1. We want to solve the Traveling Salesman Problem (TSP) by using a genetic algorithm. Please give the pseudo-code of the algorithm and key data structures in it.

Solution：

伪代码如下：

/\*说明：用遗传算法解决TSP问题\*/

/\*初始化遗传算法参数\*/

const int maxGeneration; //进化代数，即迭代次数

const int sizeGene; //种群规模

double pcross; //交叉概率选择，0和1之间

double pmutation; //变异概率选择，0和1之间

int \*bestGene; //适应度最好的染色体

double bestFitness; //记录进化中最高的适应度

struct gene //每个染色体的结构

{

int \*cityorder[]; // 路径

int fitness; // 适应度函数值

};

bool GenSelection()

/\*选择\*/

{采用转轮法进行选择}

bool GenCrossover()

/\*交叉\*/

{采用单点交叉}

bool GenMutate()

/\*突变\*/

{随机选择染色体某一位置，将其值取反}

double Fitness()

/\*适应度函数计算\*/

{double fit=F()为相应路径长度的负数;}

main()

{

//初始化种群  
 for(i=0,i< sizeGene;i++)

{

随机产生一个种群  
 计算其适应度

}

// 进化开始

t = 1 // 进化代数

do

{

选择;

交叉;

变异;

重新计算适应度函数;

bestGene = 此代进化中适应度最高的染色体

进化代数t = t + 1;

}while(t< maxGeneration) // 进化结束

bestGene存储的即为最佳路径;

}

**English Version:**

Pseudocode is as follows:

/\* Note: using genetic algorithm to solve TSP problem \*/

/\* Initialize the parameters of genetic algorithm \*/

const int maxGeneration; Evolving algebra, number of iterations

const int sizeGene; Population size

double pcross; Crossover probability choice, between 0 and 1

double pmutation; Mutation probability choice, between 0 and 1

int \*bestGene; Fitness the best chromosome

double bestFitness; Record highest evolutionary fitness

Struct gene//of each chromosome structure

{

int \*cityorder[]; Path

int fitness; Fitness function value

};

bool GenSelection()

/\* Select \*/

{Runner method was used to select}

bool GenCrossover()

/\* Cross \*/

{The use of single-point crossover}

bool GenMutate()

/\* Mutation \*/

{Random chromosome to a location, the value inverse}

double Fitness()

/\* Fitness function computation \*/

{Double fit=F () as the negative of the corresponding path length;}

main()

{

Initialize population

for(i=0,i< sizeGene;i++)

{

Random group

Calculation of the fitness

}

Evolution begins

T = 1//evolution of algebra

do

{

Choose;

Cross;

Variation;

Fitness function is recalculated;

BestGene = fitness in this generation the highest chromosome

Evolving algebra t = t + 1;

}While (t< maxGeneration)//end of evolution

BestGene store the best path;

}

2. Given a function ，where  is a real variable in [0, 31]. Please design a particle swarm optimization algorithm to find out the maximum and minimum values of .

Solution：

算法步骤：

Step1. 初始化有关参数

c1=2; //学习因子1

c2=2; //学习因子2

w=0.5; //惯性权重

MaxDT=1000; //最大迭代次数

D=1; //搜索空间维数（即未知数个数，此处未知数为x）

N=40; //初始化群体个体数目

Step2.初始化群体中的个体

Step2.1 按随机方法初始化粒子群中各粒子的位置和速度，其中保证位置是在[0，31]之间。

Step2.2 计算各个粒子对应的函数值：。

Step2.3 初始化Pi和Pg，Pi为每个粒子的当前位置，Pg为所有粒子中对应最大函数值的粒子的位置。

Step3. 进入循环，按照以下公式依次迭代改变每个粒子，同时保证每次改变后的位置值均在[0，31]之间，直到迭代次数大于MaxDT：





Step4. 取所有迭代中出现过的最好解作为结果。

English Version:

Solution:

Algorithm steps are:

Step1. Initialize parameters

c1=2; Factor 1

c2=2; Factor 2

w=0.5; Inertia weight

MaxDT=1000; Maximum number of iterations

D=1; Dimension of the search space (that is, number of unknowns, unknown for x here)

N=40; Initial number of groups of individuals

Step2. initialize groups of individuals

Step2.1 Summary of stochastic method to initialize the particle swarm in the particle'sposition and velocity, which ensure the location is in [0,31] in between.

Step2.2 calculate the value of each particle corresponds to the function.

Step2.3 Initializes the Pi and Pg,Pi for the current position of each particle, Pg for allparticle corresponds to the maximum value of the position of the particle.

Step3. Enter the loop, according to the formula followed by iterative change eachparticle, while ensuring value after each change of location in [0,31], and until thenumber of iterations than MaxDT:

Step4. All appear in the iteration as a result has been the best solution.

3. Use *Q*-learning method to solve the following 2-rank Hanoi problem：

 

Solution：

1) to obtain the corresponding state-space. Assuming reward value of 100 to thedestination state action, the remaining reward value of 0. Cumulative discount forreturn, the discount rate is 0.9. Calculate the Q-value, and on the correspondingposition marked on the State space.

2）According to the above, could be envisaged-agent from (1,1) State, in accordance withthe principle of maximum Q-, executed in each State to select the appropriate actions,change its status until the target State. The end result is (1,1) ◊ (3,1) ◊ (3,2) ◊ (2,2).

3）These problem-solving methods and problem-solving method in the fundamentaldifference: the agent only care about optimal mapping relationships between Stateand behavior. By roaming on the environment to adapt to the environment, so thatyou can get from an arbitrary initial state to the goal State of the best actionsequences, accordingly to achieve the solution. Rather than when given an initialstate, then search and optimization possible path.