

3. ВЗАИМОДЕЙСТВИЕ ОСНОВНЫХ ФАКТОРОВ ЭВОЛЮЦИИ ПОПУЛЯЦИИ В ТЕЧЕНИЕ ЖИЗНЕННОГО ЦИКЛА

3.1. Размножение особей, поддерживающее наследственную преемственность “потомками” признаков “родителей”

Будем считать, что популяция $P^t = (a_1^t, \dots, a_v^t)$ представляет собой *репродукционную группу* - совокупность из v особей, любые две из которых $a_k^t, a_l^t \in P^t, k \neq l$ могут размножаться, выступая в роли “родителей” (a_k^t - “мать”; a_l^t - “отец”). Здесь под *размножением* понимается свойство особей $a_k^t \in P^t$ воспроизводить одного или нескольких себе подобных непосредственных “потомков” (“детей”) $b_i^t, i \geq 1$ и обеспечивать у них непрерывность и наследственную преемственность качественных признаков “родителей”.

Таким образом, этот фактор эволюционного развития популяции приводит к получению новой генетической информации, содержащей различные комбинации аллельных форм генов “родительских” генотипов.

В терминах экстремальной задачи однокритериального выбора (1.3) “*воспроизводство себе подобных*” можно интерпретировать как возможность построения по заданным допустимым решениям $\bar{x}^k, \bar{x}^l \in D$ нового допустимого решения $\bar{x}^i \in D$, а “*непрерывность и наследственную преемственность*” - как возможность использования аллельных форм в виде бинарных комбинаций $e_0(\beta_i)$, содержащихся в генотипах “родителей” $E(\bar{x}^k)$ и $E(\bar{x}^l)$, для формирования генотипа $E(\bar{x}^i)$ “потомка”, тем самым обеспечивая передачу наследственных признаков особей от поколения к поколению на уровне обмена генами.

Рассмотрим механизм размножения двух “родительских” особей $a_k^t, a_l^t \in P^t$ путем *сигнамии* (оплодотворения) их *репродуктивных клеток* - “материнской” гаметы (яйцеклетки) $E(a_k^t)$ и “отцовской” гаметы (сперматозоида) $E(a_l^t)$, каждая из которых является *гаплоидом* (одинарным набором непарных хромосом $E(\bar{x}^k)$ и $E(\bar{x}^l)$, соответственно).

В процессе сигнамии образуется “родительская” *зигота* - оплодотворенная клетка, способная развиваться в новую особь с передачей наследственных признаков (генетической информации) от “родителей” их “потомкам”. Зигота, в отличие от гамет, является *диплоидом*, содержащим одну пару из двух неотличимых одна от другой

хромосом, которые происходят от “родительских” гамет: одна от “материнской” гаметы, а другая от “отцовской” гаметы. Такие хромосомы называются *гомологичными хромосомами*. В гомологичных хромосомах для всех признаков имеется по два гена, называемых *аллельными генами*. Аллельные гены принадлежат одному и тому же локусу. В этом смысле локус принадлежит уже не отдельной хромосоме, а совокупности из двух гомологичных хромосом. Каждый локус содержит не менее двух аллелей, которые могут быть как одинаковыми, так и различными. Необходимо заметить, что гены “родительских” гамет могут существовать более чем в двух аллельных формах, хотя каждая зигота может быть носителем только двух форм аллелей (А или а).

Зиготы, содержащие в аллельных генах гомологичных хромосом одинаковые аллели (АА или аа), называются *гомозиготами*, а содержащие разные аллели (Аа или аА), называются *гетерозиготами*. Очевидно, что введенные понятия “гомозигота” и “гетерозигота” определяются относительно конкретного локуса, содержащего аллельный ген.

В результате акта синапсии аллели “родительских” гамет могут меняться местами в аллельных генах гомологичной хромосомы, что позволяет рассматривать следующие ситуации образования зигот (рис. 3.1.):

- ген из “отцовской” хромосомы переходит в “материнскую” хромосому;
- ген из “материнской” хромосомы переходит в “отцовскую” хромосому;
- происходит взаимный обмен генами между “материнской” и “отцовской” хромосомами;
- “отцовская” и “материнская” хромосомы остаются без изменения.

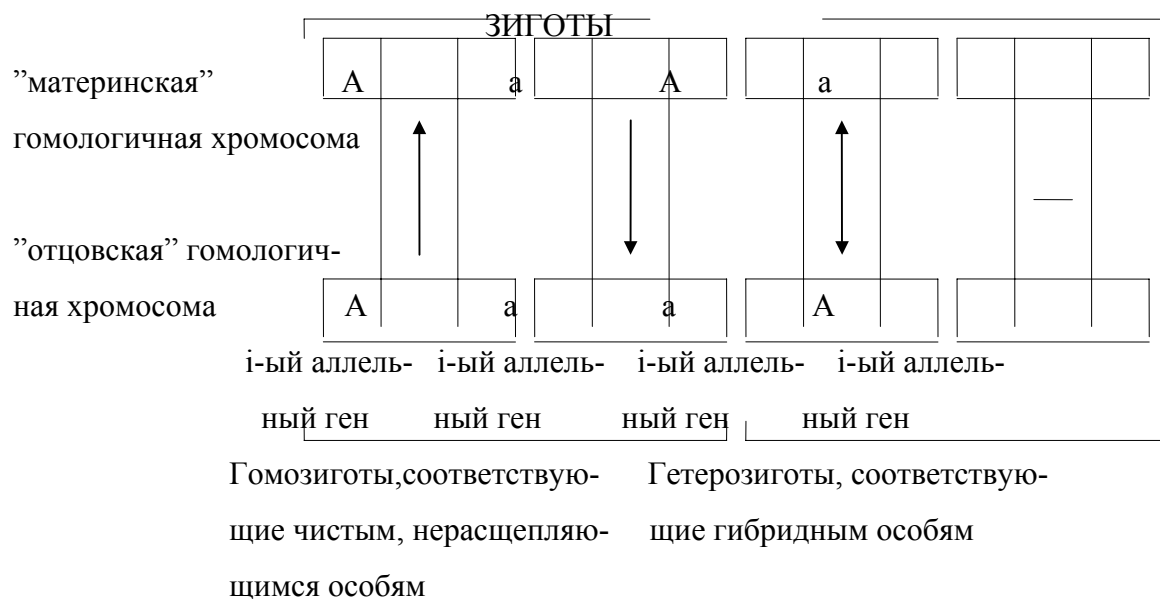


Рис. 3.1. Ситуации образования зигот (а, А - аллели, содержащиеся в i-ом локусе, соответственно, “материнской” и “отцовской” гамет).

Таким образом, при образовании зигот происходит независимое и случайное расхождение “родительских” генов по аллельным генам гомологичных хромосом зиготы независимо от того, у какой из “родительских” гамет они присутствовали до оплодотворения.

Заключительным этапом размножения особей является акт *мейоза* - процесс образования гамет из “родительской” зиготы путем независимого расхождения гомологичных хромосом по дочерним гаметам, воспроизводящим “потомство”. Одна диплоидная зигота может дать начало четырем гаплоидным гаметам (гамете, тождественно воспроизводящей “отцовскую” гамету; гамете, тождественно воспроизводящей “материнскую” гамету; гамете, являющейся “отцовской” гаметой, в которой в i -ом локусе находится аллель i -го гена из “материнской” гаметы; гамете, являющейся “материнской” гаметой, в которой в i -ом локусе находится аллель i -го гена из “отцовской” гаметы).

Процесс размножения двух особей должен удовлетворять следующим законам наследственности Менделя [4].

1. *Первому закону Менделя (закону расщепления) о наследовании альтернативных проявлений одного и того же признака, который формулируется следующим образом:*

“Два гена, определяющие тот или иной признак, не сливаются и не растворяются один в другом, но остаются независимыми друг от друга, расщепляясь при формировании гамет”.

Согласно этому закону гены (или соответствующие им признаки “родителей”), имеющие одинаковые аллели $[e_0^{OT}(\beta_i) = e_0^M(\beta_i)]$, сохраняют свои значения в “потомстве”, т.е. передаются с вероятностью, равной 1, “потомку” по наследству. Гены “родителей”, имеющие разные аллели $[e_0^{OT}(\beta_i) \neq e_0^M(\beta_i)]$, передаются “потомку” по наследству с вероятностью, равной 0,5, т.е. половина гамет оказывается носителем аллели $e_0^{OT}(\beta_i)$, а другая половина - аллели $e_0^M(\beta_i)$.

2. *Второму закону Менделя (закону независимого расщепления) о независимости комбинирования признаков, который формулируется следующим образом:*

“Родительские гены, определяющие различные признаки, наследуются независимо друг от друга”.

Согласно этому закону рекомбинация (обмен) генов в акте синапсии может происходить либо в каком-то одном аллельном гене, либо в нескольких аллельных генах одновременно, т.е. передача аллелей от “родителей” “потомству” может происходить в каждом аллельном гене независимо друг от друга. При этом может оказаться, что гаметы

“потомков” либо совпадают с “родительскими” гаметтами, либо отличаются от них в одном или нескольких локусах.

Подробно вопросы реализации процесса размножения особей будут рассмотрены в разделе 5.

3.2. Приобретение особями новых качественных признаков

В результате размножения воспроизводятся “потомки”, обладающие свойством преемственности наследственных признаков (генов) “родителей”. При этом генотипы “потомков”, как правило, содержат новые сочетания аллельных форм генов “родителей”, ведущие к новым количественным признакам “потомков” (фенотипу и степени приспособленности). Однако, генетическая информация, содержащаяся в хромосомном наборе “родителей” и “потомков”, не меняется, т.к. в результате размножения особей путем синамизии и мейоза частоты аллелей остаются постоянными, а меняются только частоты генотипов. Источником генетической изменчивости особей являются *мутации* - изменения качественных признаков особей в результате появления новых аллельных форм в отдельных генах или целиком во всей хромосоме. Тем самым в каждом поколении мутации поставляют в хромосомный набор популяции множество различных генетических вариаций, присущих особям, которых в дальнейшем будем называть *мутантами* $m_k^t, k \geq 1$.

Процесс изменения содержания генов в хромосоме особей путем мутаций называется *мутагенезом*. По сути дела, этот фактор эволюции популяции является источником новой генетической информации, не содержащейся ранее в генах генотипов “родителей” и их “потомков”.

Мутации являются случайными в том смысле, что не зависят ни от генетического кода особи, содержащегося в ее генотипе, ни от количественных значений фенотипа и степени приспособленности особи. Они происходят спонтанно с определенными вероятностями, заменяя в одном или нескольких локусах тех или иных генов аллельные формы последних новыми значениями аллелей, которые принадлежат генофонду и отличаются от аллелей всех “родительских” генотипов в том же самом локусе (гене).

Мутации происходят независимо от того, приносят ли они особи вред или пользу. Они не направлены на повышение или понижение степени приспособленности особи, а только производят структурные изменения в аллельных формах генов, меняя тем самым частоту аллелей по отдельным локусам в хромосомном наборе текущего поколения, что, в

свою очередь, приводит к изменению количественных признаков особи. В принципе, комбинация мутаций может привести к возникновению новых форм аллелей в некоторых генах генотипа мутанта, которые обеспечивают увеличение его степени приспособленности к внешней среде.

Эволюция популяции в течение смены нескольких поколений в смысле изменения генетической наследственности представляет из себя процесс одновременного и постепенного изменения как частот, так и форм аллелей в различных локусах хромосомы. При этом аллели действуют на количественные признаки не изолированно друг от друга. Так, влияние того или иного аллеля на степень приспособленности особи зависит от присутствия или отсутствия в его генотипе других аллелей. Набор аллелей каждого локуса взаимно приспособлен (*коадаптирован*) с набором аллелей других локусов. Поэтому изменение частот аллелей в одном локусе влечет за собой изменение частот аллелей и в других локусах.

Наиболее простым видом мутаций является *точечная мутация*, связанная с изменением аллеля “родительского” гена в одном из θ бит генной информации (0 заменяется на 1 или 1 заменяется на 0).

Определим *интенсивность процесса мутагенеза* в t -м поколении как среднее число точечных мутаций $M_t(t)$, которые могут произойти в хромосомном наборе t -ой популяции P^t :

$$M_t(t) = v \times (n \times \theta) \times P_m, \quad (3.1)$$

где v - численность популяции P^t ;

$n \times \theta$ - длина хромосомы, равная числу битов в бинарной строке $E(a_k^t)$;

P_m - *вероятность точечной мутации*, определяемая как число возможных однобитовых изменений на 100 бит генетической информации.

Обычно вероятность точечной мутации в популяции очень мала ($P_m=0.01$ или $P_m=0.001$), что приводит к невысокому темпу возникновения мутаций. Например, при $v=10$, $n=12$, $\theta=32$ и $P_m=0.01$ получаем, что в среднем в каждом поколении будет приходить 38 точечных мутаций.

Подробно вопросы реализации процесса мутагенеза будут рассмотрены в разделе 6.

3.3. Естественный отбор наиболее приспособленных особей

Третьим фактором эволюции является *естественный отбор* - процесс, способствующий повышению степени приспособленности особей и предотвращающий разрушительные последствия, которые могут возникнуть в результате мутаций.

Этот процесс можно рассматривать с двух позиций.

Во-первых, как преобразующий фактор, осуществляющий формирование различных форм генотипов вследствие различной степени приспособленности особей, что приводит к образованию популяций с новыми варибельными признаками, характеризующими популяцию в целом.

Во-вторых, как механизм выживания особей, наиболее приспособленных к условиям внешней среды, и гибели других, имеющих низкую степень приспособленности, в ходе смены поколений. Здесь под *гибелью (элиминацией)* особей понимается удаление отдельных особей из популяции в последующих поколениях и, следовательно, отстранение их от участия в размножении и воспроизводстве “потомства”.

Естественный отбор складывается из процессов взаимодействия особей с внешней средой и внутренних процессов, влияющих на вероятность выживания особей. Имеет смысл говорить лишь об *индивидуальном естественном отборе*, основанном на соперничестве внутри популяции одних особей с другими за право образовать “родительскую” пару, что, в свою очередь, приводит к уменьшению “потомства” у менее приспособленных особей. Особи, обладающие более высокой степенью приспособленности, имеют и большую вероятность по сравнению с другими особями выжить и оставить “потомство”. Вследствие этого полезные признаки, определяемые генами таких особей, будут накапливаться и повторяться в ряде поколений, а вредные или менее полезные признаки будут вытесняться и элиминировать.

На генетическом уровне естественный отбор можно рассматривать как процесс, посредством которого проводится дифференциация генетической информации и наиболее удачные (в смысле обеспечения высокой степени приспособленности) гены и генотипы передаются из поколения в поколение. Однако, генетические коды генотипов, возникающие в результате мутаций, передаются “потомкам” не с одинаковой вероятностью, ибо частоты некоторых из них могут сильно отличаться друг от друга. Если генотипы мутантов не поддерживаются в популяции со значительной частотой, то это означает, что они не приносят пользы своим обладателям. В то же время, если в результате мутаций появляются какие-то удачные аллели, то их частота под действием естественного отбора будет постепенно увеличиваться.

Окончательным результатом естественного отбора может быть либо полная элиминация того или иного аллеля, либо возникновение устойчивого полиморфизма, когда в хромосомном наборе популяции одновременно присутствует одна, две или более аллелоформ в одном локусе. Чем больше число изменчивых локусов и чем большим набором аллелей представлен такой локус, тем выше вероятность изменения частоты одних аллелей за счет других. Для этого разумеется необходимо, чтобы происходил естественный отбор, благоприятствующий увеличению степени приспособленности особей, составляющих популяцию последующего поколения, и чтобы генетической изменчивости были подвержены именно те гены, которые оказывают существенное влияние на степень приспособленности.

В заключение сформулируем *фундаментальную теорему естественного отбора (теорему Р. Фишера)* [4]:

“Скорость возрастания степени приспособленности популяции в любом t -ом поколении равна генетической дисперсии степени приспособленности в том же поколении.”

$$S^2(t) = \frac{1}{(v-1)} \sum_{i=1}^v [\mu(a_i^t) - \mu_{cp}(t)]^2, \quad (3.2)$$

где $\mu_{cp}(t) = \frac{1}{v} \sum_{i=1}^v \mu(a_i^t)$.

Здесь *генетическая дисперсия* является дисперсией распределения особей $a_k^t \in P^t$ по степени приспособленности $\mu(a_i^t)$ в составе одной и той же популяции P^t . Численное значение генетической дисперсии обуславливается различиями в генотипах отдельных особей.

Подробно вопросы реализации процесса естественного отбора будут рассмотрены в разделе 7.

3.4. Базовая структура генетического алгоритма, моделирующего эволюционное развитие популяции

Обобщая вышесказанное, цель эволюции первоначально заданной популяции $P^0 = (a_1^0, \dots, a_v^0)$ в течение жизненного цикла T , можно сформулировать следующим образом.

Отношения между особями и внешней средой, приводящие к избирательной элиминации (“гибели”) менее приспособленных и выживанию более приспособленных

особей, должны быть построены таким образом, чтобы в течение смены поколений в хромосомном наборе популяции накапливались такие новые качественные признаки (гены и генотипы), которые обеспечивают увеличение средней степени приспособленности особей по популяции в целом:

$$\text{MAX}_{P^t} \left\{ \frac{1}{v} \sum_{i=1}^v \mu(a_i^t) \right\} \quad (3.3)$$

При этом генотипы особей $a_k^t \in P^t$, $t=0,1,2, \dots, T$, заданные с помощью бинарных строк $E(a_k^t)$, в процессе своего преобразования в результате факторов эволюционного развития популяции по Дарвину в каждом поколении должны обладать следующими свойствами:

- *наследственности*, которая закрепляет у “потомков” лучшие признаки, полученные от “родителей” в результате их размножения;
- *изменчивости*, которая служит основой образования новых признаков за счет изменения генетического состава популяции в результате мутаций;
- *соревновательности*, которая определяет направление генетических изменений в популяции в результате естественного отбора по степени приспособленности особей к условиям внешней среды.

В дальнейшем под *генетическим алгоритмом* будем понимать алгоритмический подход к решению экстремальных задач однокритериального выбора, основанный на моделировании основных факторов эволюционного развития популяции.

Большую роль в развитии генетических алгоритмов сыграли I.Holland [5], D.Goldberg [6] и L.Davis [7], которые заложили и развили теоретические основы генетического подхода к решению задач оптимизации. Не останавливаясь на обзоре этих работ, приведем обобщенную схему генетического алгоритма, структура которого является типичной для широкого круга публикаций по этому вопросу.

Базовая структура “Генетического алгоритма” :

1. *Формирование начальной популяции* P^0 из v особей (a_1^0, \dots, a_v^0) :

1.1. Генерация хромосомного набора из v бинарных строк $E(a_k^t)$, удовлетворяющих требованиям, предъявляемым к символьной модели исходной экстремальной задачи.

1.2. Преобразование бинарных строк $E(a_k^t)$ в соответствующие им векторы управляемых переменных $\bar{X}^k \in D$ и вычисление степени приспособленности $\mu(a_k^t)$ для каждой особи a_i^0 , обладающей генотипом

$E(\bar{X}^i)$.

1.3. Особи $a_k^0, k = \overline{1, v}$ образуют начальную популяцию P^t для поколения $t=0$.

2. *Воспроизводство “потомков” с наследственными признаками “родителей”:*

2.1. Выбор конкретной “родительской” пары $(a_k^t, a_l^t) \in P^t$

для участия в процессе размножения.

2.2. Выбор схемы размножения.

2.3. Построение по выбранной схеме из генотипов “родителей”

$E(a_k^t), E(a_l^t)$ генотипов их “потомков” $E(b_i^t), i \geq 1$, сохраняющих наследственные признаки “родителей”.

2.4. Преобразование бинарных строк $E(b_i^t)$ в соответствующие векторы управляемых переменных $\bar{X}^i \in D$ и вычисление степени приспособленности “потомков”, обладающих генотипами $E(\bar{X}^i)$.

2.5. Вычисления повторяются с шага 2.1. до тех пор, пока не будет воспроизведено заданное число “потомков”.

3. *Мутагенез, приводящий к генетическим изменениям “родительских” признаков:*

3.1. Выбор типа мутации.

3.2. Построение по генотипу $E(a_k^t)$ одной из особей $a_k^t \in P^t$ генотипа $E(m_i^t)$ особи-”мутанта” m_i^t с помощью конкретного типа мутации.

3.3. Преобразование бинарной строки $E(m_i^t)$ в соответствующий вектор управляемых переменных $\bar{X}^i \in D$ и вычисление степени приспособленности особи-”мутанта” m_i^t , обладающей генотипом $E(\bar{X}^i)$.

3.4. Вычисления повторяются с шага 3.1. до тех пор, пока не будет создано заданное число “мутантов”.

4. Естественный отбор:

4.1. Определение среди “родителей”, “потомков” и “мутантов” особей, образующих репродуктивную группу, которая примет участие в естественном отборе.

4.2. Выбор схемы естественного отбора.

4.3. Формирование по выбранной схеме хромосомного набора популяции следующего поколения $P^{t+1} = (a_1^{t+1}, \dots, a_v^{t+1})$ из особей, принадлежащих репродукционной группе.

5. Проверка условий окончания процесса эволюции популяции P .

Если условия окончания процесса эволюции не выполнены, то происходит смена поколений и все вычисления для популяции следующего $(t+1)$ - го поколения P^{t+1} повторяются с шага 2.

В качестве условий окончания процесса эволюции популяции может использоваться одно из следующих неравенств:

$$t > T \quad (3.4)$$

или

$$D_b(t) = 0. \quad (3.5)$$

Выполнение неравенства (3.4) означает, что эволюция популяции закончена в связи с тем, что она исчерпала свой жизненный цикл; окончание эволюции популяции при равенстве побитового разнообразия текущей популяции P^t нулю означает, что все генотипы в хромосомном наборе популяции P^t совпадают между собой.

В заключение данного раздела приведем отличия генетических алгоритмов от поисковых методов оптимизации [6].

1. Генетические алгоритмы осуществляют *прямое манипулирование бинарными строками* $E(\bar{X})$, с помощью которых закодированы в двоичном коде допустимые решения $\bar{X} \in D$, а не самими внутренними параметрами $x_i, i = \overline{1, n}$, заданными действительными числами.

2. Генетические алгоритмы в каждом t -ом поколении *оперируют одновременно со всей совокупностью из v допустимых решений* $\bar{X}^k \in D$, образующих популяцию P^t , с целью получения хромосомного набора популяции следующего поколения P^{t+1} .

Таким образом генетические алгоритмы на каждой итерации, совпадающей с текущим поколением, *позволяют определять v новых допустимых решений*, в то время как классические методы поиска [8] на каждой итерации определяют единственное новое

допустимое решение. Например, градиентный метод минимизации реализуется с помощью рекуррентного выражения:

$$\bar{x}^{k+1} = \bar{k}^k - \lambda \nabla Q(\bar{k}^k), k=0,1,2,\dots \quad (3.6)$$

где \bar{x}^0 - начальное приближение;

$\nabla Q(\bar{k}^k)$ - градиент минимизируемой функции;

λ - шаг вдоль градиента.

3. Генетические алгоритмы основаны на вероятностных схемах преобразования бинарных строк $E(\bar{X})$, составляющих хромосомный набор популяции P^t , которые *моделируют биологические механизмы популяционной генетики* [4].

4. Генетические алгоритмы - это *методы нулевого порядка*, стратегия поиска, в которых построена только на вычислении значений критерия оптимальности Q и не требует знания дополнительной информации о производных, константе Липшица и т.д., что характерно для градиентных и квази-ньютоновских методов [8].

5. Генетические алгоритмы являются *робастными методами по отношению к виду минимизируемой функции*, т.к. при их применении не требуется, чтобы критерий оптимальности был непрерывным, дифференцируемым, унимодальным и т.д. Они осуществляют поиск оптимального решения по одной и той же стратегии как для унимодальных, так и для многоэкстремальных функций.