популяции P_i) точное число экземпляров для каждой строки y (в популяции P_j) равно

$$\prod_{y=0}^{2^l-1}[p_i(y)]^{Z_{y,j}}.$$

Вероятность, что популяция P_j сформирована из популяции P_i, есть произведение двух предыдущих выражений (соответствующее полиномиальному распределению (multinomial distribution)):

$$Q_{i,j}=\frac{n!}{Z_{0,j}!\,Z_{1,j}!\ldots Z_{2^l-1,j}!}\prod_{y=0}^{2^l-1}[p_i(y)]^{Z_{y,j}}=n!\prod_{y=0}^{2^l-1}\frac{[p_i(y)]^{Z_{y,j}}}{Z_{y,j}!}$$

Единственное, что остается сделать – это вывести выражение для $p_i(y)$, обозначающего вероятность получения строки y в результате однократного шага (селекция + рекомбинация), примененного к популяции P_i. Для этого мы можем использовать определенные выше матрицы F и $\mathbf{\Gamma}$. Тогда $p_i(y)$ просто равно ожидаемой доле строк y в рассматриваемой популяции, полученной из P_i в результате работы простого ГА. Доля строк y в популяции P_i равна $\left(\vec{f}_i/\left|\vec{f}_i\right|\right)_y$, где $\left|\vec{v}_i\right|$ обозначает суммирование по всем компонентам вектора $\vec{v}_i$, а $(A)_y$ – y-ю компоненту вектора $\vec{v}_i$. Вероятность выбора на каждом шаге селекции строки y равна:

$$\left(\frac{F\vec{f}_i}{\left|F\vec{f}_i\right|}\right)_y,$$

а ожидаемая доля строк y в следующем поколении:

$$\left[\mathbf{\Gamma} \left(\frac{F\vec{f}_i}{\left|F\vec{f}_i\right|} \right) \right]_y .$$

Поскольку $p_i(y)$ соответствует ожидаемой доле строк y в популяции следующего поколения, мы можем записать окончательное выражение для $Q_{i,j}$:

$$Q_{i,j} = n! \prod_{y=0}^{2^l-1} \frac{\left[\mathbf{\Gamma} \left(F\vec{f}_i / \left|F\vec{f}_i\right| \right)_y \right]^{Z_{y,j}}}{Z_{y,j}!} .$$

Матрица $Q_{i,j}$ задает точную модель простого ГА, работающего с конечной популяцией.

Никс и Вос использовали теорию цепей Маркова, чтобы получить множество результатов, касающихся данной модели. Они показали, к примеру, что при $n \to \infty$, траектория марковской цепи сходится к итерациям G (либо G_p) с вероятностью сколь угодно близкой к 1. Это означает, что для очень больших значений n модель конечной популяции ведет себя подобно модели для ГА с бесконечной популяцией. Они также продемонстрировали, что если G_p имеет единственную устойчивую точку, то при $n \to \infty$ ГА асимптотически бесконечно долго находится в этой точке. В случае, если G_p имеет более одной устойчивой точки, то при $n \to \infty$ время, когда ГА «отсутствует» в любой из этих точек, стремится к 0. Для детального доказательства данных утверждений см. [Nix, Vose, 1991].

В работе [Vose, 1993] Вос расширил описанные модели как для конечной, так и для бесконечной популяции. Он представил геометрическую интерпретацию этим моделям, определив «поверхность ГА», на которой располагаются траектории популяционной динамики. Здесь я не буду приводить подробностей этих расширений, кроме главного результата, в котором показано, что при $n \to \infty$ доля времени, которое ГА проводит в окрестности точек неустойчивого равновесия, асимптотически стремится к 0, а время, когда ГА близок к устойчивым фиксированным точкам, асимптотически приближается к 1. В терминах теории динамических систем это означает, что ГА асимптотически сходится к точкам устойчивого равновесия, имеющих наибольшие области притяжения. В случае $n \to \infty$ вероятность, что ГА будет «находиться» где-либо еще стремится к 0. Следствием данного результата Воса является то, что кратковременное поведение ГА обуславливается начальной популяцией, когда определяется устойчивая точка, к которой ГА приближается в начале работы. Но долговременная динамика ГА зависит только от структуры поверхности ГА, которая определяет, какая устойчивая точка имеет наибольшую область притяжения.

Каким образом можно использовать представленные формальные модели? Поскольку они являются наиболее подробными моделями ГА, они, в принципе, могут быть использованы, чтобы предсказать любой аспект поведения ГА. Однако, на практике данные модели не могут использоваться для подробного исследования и прогнозирования динамики ГА в силу все той же своей детальности – используемые в моделях матрицы становятся слишком большими. Например, даже для очень «скромного» ГА при, допустим, $l = 8$ и $n = 8$, определенная Никсом и Восом марковская матрица переходов Q

имеет более 10^{29} элементов, и их количество очень быстро растет с увеличением значений l и n. В силу этого обстоятельства подробное моделирование ГА с использованием матриц подобного размера попросту невозможно.

Это не значит, что такие модели бесполезны. Как мы убедились ранее, есть некоторые общие свойства динамики ГА, которые могут быть продемонстрированы с использованием данных моделей, такие как наличие на «поверхности ГА» точек устойчивого равновесия или свойство асимптотической динамики ГА, зависящей от этих устойчивых точек. Такие свойства позволяют проанализировать поведение ГА, пусть и с некоторыми ограничениями. Многие из этих свойств остаются на уровне предположений, догадок и обсуждаются Восом и его коллегами; также до сих отсутствует детальное понимание характеристик поверхности ГА при совместном использовании F и $\mathbf{\Gamma}$. Попытка уяснить эти характеристики является достойным (и все еще открытым) направлением исследований.

СПИСОК ИСТОЧНИКОВ

[**Davis, Principe, 1991**] Davis T. E., Principe J. C. A simulated annealing−like convergence theory for the simple genetic algorithm // In Belew R.K., Booker L.B. (eds.): Proceedings of the Fourth International Conference on Genetic Algorithms. Morgan Kaufmann, 1991.

[**Ewens, 1979**] Ewens W.J. Mathematical Population Genetics. Springer−Verlag, 1979.

[**Feller, 1968**] Feller W. An introduction to probability theory and its application. Vol. 1. Third edn. Wiley, 1968.

[**Goldberg, 1987**] Goldberg D. E. Simple genetic algorithms and the minimal deceptive problem // In Davis L.D. (ed.): Genetic Algorithms and Simulated Annealing. Morgan Kaufmann, 1987.

[**Goldberg, Segrest, 1987**] Goldberg, D. E., and Segrest, P. 1987. Finite Markov chain analysis of genetic algorithms // In Grefenstette J.J. (ed.): Genetic Algorithms and Their Applications: Proceedings of the Second International Conference on Genetic Algorithms. Erlbaum, 1987.

[**Horn, 1993**] Horn J. Finite Markov chain analysis of genetic algorithms with niching // In S. Forrest (ed.): Proceedings of the Fifth International Conference on Genetic Algorithms. Morgan Kaufmann, 1993.

[**Nix, Vose, 1991**] Nix A.E., Vose M.D. Modeling genetic algorithms with Markov chains // Annals of Mathematics and Artificial Intelligence, 1991, no. 5, pp. 79–88.

[**Vose, 1993**] Vose M.D. Modeling simple genetic algorithms // In Whitley L.D. (ed): Foundations of Genetic Algorithms 2. Morgan Kaufmann, 1993.

[**Vose, Liepins, 1991**] Vose M. D., Liepins G. E. Punctuated equilibria in genetic search // Complex Systems, 1991, no. 5, pp. 31–44.

[**Whitley, 1993a**] Whitley L. D. An executable model of a simple genetic algorithm // In Whitley L.D. (ed): Foundations of Genetic Algorithms 2. Morgan Kaufmann, 1993.