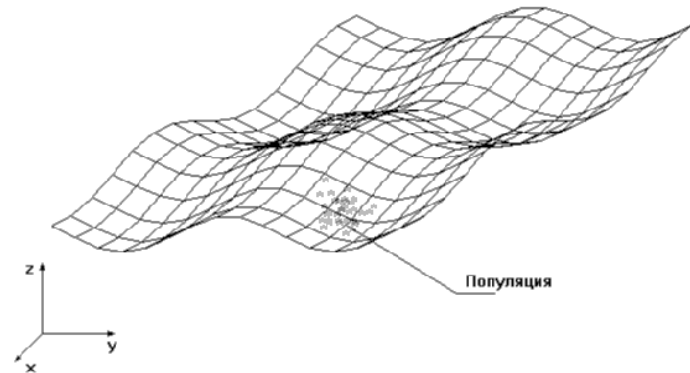
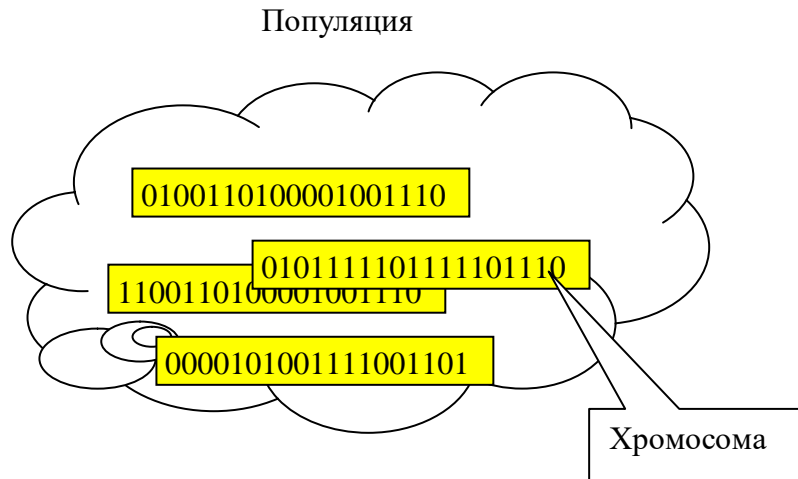


Карпов В.Э.

Генетические алгоритмы

Введение

- ГА – это стохастические, эвристические оптимизационные методы, впервые предложенные Холландом в 1975 году. Считается, что они основываются на идее эволюции с помощью естественного отбора, выдвинутой Дарвином.
- ГА работают с совокупностью (популяцией) “особей”, каждая из которых представляет возможное решение данной проблемы.
- Классический ГА требует кодирования параметров или переменных, представляющих собой решение некоторой задачи, в виде двоичной строки - **хромосомы**.



Суть ГА

- Каждая особь оценивается мерой ее «приспособленности» согласно тому, насколько «хорошо» соответствующее ей решение задачи.
- Наиболее приспособленные особи получают возможность «воспроизводить» потомство с помощью «перекрестного скрещивания» с другими особями популяции.
- Иногда происходят мутации \Rightarrow появление новых особей, которые сочетают в себе некоторые характеристики, наследуемые ими от родителей.

-
- Из поколения в поколение, хорошие характеристики распространяются по всей популяции. В конечном итоге популяция будет сходиться к субоптимальному решению задачи.
 - Полагается, что преимущество ГА состоит в том, что он находит **приблизительные оптимальные** решения за относительно короткое время.



Компоненты ГА

- *Хромосома*. Вектор (последовательность) генов, представляющее решение рассматриваемой задачи.
- *Начальная* популяция хромосом.
- *Набор операторов* для генерации новых решений из предыдущей популяции.
- *Целевая функция* для оценки пригодности (fitness) решений.

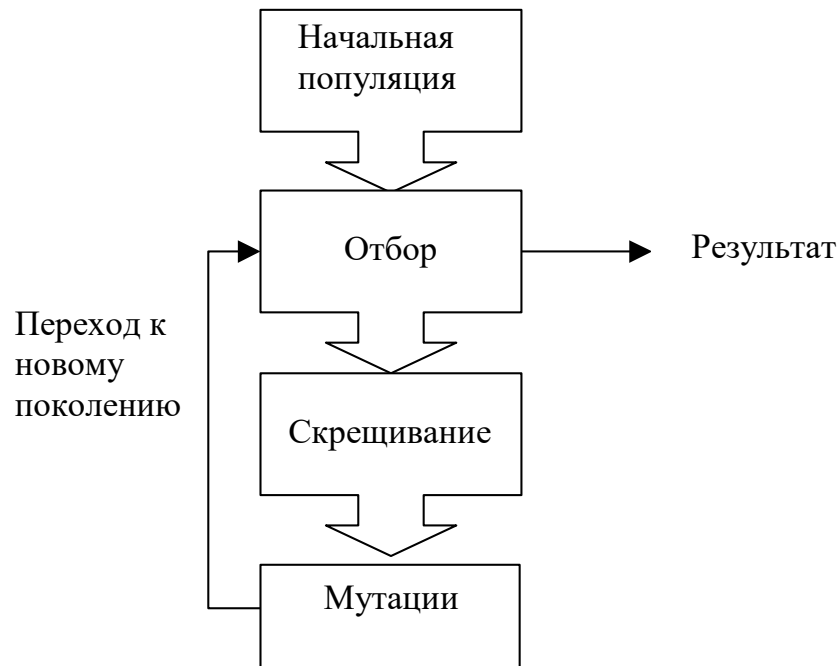
Соответствие терминологии естественного отбора в природе и ГА

<i>Приспособленность индивидуума</i>	Значение целевой функции на этом индивидууме.
<i>Выживание наиболее приспособленных</i>	Популяция следующего поколения формируется в соответствии с целевой функцией. Чем приспособленнее индивидуум, тем больше вероятность его участия в кроссовере, т.е. размножении

Алгоритм работы ГА

Работа ГА - итерационный процесс, который продолжается до тех пор, пока не выполнятся заданное число поколений или какой-либо иной критерий останова.

На каждом поколении ГА реализуется отбор (пропорционально приспособленности), скрещивание и мутация.



Алгоритм

-- генетический алгоритм

НАЧАЛО

1. Создать начальную популяцию
 2. Оценить приспособленность каждой особи
 3. останов := FALSE
 4. ПОКА НЕ останов ВЫПОЛНЯТЬ
 5. -- создать популяцию нового поколения
 6. ПОВТОРИТЬ (размер_популяции/2) РАЗ
 7. -- цикл воспроизводства
 8. Выбрать две особи с высокой приспособленностью из предыдущего поколения
 9. Скрестить выбранные особи и получить двух потомков
 10. Оценить приспособленности потомков
 11. Поместить потомков в новое поколение
 12. КОНЕЦ ЦИКЛА
 13. ЕСЛИ популяция сошлась, ТО останов := TRUE
 14. КОНЕЦ ЦИКЛА
- КОНЕЦ

Операторы ГА

Стандартные операторы для всех типов ГА:

- *Селекция (reproduction)*. Отбор хромосом в соответствии со значениями их функции пригодности.
- *Скращивание (crossover)*. Обмен частями хромосом между двумя (может быть и больше) хромосомами в популяции. Может быть однотоочечным или многотоочечным.
- *Мутация (mutation)*. Стохастическое изменение части хромосом.

Селекция (отбор)

Оператор селекции (reproduction, selection) осуществляет отбор хромосом в соответствии со значениями их функции приспособленности.

- **Турнирный отбор** (tournament selection) реализует n турниров, чтобы выбрать n особей. Каждый турнир построен на выборке k элементов из популяции, и выбора лучшей особи среди них. Наиболее распространен турнирный отбор с $k=2$.
- **Пропорциональный отбор** (fitness proportionate selection). Иногда называется **методом рулетки** (roulette-wheel selection) Основан на включении особи в новую популяцию по вероятностному принципу - отбирает особей с помощью N «запусков» рулетки. Колесо рулетки содержит по одному сектору для каждого члена популяции. Размер i -го сектора пропорционален соответствующей величине $P_{sel}(i)$:

$$P_{sel}(i) = \frac{f(i)}{\sum_{i=1}^N f(i)}$$

Здесь $f(i)$ – значение функции приспособленности i -й особи.

При таком отборе члены популяции с более высокой приспособленностью с большей вероятностью будут чаще выбираться, чем особи с низкой приспособленностью.

Обычно дает плохие результаты.

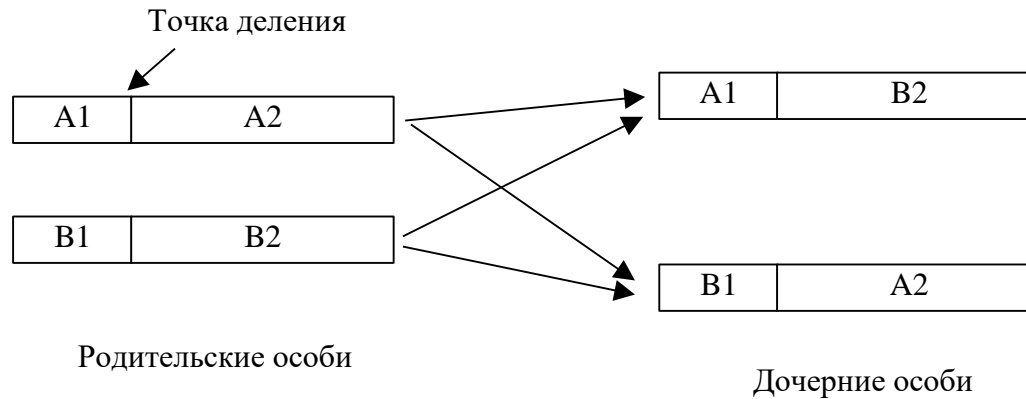
Селекция. Продолжение

- **Элитный отбор.** Построение новой популяции только из лучших особей репродукционной группы, объединяющей в себе родителей, их потомков и мутантов. Считается весьма слабым с точки зрения эффективности поиска (потенциальная опасность преждевременной сходимости).
- **Отбор вытеснением.** Отбор носит бикритериальный характер: то, будет ли особь из репродукционной группы заноситься в популяцию нового поколения, определяется не только величиной ее приспособленности, но и тем, есть ли уже в формируемой популяции следующего поколения особь с аналогичным хромосомным набором. Таким образом:
 - не теряются лучшие найденные решения, обладающие различными хромосомными наборами,
 - в популяции постоянно поддерживается достаточное генетическое разнообразие.
 - Вытеснение в данном случае формирует новую популяцию скорее из далеко расположенных особей, вместо особей, группирующихся около текущего найденного решения. Этот метод особенно хорошо себя показал при решении многоэкстремальных задач

Скрещивание

Кроссовер

- Одноточечный
- Многоточечный

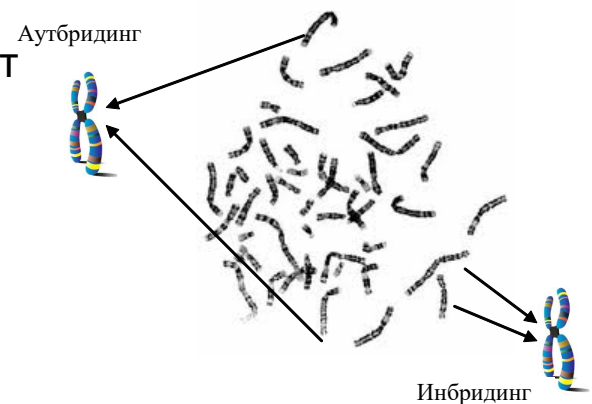


Выбор родительской пары

- **Случайный выбор** родительской пары («панмиксия» - свободное скрещивание). Любая особь может стать членом нескольких пар.
- **Селективный способ.** «Родителями» могут стать только те особи, значение приспособленности которых не меньше среднего значения приспособленности по популяции. Такой подход обеспечивает более быструю сходимость алгоритма.
- **Инбридинг и аутбридинг.** Оба метода построены на формировании пары на основе близкого и дальнего «родства» соответственно.

• **Инбридинг.** Первый член пары выбирается случайно, а вторым с большей вероятностью будет максимально близкая к нему особь. Концентрация поиска в локальных узлах, что приводит к разбиению популяции на отдельные локальные группы локальных экстремумов.

• **Аутбридинг.** Формирует пары из максимально далеких особей. Направлен на предупреждение сходимости алгоритма к уже найденным решениям.

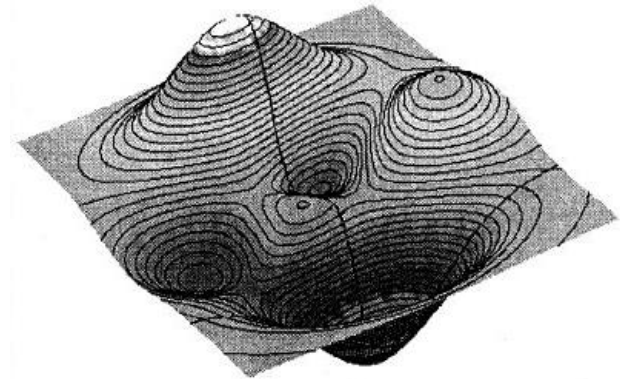
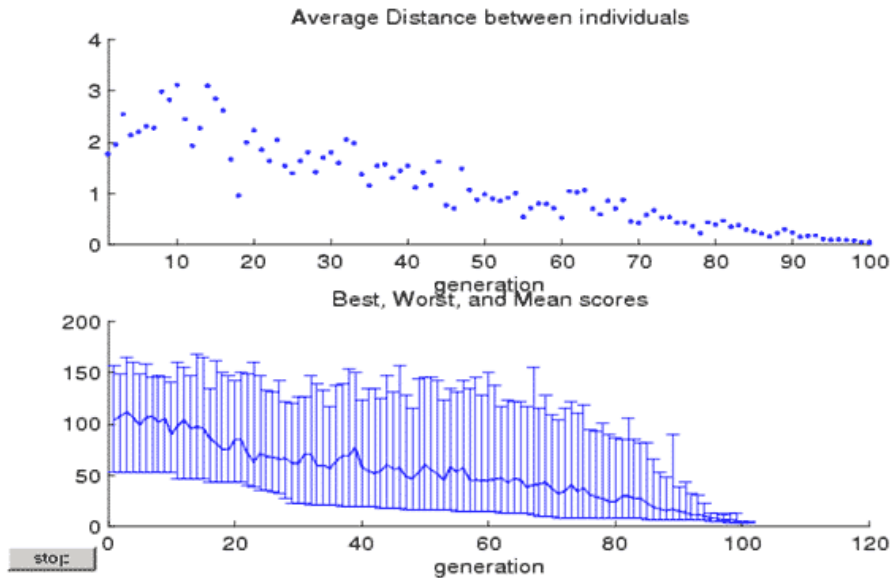


Мутации

- Основой появления новых видов хромосом является кроссовер. Мутации в ГА носят *вспомогательный* характер. Они заключаются в случайном изменении разрядов битовых строк.
- Мутации могут носить и более сложный характер – быть кратными, циклическими и проч.

Количественные оценки параметров ГА

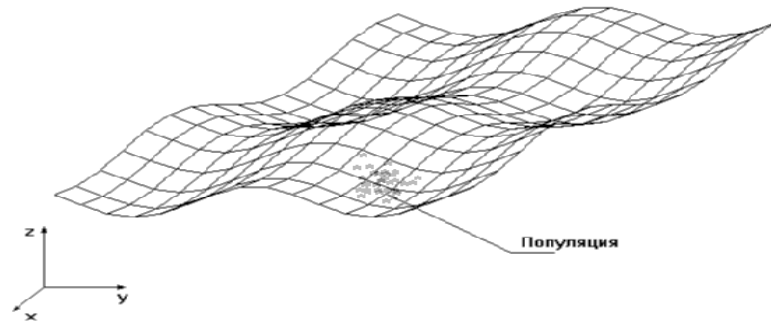
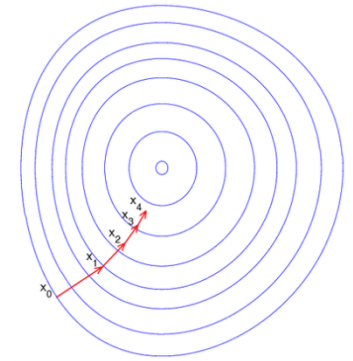
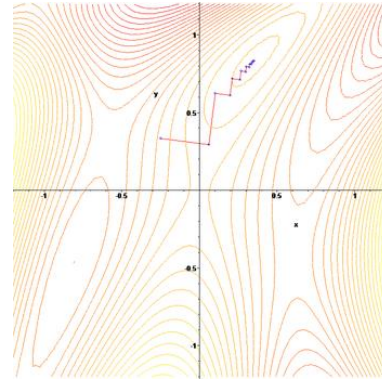
- Эффективность поиска чаще всего определяется как отношение среднего значения приспособленности популяции к максимальному значению после некоторого количества вычислений.
- Обычно исследуются несколько десятков поколений. Численность популяции должна быть достаточно большой – от нескольких десятков до нескольких сотен особей.
- Формирование генотипов особей начальной популяции желательно проводить по принципу максимального разнообразия. Максимальное побитовое разнообразие призвано обеспечить максимальное богатство генетического материала в начальной популяции.



Типичные задачи ГА

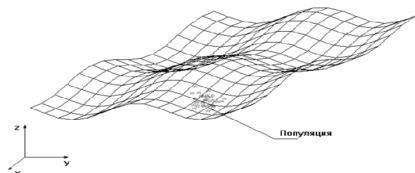
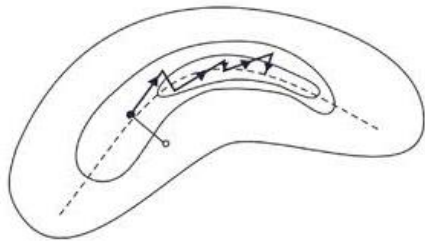
3 класса задач, типичных для ГА:

1. задача быстрой локализации одного экстремума,
2. задача определения нескольких (или всех) глобальных экстремумов,
3. задача описания ландшафта исследуемой функции, которая может сопровождаться выделением не только глобальных, но и локальных максимумов.



Тестовые задачи

- Овражные и многоэкстремальные функции



j	Тестовая задача	Размерность	Свойства
1	De Jong 2	2	Овражная функция, один глобальный экстремум
2	De Jong 3	5	Разрывы типа «скачок», максимум достигается на гиперкубе
3	De Jong 5	2	Один глобальный на гиперкубе, 24 локальных максимума.
4	Растригина	2	96 локальных экстремумов и 4 глобальных.
5	Griewank	2	Один глобальный и множество локальных максимумов.
6	Растригина	10	Один глобальный экстремум и $10^{10} - 1$ локальных.
7	Griewank	10	Один глобальный и множество локальных максимумов

Функции

Тестовая функция j1

$$F(x_1, x_2) = \frac{100}{100(x_1^2 - x_2)^2 + (1 - x_1)^2 + 1}, -1.28 \leq x_{1,2} \leq 1.28.$$
$$F^* = F(1, 1) = 100$$

Тестовая функция j2

$$F(x_1, x_2, \dots, x_5) = \sum_{i=1}^5 \lceil x_i \rceil, -5.12 \leq x_{1,2,3,4,5} \leq 5.12.$$
$$F^* = 30 \text{ на гиперплоскости } 5.00 \leq x_{1,2,3,4,5} \leq 5.12$$

Тестовая функция j3

$$F(x_1, x_2) = 0,002 + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^2 (x_i - a_{ij})^6}, \text{ где } a_{1j} = 16[(j \bmod 5) - 2], a_{2j} = 16[(j \% 5) - 2],$$
$$F^* = F(-16, -32) = 1.002, \text{ при } -65.536 \leq x_{1,2,3,4,5} \leq 65.536.$$

Тестовая функция j4

$$F(x_1, x_2) = 20 + x_1^2 + x_2^2 - 10 \cos(2\pi x_1) - 10 \cos(2\pi x_2), -5.12 \leq x_{1,2} \leq 5.12.$$
$$F^* = F(4.52299, 4.52299) = F(-4.52299, 4.52299) =$$
$$= F(-4.52299, -4.52299) = F(4.52299, -4.52299) = 80.7065.$$

Функции 5-7

Тестовая функция j5

$$F(x_1, x_2) = \frac{1}{\frac{x_1^2 + x_2^2}{200} - \cos(x_1) \cos\left(\frac{x_2}{\sqrt{2}}\right) + 2}, -20 \leq x_{1,2} \leq 20$$

$$F^* = F(0, 0) = 1.0.$$

Тестовая функция j6

$$F(x_1, x_2, \dots, x_{10}) = \sum_{i=1}^{10} \left(10 \cos(2\pi x_i) - x_i^2 \right) - 100$$

$$F^* = F(0, 0, \dots, 0) = 0.00, \text{ при } -5.12 \leq x_{1,2, \dots, 10} \leq 5.12.$$

Тестовая функция j7

$$F(x_1, x_2, \dots, x_{10}) = 10 - \left\{ \sum_{i=1}^{10} \frac{x_i^2}{4000} - \prod_{i=1}^{10} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1.0 \right\}, -512 \leq x_{1,2, \dots, 10} \leq 512.$$

$$F^* = F(0, 0, \dots, 0) = 10$$

Недостатки ГА

В ГА декларируется:

1. Способность простого представления (битовых строк) для кодирования сложных структур.
2. Мощность простого преобразования для улучшения таких структур.
3. ГА - простая модель эволюции в природе. В нем используются как аналог механизма генетического наследования, так и аналог естественного отбора.
4. Генотип определяет все.

Но:

- Каждому генотипу может соответствовать целое множество фенотипов. И наоборот.
- Мутации в ГА - это лишь дополнительный механизм. На самом деле именно мутации являются основой генетического механизма эволюции. Более того, как раз принципы размножения оказываются вторичными по сравнению с общим принципом вариабельности экземпляров вида, основанной на случайных мутациях.
- Основная техническая проблема ГА заключается в форме представления решения задачи. Язык описания в виде цепочек символов является слишком ограниченным и негибким для того, что бы можно было описать на нем сколько-нибудь сложные понятия. Это слишком сужает область применения ГА.
- Код Грея предпочтительнее обычного двоичного тем, что обладает свойством непрерывности бинарной комбинации: изменение кодируемого числа на единицу соответствует изменению кодовой комбинации только в одном разряде.

Код Грея

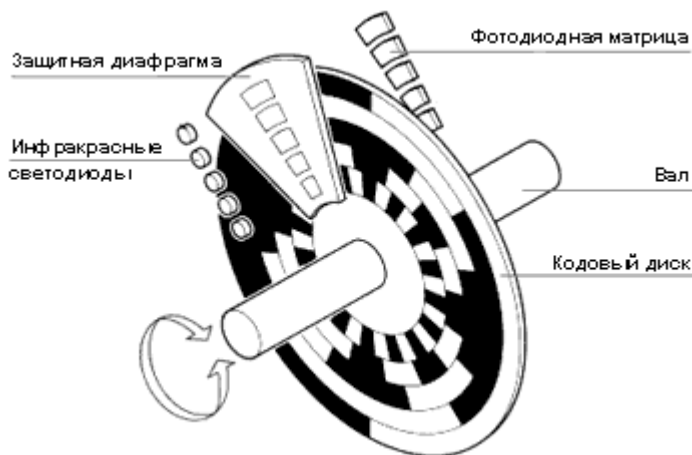
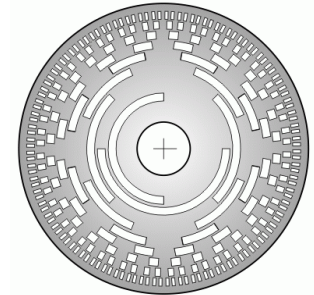
Особенность кода Грея - высокая **помехозащищенность**, т.к. при его использовании соседние целые числа отличаются друг от друга только на один бит (в одном разряде).

Для кода Грея декодирующая функция имеет вид:

$$V(S_{\Gamma}) = a + \frac{D(S_{\Gamma})}{2^n - 1} (b - a)$$

где a и b – границы отрезка $[a, b]$, включающего в себя значения варьируемого параметра V ; $D(S_{\Gamma})$ – десятичное представление кода Грея $S_{\Gamma} = (\sigma_n, \sigma_{n-1}, \dots, \sigma_2, \sigma_1)$, вычисляемое по формуле

$$D(S_{\Gamma}) = \sum_{i=1}^n \left(\bigoplus_{j=1}^i \sigma_j \right) 2^{i-1}, \text{ где } \bigoplus - \text{сумма по модулю 2.}$$

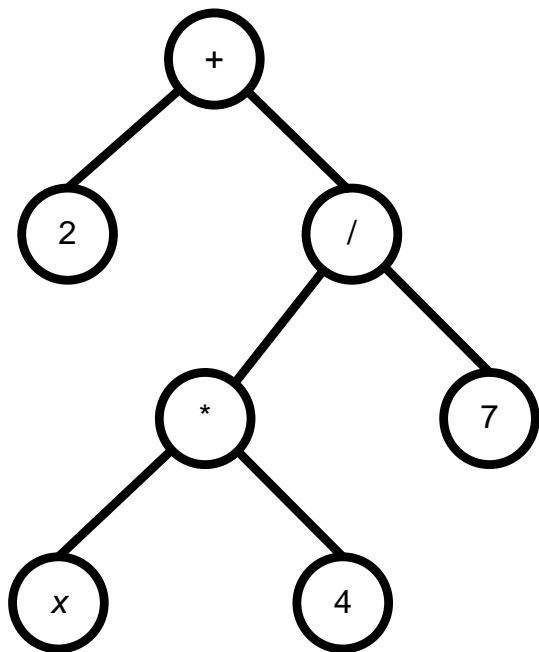


Число V	Двоичный код (s)	Код Грея (s_{Γ})
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100
9	1001	1101
10	1010	1111
11	1011	1110
12	1100	1010
13	1101	1011
14	1110	1001
15	1111	1000

Генетическое программирование

Коза (Koza), 1992 г.

- В отличие от ГА в ГП все операции производятся не над строками, а над **деревьями**. При этом используются такие же операторы, как и в ГА: селекция, скрещивание и мутация.
- В ГП хромосомами являются **программы**. Программы представлены в виде деревьев с функциональными (промежуточными) и терминальными (конечными) элементами.



Функция (программа)
 $2+x*4/7$

Решение задачи в ГП

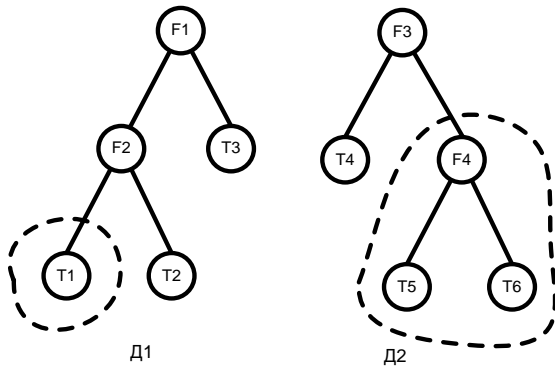
Необходимо определить:

- Множество терминальных элементов.
- Множество функциональных элементов.
- Мету приспособленности (fitness).
- Параметры, контролирующие эволюцию.
- Критерий останова эволюции.

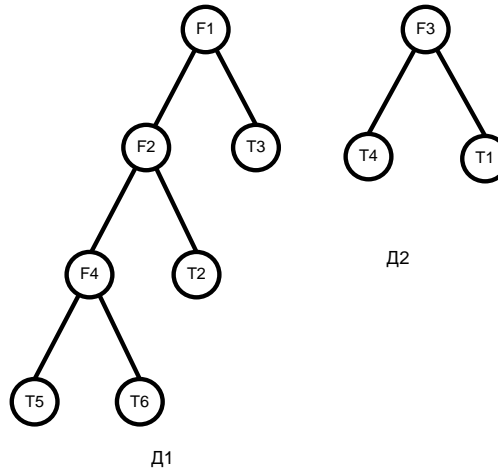
Особенности ГП

- Размер хромосомы меняется.
- Для предотвращения чрезмерного разрастания дерева вводят максимальное количество функциональных элементов в дереве или максимальную глубину дерева.
- В случае превышения лимита при скрещивании вместо конфликтного дерева копируется родительское дерево.

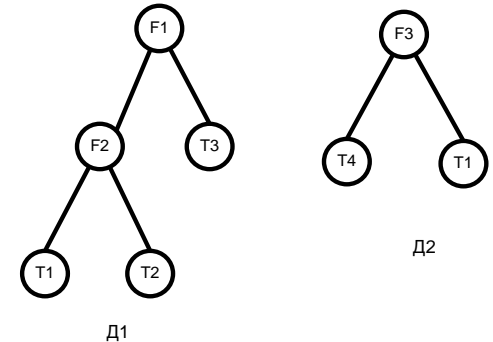
Исходные деревья



Деревья до разрешения конфликта



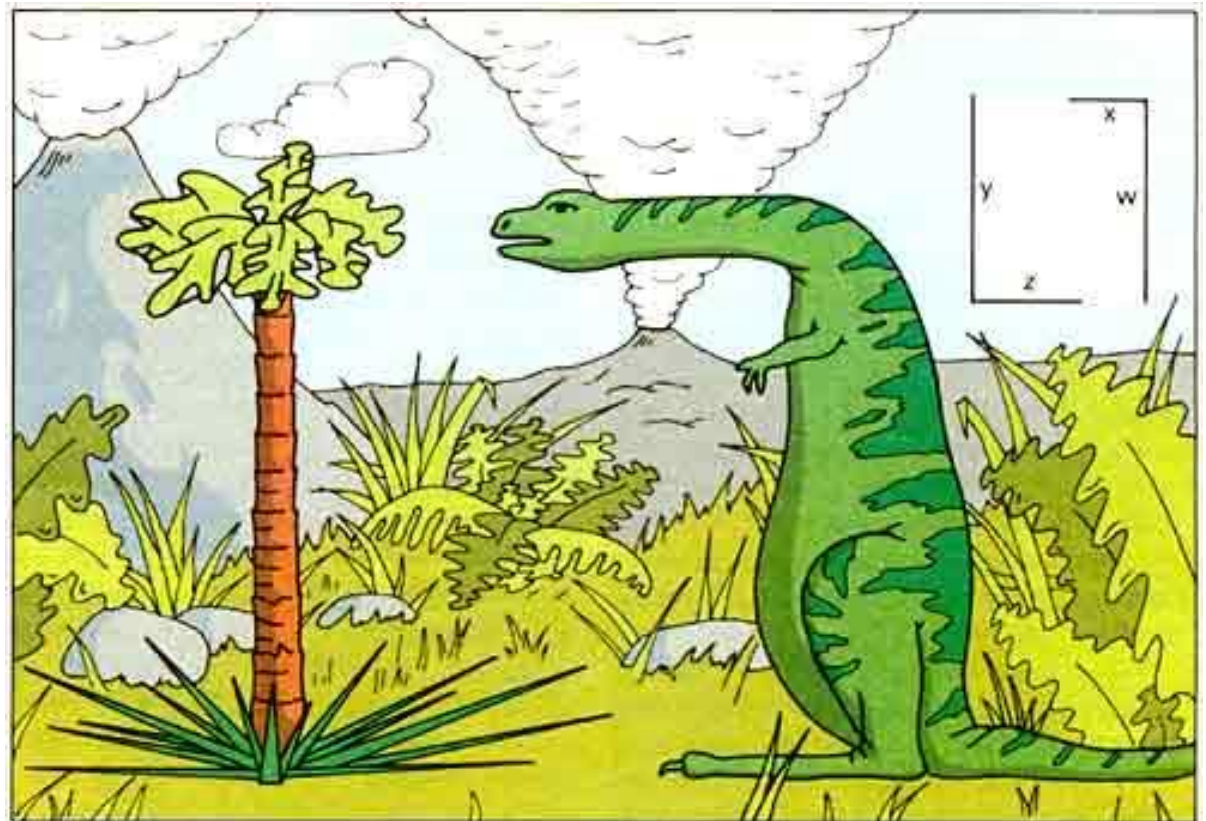
Деревья после разрешения конфликта



Пример 1.

МИРЕ НАУКИ. (*Scientific American*) • 1988 No4. стр. 81-85

А.К. Дьюдни «Моделирование эволюции в мире биоморфов»



• *Взаимозависимая эволюция шипофита и гнутаозавра*

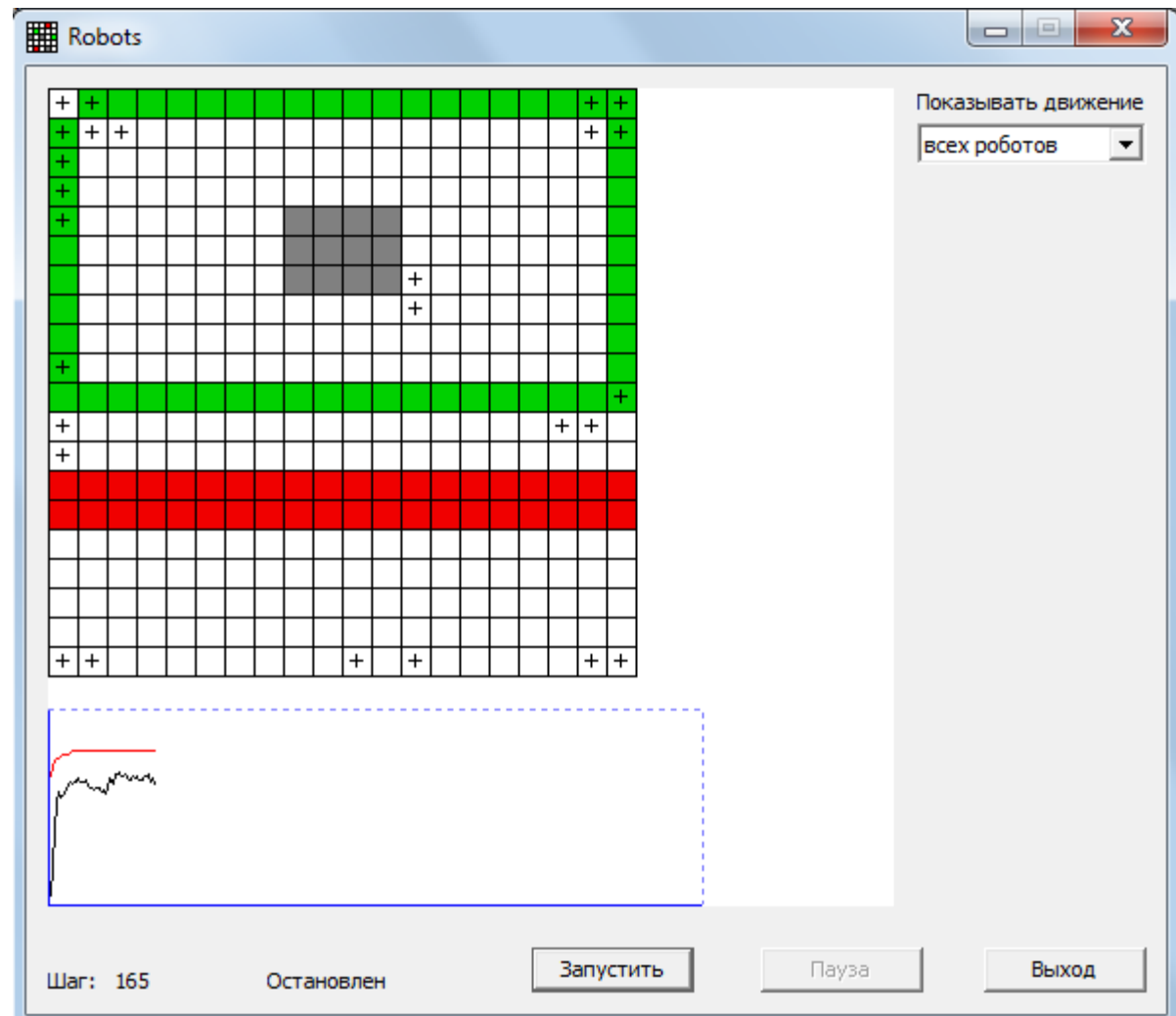
Программа

The screenshot displays the 'Гнутазагры' (GnutaZagry) software interface, which simulates the evolution of plants and animals. The interface is divided into several sections:

- Растения (Plants):** A bar chart showing the population of five plant species over time. The bars are green. Below the chart are five small diagrams of plant structures, each with a downward-pointing arrow, indicating a mutation path. To the right, a dashed box contains a single downward-pointing arrow.
- Животные (Animals):** A bar chart showing the population of five animal species over time. The bars are green. Below the chart are five small diagrams of animal structures, each with a rightward-pointing arrow, indicating a mutation path. To the right, a dashed box contains a rightward-pointing arrow.
- Мутации (Mutations):** A control panel with a text input field set to '0.0500' and a checkbox labeled 'Остановить эволюцию' (Stop evolution), which is currently unchecked.
- Simulation Area:** A large dashed box on the right side of the window, containing a complex diagram of evolutionary paths. It shows a vertical line with a horizontal branch to the left, and a horizontal line with a vertical branch to the right, both ending in downward-pointing arrows.
- Buttons:** At the bottom, there are buttons for 'Шаг: 572' (Step: 572), 'Остановлен' (Stopped), 'Запустить' (Start), 'О программе...' (About...), and 'Выход' (Exit).

Пример 2. Задача о роботе, собирающем еду

- На поле - пища, стена или мина
- По краям мир огорожен стеной.
- Задача робота – собрать всю еду за наименьшее количество шагов



Устройство работа

Функция приспособленности работа:

$$f = \frac{\sum enc}{\sum enc - \sum pin}$$

Ген работа (8 бит):

n	n	n	s	s	a	a	a
---	---	---	---	---	---	---	---

Где nnn – номер датчика (0...7), ss – состояние датчика, aaa – действие (движение).

Состояния датчика:

Код (ss)	В клетке находится
00	Еда
01	Стена
10	Мина
11	состояние не используется

Действия:

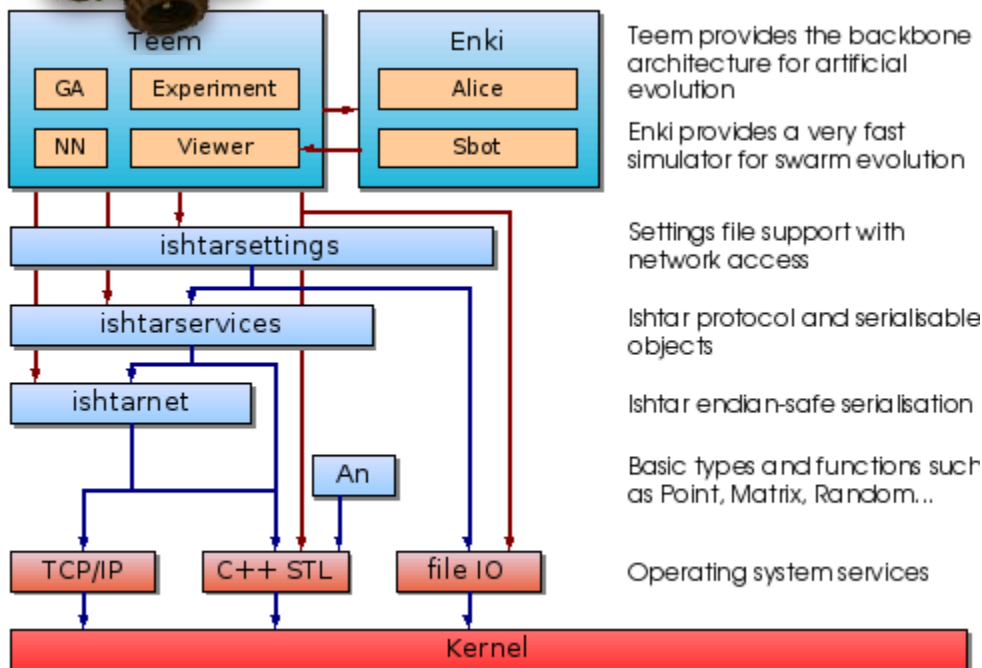
Код (aaa)	Направление движения	Физический эквивалент действия
000	вверх-влево	правый двигатель вперёд, затем оба двигателя вперёд, затем левый двигатель вперёд
001	Влево	правый двигатель вперёд и левый двигатель назад, затем оба двигателя вперёд, затем левый двигатель вперёд и правый двигатель назад
010	вниз-влево	правый двигатель назад, затем оба двигателя назад, затем левый двигатель назад
011	Вверх	оба двигателя вперёд
100	Вниз	оба двигателя назад
101	вверх-вправо	левый двигатель вперёд, затем оба двигателя вперёд, затем правый двигатель вперёд
110	Вправо	левый двигатель вперёд и правый двигатель назад, затем оба двигателя вперёд, затем правый двигатель вперёд и левый двигатель назад
111	вниз-вправо	левый двигатель назад, затем оба двигателя назад, затем правый двигатель назад

Пример 3. ГА и НС

Laboratory of Intelligent Systems (Политехническая школа, Лозанна, Швейцария). Группа из 10 роботов. Состязания за пищу.



s-бот: 12x15 см.; процессор Xscale 400 MHz, 64 MB ОЗУ, 32 MB флеш, 12 PIC микроконтроллеров для низкоуровневой обработки.



Искусственная эволюция проходила в симулирующей среде [Enki](#), где моделировались роботы. Затем был использован эволюционный фреймворк роботов [Teem](#) для эволюции лучших «контроллеров», которые переносились на реальные роботы.

Задача

- Задача - поиск «источника пищи» (светящееся кольцо)
- На другом конце арены - более темное кольцо («отравленное»)
- Роботы: ИНС. 11 сенсорных нейронов и 3 мотонейрона (колеса и свет). Нейроны соединены с помощью 33 синапсов, и мощность сигналов каждого синапса контролировалась одним восьмибитным геном. Т.о., каждый робот имеет **264-битный геном**, который контролирует, как он будет реагировать на информацию, поступающую с сенсоров.

