
Algoritmos Genéticos

y

sus

Aplicaciones

Angel Kuri M.
Centro de Investigación en Computación
Instituto Politécnico Nacional

oct. 2000

Computación Evolutiva

1 Computación Evolutiva

» Algoritmos Genéticos

2 Aplicaciones

2.1 Optimización Numérica

2.2 Asignación de Segmentos en Bases de Datos Distribuidas

Computación Evolutiva

1. Simulación parcial del proceso de selección natural
2. No requiere de un modelo matemático para atacar un problema dado
3. No alcanza resultados óptimos sino cercanos al óptimo en poco tiempo

Algoritmo Evolutivo

- 1. Generar un conjunto de posibles soluciones**
- 2. Evaluar el desempeño del sistema para cada individuo.**
- 3. Verificar si ya se ha alcanzado algún criterio de convergencia.**
 - Si: Terminar.**
 - No: Proceder con el paso 4.**
- 4. Seleccionar a los mejores individuos de acuerdo con su evaluación.**
- 5. Modificar a cada uno de los individuos en base a su desempeño para obtener un nuevo conjunto de posibles soluciones.**
- 6. Proceder con el paso 2.**

Programa de Ilustración

- El programa que va a ilustrarse forma parte del libro
- “A Comprehensive Approach to Genetic Algorithms in Optimization and Learning”
- El software puede bajarse de 148.204.45.189/galsk/

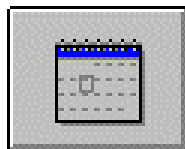
Chapter 1



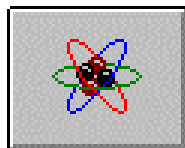
Chapter 2



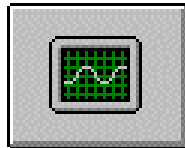
Chapter 3



Chapter 4



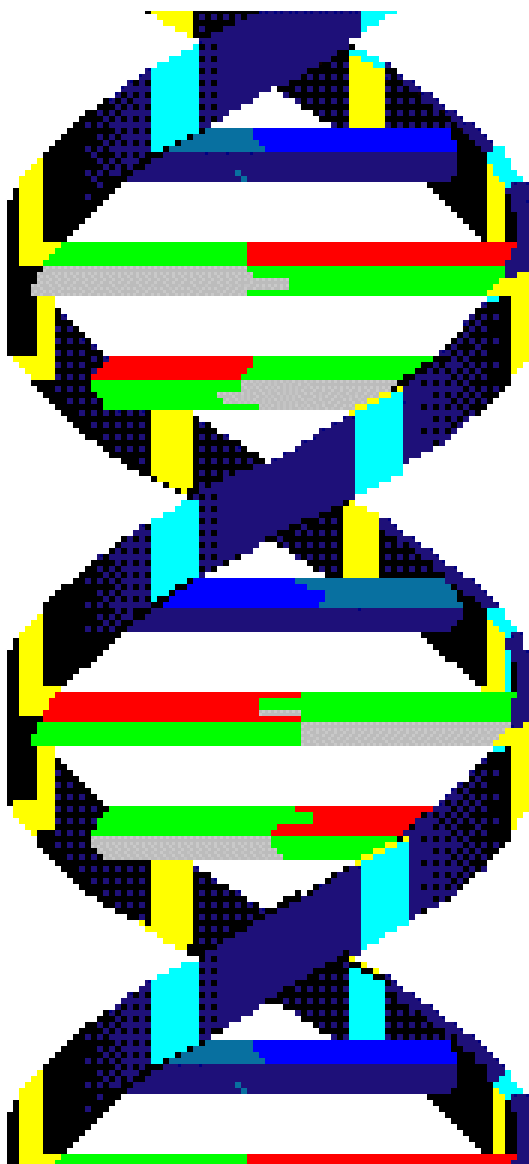
Chapter 5



Chapter 6



Chapter 7



Introduction

Preliminary
Concepts

Evolutionary
Computation

Theoretical
Considerations

GAs Applied to
the Optimization
of Functions

GAs in Learning

A Universal
Genetic Algorithm

THIS IS AN UNPROTECTED COPY

Ins



7:46 AM

SIMPLE GENETIC ALGORITHM

Edit

Simple Genetic Algorithm (Pc=1; Pm=0.05)

Function
to Optimize

- ☒ $2X^2 - 3$
☐ $3X^3 - 4X^2 + X - 4$
☐ $(X^2 - 2) / X$

Reset
Algorithm



Generate

Initial Population

Evaluate Population

Cumulative Distribution Function

New Population

Crossover

Mutation

Next Generation

Show

Generation Number

0

Run Automatically

Show Fitness

Show History

End

COMPL

EXPIRATION DATE: 21/02/2000

Ins

Num

 Inicio

 Administrad...

 Ciclo de IA ...

 Microsoft P...

 CALEND...

 SIMPLE ...

 11:36 AM

SIMPLE GENETIC ALGORITHM

Edit

Simple Genetic Algorithm (Pc=1; Pm=0.05)

Function
to Optimize

- ☒ $2X^2 - 3$
- ☐ $3X^3 - 4X^2 + X - 4$
- ☐ $(X^2 - 2) / X$

Generation Number

0

Reset
Algorithm

Random Number Generator Initialization

Random Number Seed

3.1416

Continue

Generate

Init

Evaluate

Cumulative

New Population

Crossover

Mutation

Next Generation

OH S

Automatically

tness

Show History

End

COMPL

EXPIRATION DATE: 21/02/2000

Ins

Num

Inicio

Administrad...

Ciclo de IA ...

Microsoft P...

CALENDA...

SIMPLE ...

11:37 AM

SIMPLE GENETIC ALGORITHM

Edit Visor

Simple Genetic Algorithm (Pc=1; Pm=0.05)

Function
to Optimize

- ☒ $2X^2 - 3$
☐ $3X^3 - 4X^2 + X - 4$
☐ $(X^2 - 2) / X$

Reset
Algorithm



Generate

Initial Population

Evaluate Population

Cumulative Distribution Function

New Population

Crossover

Mutation

Next Generation

Show

Generation Number

0

 <Esc> to Continue

Number	Binary
888	1101111000
540	1000011100
235	0011101011
590	1001001110
745	1011101001
647	1010000111
149	0010010101
534	1000010110
408	0110011000
1002	1111101010
773	1100000101
185	0010111001

COMPL

EXPIRATION DATE: 21/02/2000

Ins

Num

 Inicio

 Administrad...

 Ciclo de IA ...

 Microsoft P...

 CALEND...

 SIMPLE ...

 11:38 AM

SIMPLE GENETIC ALGORITHM

Edit Visor

Simple Genetic Algorithm (Pc=1; Pm=0.05)

Function
to Optimize

- ☒ $2X^2 - 3$
☐ $3X^3 - 4X^2 + X - 4$
☐ $(X^2 - 2) / X$

Reset
Algorithm



Generate

Initial Population

Evaluate Population

Cumulative Distribution Function

New Population

Crossover

Mutation

Next Generation

Show

Generation Number

0

<Esc> to Continue

Fitness	Binary
157708	0000001000001111101
583197	0001110011000011101
110447	0011010111101101111
696197	0101001111110000101
1110047	0001111000000011111
837215	1001100011001011111
44399	0001010110101101111
570309	0001011001111000101
332925	1010001010001111101
2008005	1101010001111000101
1195055	0100011110000101111
68447	0010000101101011111
1984029	1100100011000011101
38639	0001001011011101111

COMPL

EXPIRATION DATE: 21/02/2000

Ins

Num

Inicio

Administrad...

Ciclo de IA ...

Microsoft P...

CALENDA...

SIMPLE ...

11:38 AM

SIMPLE GENETIC ALGORITHM

Edit Visor

Simple Genetic Algorithm (Pc=1; Pm=0.05)

Function
to (

☒ $2X^2 - 3$

☐ $3X^3 - 4X^2 + X - 4$

Generation Number

0

<Esc> to Continue

Fitness	Normalized	Binary	CDF
2008005	2791229	1101010001111000101	0.03564217
1984029	2767253	1100100011000011101	0.07097819
1932575	2715799	1010111110100011111	0.10565717
1893455	2676679	1001110010001001111	0.13983661
1740975	2524199	0101001000010101111	0.17206898
1692797	2476021	0011101010001111101	0.20368615
1678109	2461333	0011001101100011101	0.23511577
1577085	2360309	0000001000001111101	0.26525537
1562909	2346133	1111101100100011101	0.29521396
1465469	2248693	1100101110001111101	0.32392830
1321935	2105159	1000010101111001111	0.35080981
1305725	2088949	0111110110001111101	0.37748432
1305725	2088949	0111110110001111101	0.40415884
1273605	2056829	0110110111100000101	0.43042320

Generate

COMPL

EXPIRATION DATE: 21/02/2000

Ins

Num

Inicio

Administ...

PROYE...

Microso...

CALEN...

Editor d...

SIMP...

4:23 PM

SIMPLE GENETIC ALGORITHM

Edit Visor

Simple Genetic Algorithm (Pc=1; Pm=0.05)

Function
to Optimize

- ☒ $2X^2 - 3$
☐ $3X^3 - 4X^2 + X - 4$
☐ $(X^2 - 2) / X$

Reset
Algorithm



Generate

Initial Population

Evaluate Population

Cumulative Distribution Function

New Population

Crossover

Mutation

Next Generation

Show

Generation Number

0

<Esc> to Continue

Number	Binary	
428	0110101100	
582	1001000110	
983	1111010111	
398	0110001110	
916	1110010100	
540	1000011100	
920	1110011000	
916	1110010100	
185	0010111001	
933	1110100101	
590	1001001110	
721	1011010001	
722	1011010010	
808	1100101000	

COMPL

EXPIRATION DATE: 21/02/2000

Ins

Num

Inicio

Administ...

PROYE...

Microso...

CALEN...

Editor d...

SIMP...

4:23 PM

SIMPLE GENETIC ALGORITHM

Edit Visor

Simple Genetic Algorithm (Pc=1; Pm=0.05)

Function
to Optimize

- ☒ $2X^2 - 3$
☐ $3X^3 - 4X^2 + X - 4$
☐ $(X^2 - 2) / X$

Reset
Algorithm



Generate

Initial Population

Evaluate Population

Cumulative Distribution Function

New Population

Crossover

Mutation

Next Generation

Show

Generation Number

0

<Esc> to Continue

Number	Binary		
870	1101100110		
600	1001011000		
440	0110111000		
876	1101101100		
940	1110101100		
421	0110100101		
916	1110010100		
582	1001000110		
188	0010111100		
537	1000011001		
997	1111100101		
982	1111010110		
871	1101100111		
301	0100101101		

COMPL

EXPIRATION DATE: 21/02/2000

Ins

Num

Inicio

Administ...

PROYE...

Microso...

CALEN...

Editor d...

SIMP...

4:24 PM

SIMPLE GENETIC ALGORITHM

Edit Visor

Simple Genetic Algorithm (Pc=1; Pm=0.05)

Function
to Optimize

- ☒ $2X^2 - 3$
☐ $3X^3 - 4X^2 + X - 4$
☐ $(X^2 - 2) / X$

Reset
Algorithm



Generate

Initial Population

Evaluate Population

Cumulative Distribution Function

New Population

Crossover

Mutation

Next Generation

Show

Generation Number

0

<Esc> to Continue

Number	Binary	
870	1101100110	
600	1001011000	
440	0110111000	
876	1101101100	
940	1110101100	
421	0110100101	
948	1110110100	
70	0001000110	
188	0010111100	
537	1000011001	
993	1111100001	
982	1111010110	
871	1101100111	
301	0100101101	

COMPL

EXPIRATION DATE: 21/02/2000

Ins

Num

Inicio

Administ...

PROYE...

Microso...

CALEN...

Editor d...

SIMP...



4:24 PM

SIMPLE GENETIC ALGORITHM

Edit

Simple Genetic Algorithm (Pc=1; Pm=0.05)

Function
to Optimize

- ☒ $2X^2 - 3$
- ☐ $3X^3 - 4X^2 + X - 4$
- ☐ $(X^2 - 2) / X$

Generation Number

1



 INDICATE THE NUMBER OF GENERATIONS YOU WISH TO SIMULATE

20

Continue

Gen

New Population

Crossover

Mutation

Next Generation

O
H
S

Show History



End

COMPL

EXPIRATION DATE: 21/02/2000

Ins

Num

 Inicio

 Administrad...

 Ciclo de IA ...

 Microsoft P...

 CALEND...

 SIMPLE ...

 11:39 AM

SIMPLE GENETIC ALGORITHM

Edit Visor

Simple Genetic Algorithm (Pc=1; Pm=0.05)

Function
to (

☒ $2X^2 - 3$

☐ $3X^3 - 4X^2 + X - 4$

Generation Number

21

<Esc> to Continue

N	Best Fitness	Average Fitness	Binary	Decimal
6	2,080,797	1,539,458	1111111100	522,240
7	2,056,389	1,486,081	1111110110	519,168
8	2,040,197	1,509,294	1111110010	517,120
9	2,040,197	1,591,934	1111110010	517,120
10	2,040,197	1,650,593	1111110010	517,120
11	2,093,055	1,707,444	1111111111	523,776
12	2,088,965	1,717,993	1111111110	523,264
13	2,040,197	1,678,616	1111110010	517,120
14	2,080,797	1,769,522	1111111100	522,240
15	2,088,965	1,616,726	1111111110	523,264
16	2,080,797	1,622,433	1111111100	522,240
17	2,084,879	1,753,589	1111111101	522,752
18	2,084,879	1,666,828	1111111101	522,752
19	2,084,879	1,496,513	1111111101	522,752
20	2,093,055	1,450,175	1111111111	523,776

Generate

COMPL

EXPIRATION DATE: 21/02/2000

Ins

Num

Inicio

Administrad...

Ciclo de IA ...

Microsoft P...

CALENDA...

SIMPLE ...

11:40 AM

Ejemplo de Aplicación

- **Optimización de una función con restricciones**

- » Función no lineal
- » Restricciones no lineales

Representación Numérica

Para representar un conjunto de números puede emplearse un formato de punto fijo

En él, cada número consta de signo, una parte entera y una parte decimal

La codificación suele hacerse con código Gray

Una Corrida de Ejemplo

Ejemplificamos el proceso maximizando la función

$$f(x, y) = x + |\sin(32x)| - |\cos(11y)|$$

sujeto a las siguientes restricciones

$$\Pi > x \geq 0$$

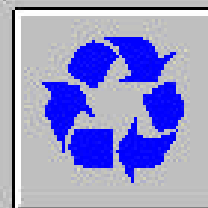
$$\Pi > y \geq 0$$

GENETIC OPTIMIZATION OF FUNCTIONS WITH/WITHOUT RESTRICTIONS

Edit Data Base Tools Genetic Optimization Help Exit Visor

Parameter Definition

Maximum Genome 20000
Number of Variables 2
Bits of the Integer Part 5
Bits of the Decimal Part 24
Present Genome 60



Previous Data



DEFINE CONDITIONS

Key In Conditions	
Number	Condition
1	$V01 \geq 0$
2	$V01 < \pi$
3	$V02 \geq 0$
4	$V02 < \pi$



Maximize



Minimize

FUNCTION: $F = V01 + \text{ABS}(\text{SIN}(32 * V01)) - \text{ABS}(\text{COS}(11 * V02))$

Accept Parameters

End

COMPUTER

VARIABLES MUST BE OF THE TYPE "VXX"

Ins

Num

Inicio

Administrad...

Ciclo de IA ...

Microsoft P...

CALENDA...

GENETI...

11:42 AM

 **Genetic Process Control**

Lower Pc	6	# of Individuals	100
Upper Pc	10	# of Generations	50
Lower Pm	4	Evaluations	5000
Upper Pm	5	Genome	60
Lower # of Descendants	1	Maximize/Minimize	1
Upper # of Descendants	2	Windows/Dos	1
Lower HC Percentage	8	Data Format	1
Upper HC Percentage	10	Random Seed	2.7182800000

External Function

C:\GALS2000\GRAYFUNC.EXE

Start Genetic Process

End



Probabilities of the Process

Crossover Probability

0.900

Mutation Probability

0.005

Continue

Type of Genetic Algorithm

ELITIST

Algoritmo Genético

Los individuos de la población inicial no necesariamente satisfacen las restricciones impuestas por el problema.

Si ninguna restricción fuese satisfecha, el *fitness* sería de “-1000,000,000”; si se satisficiera una de las restricciones sería “-75,000,000”, etc.

Función de Fitness

$$f(\vec{x}) = \begin{cases} x_1 + |\sin(32x_1) - \cos(11x_2)| & \text{si } 0 \leq x_1 \leq \Pi \\ & 0 \leq x_2 \leq \Pi \\ -10^9 + (25 \times 10^7)s & \text{en otro caso} \end{cases}$$

en donde s es el número de restricciones satisfechas y $0 \leq s \leq 4$

Algoritmo Genético

Nótese que, como estamos maximizando, lo que el algoritmo nos “dice” es que estas propuestas de solución son muy malas.

En la tabla que se ilustra a continuación aparece el número “-25,000,000” que nos indica que 3 de las 4 restricciones son satisfechas.

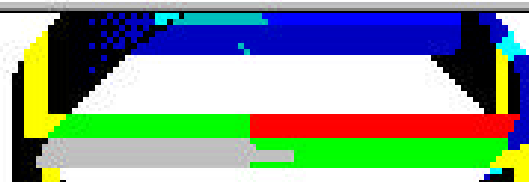


Results of the Genetic Process: Generation #3

Number	Individual	Population	Fitness
12	1	3	-250000000.000000000
11	2	3	-250000000.000000000
10	8	3	-250000000.000000000
9	16	3	-250000000.000000000
8	19	3	-250000000.000000000
7	31	3	-250000000.000000000
6	40	3	-250000000.000000000
5	49	3	-250000000.000000000
4	54	3	-250000000.000000000
3	66	3	-250000000.000000000
2	73	3	-250000000.000000000
1	96	3	-250000000.000000000
100	3	3	-500000000.000000000
99	4	3	-500000000.000000000

Type of Genetic Algorithm

ELITIST



"ESC" ENABLED

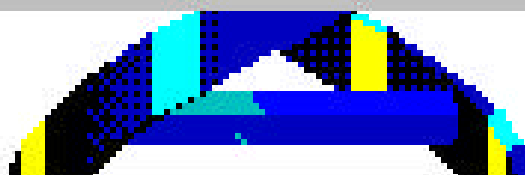
Ins

Algoritmo Genético

En la generación 8 aparece el primer individuo que satisface las 4 restricciones.

Esto induce un *fitness* aún lejano al óptimo pero claramente mejor que sus antecesores.

Como el algoritmo es elitista, siempre conserva al mejor individuo.

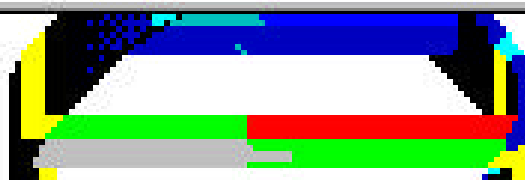


Results of the Genetic Process: Generation #8

Number	Individual	Population	Fitness
17	17	9	0.8171917293
100	100	9	-250000000.00000000
97	97	9	-250000000.00000000
96	96	9	-250000000.00000000
95	95	9	-250000000.00000000
94	94	9	-250000000.00000000
92	92	9	-250000000.00000000
89	89	9	-250000000.00000000
88	88	9	-250000000.00000000
87	87	9	-250000000.00000000
86	86	9	-250000000.00000000
85	85	9	-250000000.00000000
84	84	9	-250000000.00000000
83	83	9	-250000000.00000000

Type of Genetic Algorithm

ELITIST



Poblacn

Record: None

Record Unlocked

Ins

Inicio

Admin...

CALE...

Micro...

Micro...

GEN...

GE...

RUN...

System tray icons (volume, network, clock).

1:23 PM

Algoritmo Genético

Para la generación 15 la mayoría de las soluciones propuestas (individuos) arrojan *fitnesses* que satisface las restricciones.

Nótese que el mejor individuo es considerablemente mejor que en la tabla anterior y que, a medida que el AG procede, las soluciones son cada vez más grandes y cercanas al óptimo.

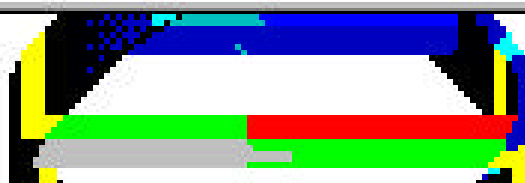


Results of the Genetic Process: Generation #15

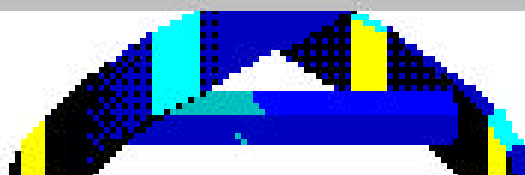
Number	Individual	Population	Fitness
41	41	16	2.3390775628
1	1	16	2.1712886520
33	33	16	1.9547522627
18	18	16	1.8512953267
40	40	16	1.5293874778
87	87	16	1.5293537105
78	78	16	1.4557058100
72	72	16	1.4456584709
77	77	16	1.3447633091
34	34	16	1.3414758006
75	75	16	1.3335761732
9	9	16	1.3255457715
17	17	16	1.2807845207
96	96	16	1.2636143918

Type of Genetic Algorithm

ELITIST



Ins [] []

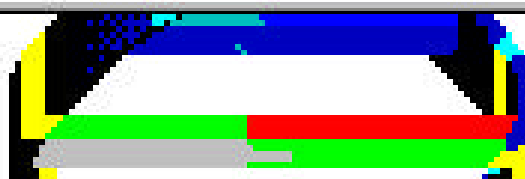


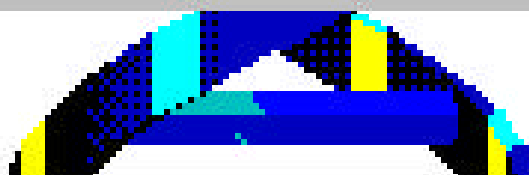
Results of the Genetic Process: Generation #21

Number	Individual	Population	Fitness
12	12	22	3.5540517202
1	1	22	3.5540517202
83	83	22	3.4793181877
85	85	22	2.8842320609
73	73	22	2.8220311868
15	15	22	2.7490822271
86	86	22	2.3463814198
74	74	22	2.3386295353
72	72	22	2.3285946582
76	76	22	2.1308106204
75	75	22	2.1308106204
36	36	22	2.1308106204
10	10	22	2.1308106204
13	13	22	2.1266033119

Type of Genetic Algorithm

ELITIST



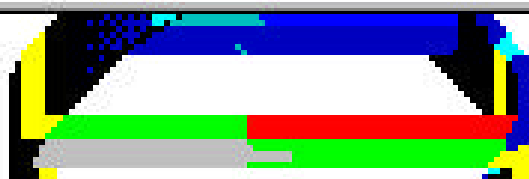


Results of the Genetic Process: Generation #32

Number	Individual	Population	Fitness
1	1	33	3.7399599919
41	41	33	3.6828059118
10	10	33	3.6130782562
64	64	33	3.6110760648
56	56	33	3.4052390621
40	40	33	3.0035284162
22	22	33	2.9685393799
24	24	33	2.8951989882
21	21	33	2.8940436688
93	93	33	2.8845280594
50	50	33	2.7486889074
58	58	33	2.7483197179
12	12	33	2.4089033009
34	34	33	2.4049648780

Type of Genetic Algorithm

ELITIST



Algoritmo Genético

En la siguiente figura puede observarse el *fitness* del mejor individuo al final de 50 generaciones.

También se muestra el genoma correspondiente.

 **C:\GALS2000\POBLACN.DBF**

F_recno	F_fitness	F_genoma
7	2.8579809776	Memo
8	1.2193982626	Memo
9	2.0336040162	Memo
10	4.0812366441	Memo
11	1.9776089273	Memo
12	1.2089124117	Memo
13	2.6067315942	Memo
14	1.2091950860	Memo
15	3.6289948594	Memo
16	1.9366997465	Memo
17	2.1627829514	Memo

Poblacn.f_genoma

000010100111011011101100011000000001001101101010100000010101

Algoritmo Genético

Ahora queremos interpretar los valores de cada una de las variables contenidas en el individuo.

Eso se logra como se muestra en la siguiente figura.

 **Interpret the Genome**

Select Results Directory

Results File: **RESULTS.DBF**

BCD

Gray

Length of the Genome

60

Number of Variables

2

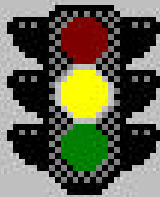
Number of Integers

5

Number of Decimals

24

See Optimized Function



See Conditions



See Variables

End

COMPUTER EXERCISES

ResultsRecord: 1/1Record Unlocked

InsNum

 Inicio

 Administrador d...

 CALENDARIO ...

 GENETIC O...

 Microsoft Power...



Algoritmo Genético

Al activar el botón “**See Variables**”, podemos ver:

- a) El máximo fitness reportado
- b) Los valores de las variables

Nótese que los valores:

- a) Satisfacen las restricciones
- b) Maximizan la función

GENETIC OPTIMIZATION OF FUNCTIONS WITH/WITHOUT RESTRICTIONS

Edit Data Base Tools Genetic Optimization Help Exit

Interpret the Genome

Results File: RESULTS.DBF

☐

☒

Length of the Genome	60
Number of Variables	2
Number of Integers	5
Number of Decimals	24

FITNESS: 4.0812366441	
Variable	Value
V01	3.0891427994
V02	1.8562026620

COMPUTER EXERCISE...

Variables

Record: 1/2

Exclusive

Ins

Num

Inicio

Administrador d...

CALENDARIO ...

GENETIC O...

Microsoft Power...

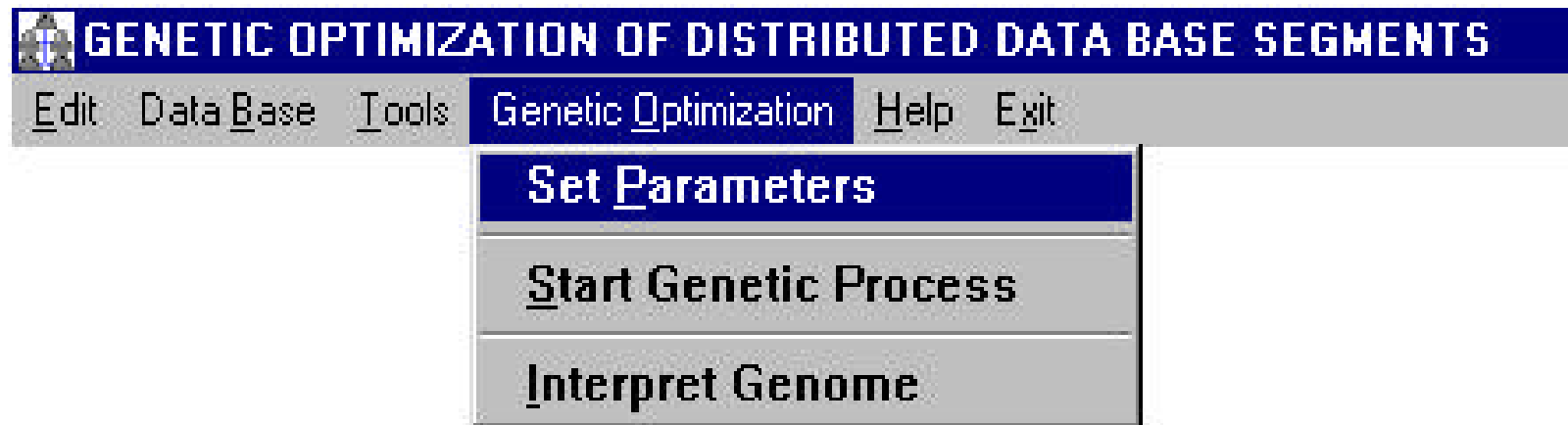
12:22 PM

Aplicaciones

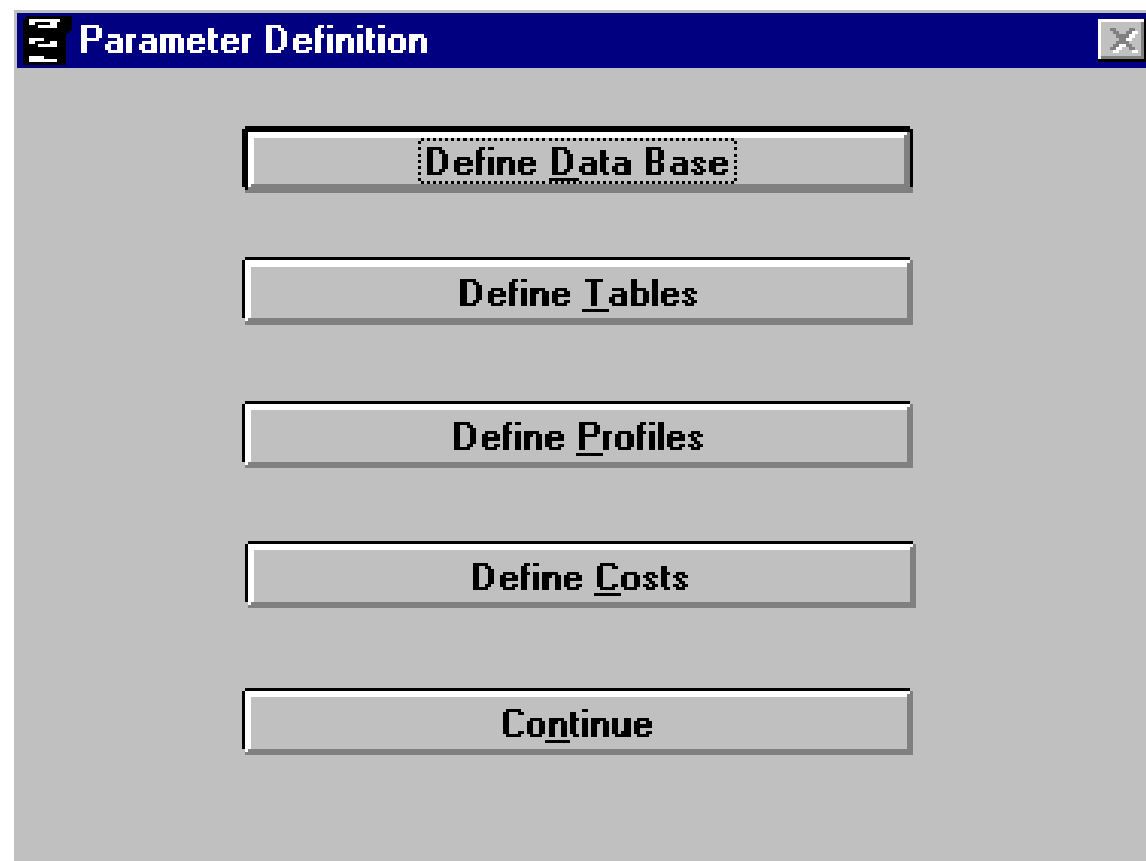
- **Segmentación Automática de Bases de Datos Distribuidas**

- » Segmentación horizontal
- » Segmentación vertical
- » Segmentación mixta

Bases de Datos Distribuidas



BDDs



Define Data Base



How many Tables?

2

Define Names



Control+F4 to continue



Number	Name	Size	
1	TAB1	1000	
2	TAB2	1000	

Continue

Define Tables

Select Table Number:

1

TAB1

How many Attributes?

2

Define Attributes

Control+F4 to continue		
Table	Number	Name
TAB1	1	AT1
TAB1	2	AT2

Continue

Define Profiles

How many Sites?

2

Define Profiles

Control+F4 to continue

"Site"	Table	Table #	Atrib.
1	TAB1	1	
1	TAB1	1	
1	TAB2	2	
1	TAB2	2	
2	TAB1	1	
2	TAB1	1	

Continue

C_PROFIL								
F_site	F_table	F_tabnum	F_atnumber	F_atname	F_atperc	F_atsize	F_tuple	
1	TAB1	1	1	AT1	0.2	20	Site01.001	
1	TAB1	1	2	AT2	0.3	30	Site01.001	
1	TAB2	2	1	AT1	0.1	40	Site01.002	
1	TAB2	2	2	AT2	0.3	50	Site01.002	
2	TAB1	1	1	AT1	0.1	60	Site02.001	
2	TAB1	1	2	AT2	0.1	70	Site02.001	
2	TAB2	2	1	AT1	0.1	80	Site02.002	
2	TAB2	2	2	AT2	0.2	90	Site02.002	

BDDs

Define Costs			
Site01	Site02		
0	100		

BDDs

Parameters for Genome

Minimum Number of Segments

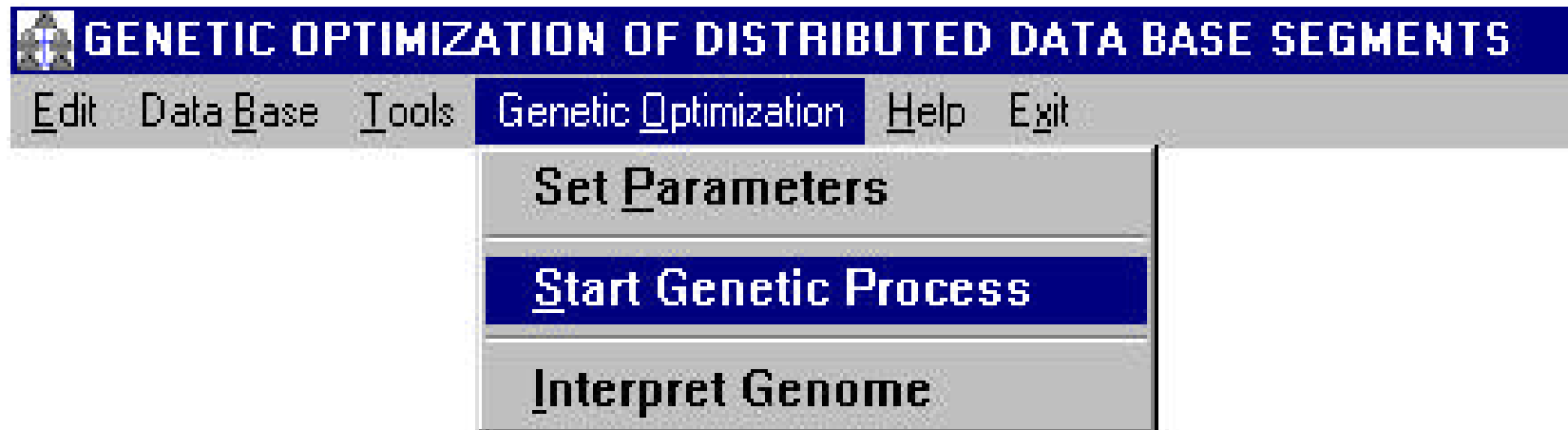
2

Maximum Number of Segments

8

Continue

BDDs



BDDs

Genetic Process Control

Lower Pc	9	# of Individuals	50
Upper Pc	10	# of Generations	25
Lower Pm	5	Evaluations	1250
Upper Pm	5	Genome	40
Lower # of Descendants	1	Maximize/Minimize	2
Upper # of Descendants	2	Windows/Dos	1
Lower HC Percentage	9	Data Format	1
Upper HC Percentage	10	Random Seed	3.0000000000

External Function

C:\FOX_APPS\SEGMENTS\BASE\SEG_FUNC.EXE

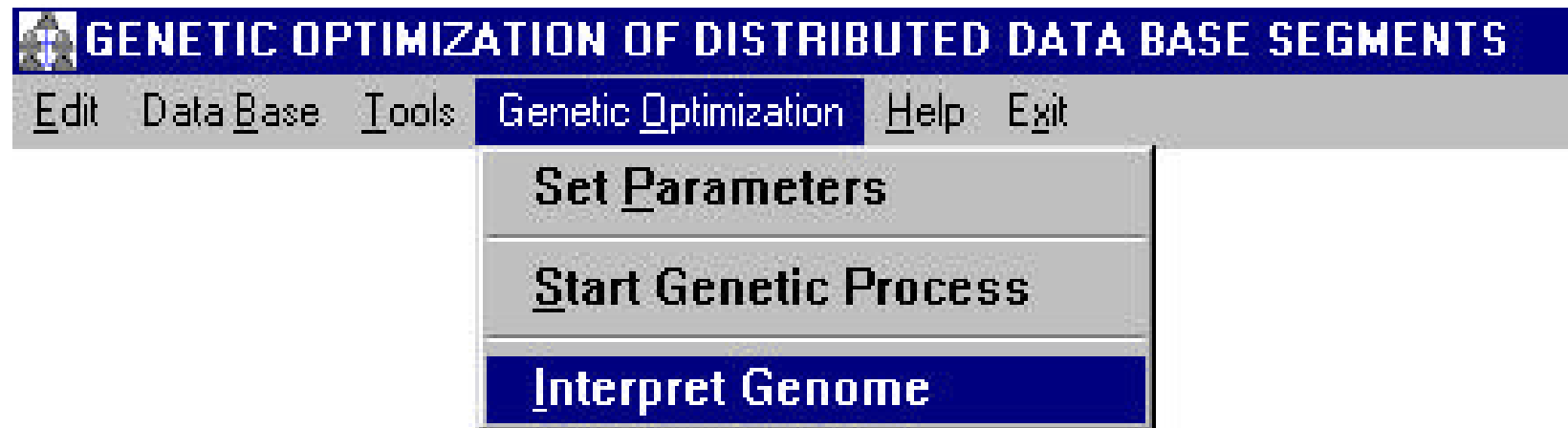
Start Genetic Process

End

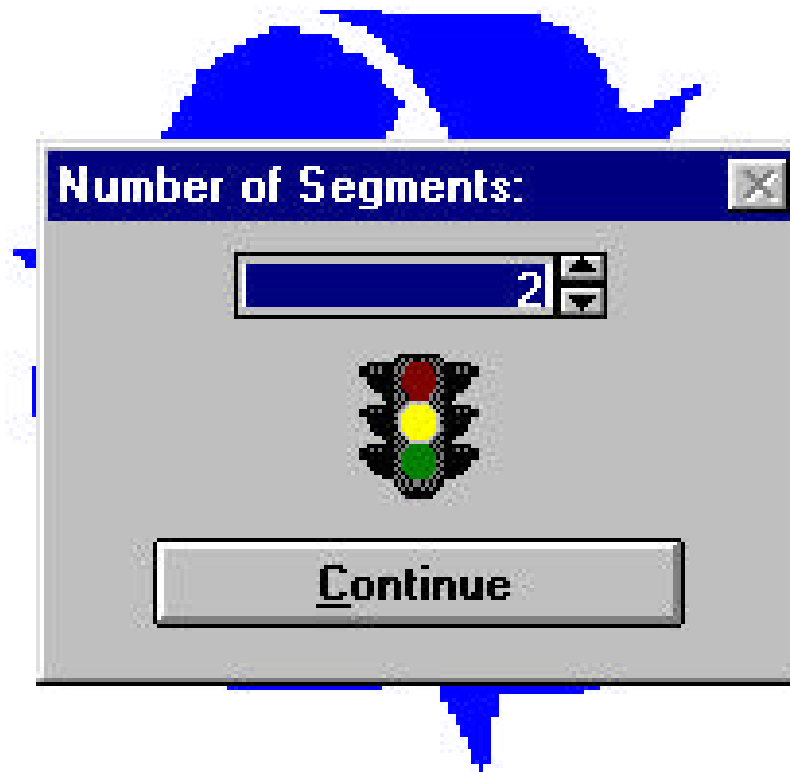
BDDs

- Ejemplificamos los resultados con un sistema un poco más complejo que el señalado en las láminas anteriores
- En lo que sigue, el número de “sites” es 3.

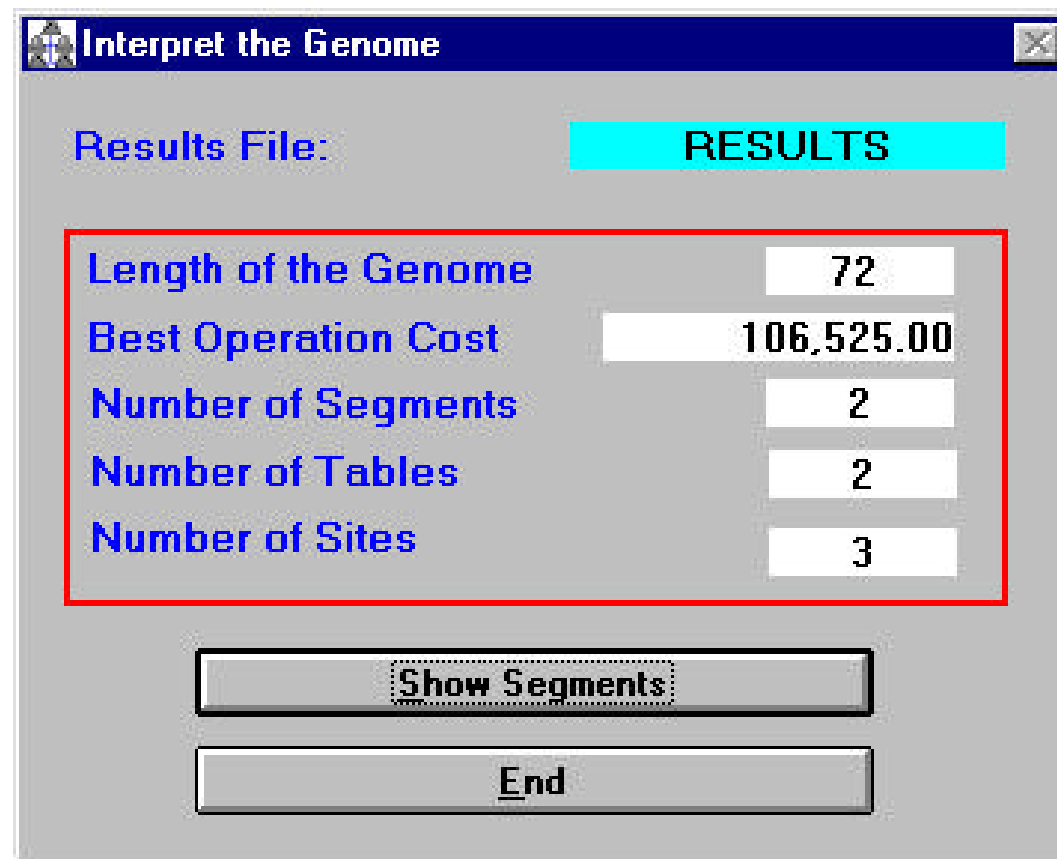
BDDs



BDDs



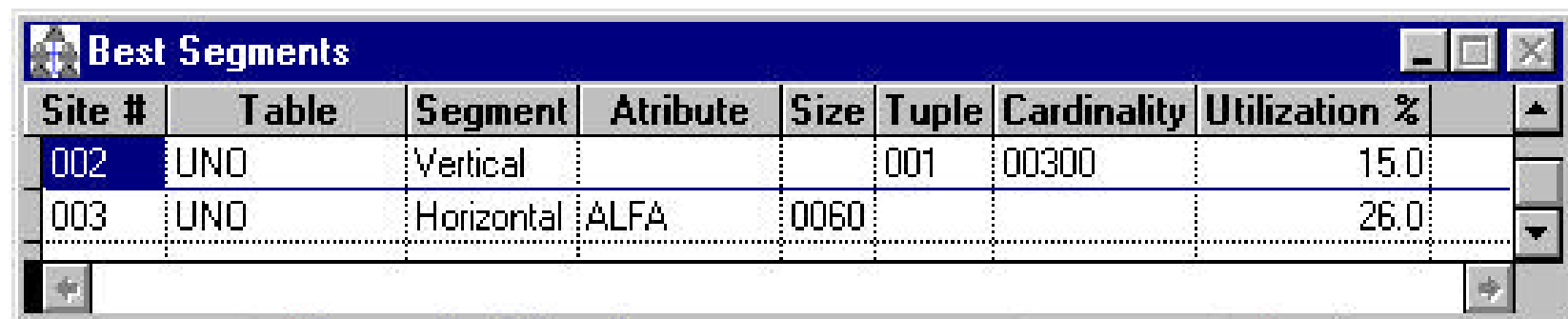
BDDs



The screenshot shows a software window titled "Interpret the Genome". Inside, there is a section labeled "Results File:" followed by a cyan button labeled "RESULTS". Below this, a red rectangular box highlights a table of results. The table has two columns: labels on the left and numerical values on the right. The labels are "Length of the Genome", "Best Operation Cost", "Number of Segments", "Number of Tables", and "Number of Sites". The corresponding values are 72, 106,525.00, 2, 2, and 3. Below the table, there are two buttons: "Show Segments" and "End".

Label	Value
Length of the Genome	72
Best Operation Cost	106,525.00
Number of Segments	2
Number of Tables	2
Number of Sites	3

BDDs



The screenshot shows a window titled "Best Segments" with a table containing the following data:

Site #	Table	Segment	Attribute	Size	Tuple	Cardinality	Utilization %
002	UNO	Vertical			001	00300	15.0
003	UNO	Horizontal	ALFA	0060			26.0

Conclusiones

- Los algoritmos genéticos son herramientas de amplia aplicación
- La metodología puede ser generalizada
- Su programación es fácil