

Genetic Algorithms (GAs)

- heuristic methods--do not guarantee optimality
- iterative methods
- simulate biological evolution, which has enabled "nature" to develop species remarkably welladapted to their environment.

Ľ

GAs

- operate on a population of *individuals* which represent potential solutions to the problem
- from the population, a set of "good" individuals are selected to *mate* & form a new *generation*
- the quality of subsequent generations will (it is hoped!) gradually improve and approach optimality

©Dennis Bricker, U. of Iowa, 1997

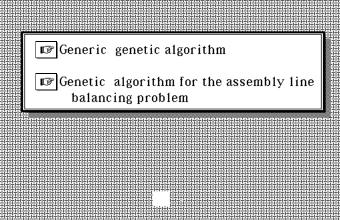
©Dennis Bricker, U. of Iowa, 1997

ENCODING

Each individual in the population consists of a string, which "resembles" a biological *chromosome* composed of *genes*.

These genes may take on values from a finite set of digits or characters, called *alleles*. (Most frequently, the alleles are binary digits.)

Example: from the set of alleles {A,B,C} we might produce the individual ABCABBCABBC





- development & theoretical foundation provided
 by:
 - John Holland, *Adaption in Natural and Artificial Systems,* The University of Michigan Press,1975.
- further developed & analyzed by Holland's student:

David Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning,* Addison-Wesley, 1989.

©Dennis Bricker, U. of Iowa, 1997

Encoding
Fitness measure
Selection
Scaling fitness
Crossover
Mutation
Generic GA Algorithm

©Dennis Bricker, U. of Iowa, 1997



For each individual i, we define a *fitness* value f(i) to measure how good it is.

Individuals representing good solutions will have a larger fitness value than those which correspond to poor solutions.

Used to determine which individuals *survive*: to produce the next generation.



"Survival of the fittest..." Opennis Bricker, U. of Iowa, 1997

SELECTION

From the population at each generation, a GA randomly selects the fittest individuals to survive and mate to produce the next generation.

The simplest method for doing this is "stochastic sampling with replacement", in which individual

i is selected with probability p_i

$$= \frac{f(1)}{\sum_{j=1}^{N} f(j)}$$

This is repeated until N individuals have been selected. ⊈_

©Dennis Bricker, U. of Iowa, 1997

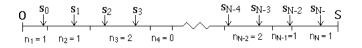


1. Compute S, the sum of the fitness values for the population.

2. Map each fitness value in random order to contiguous segments of the interval [0,S] on the real line, such that each segment has length equal to its corresponding fitness value.

0			∧S
f(1) f(2)	f(3)	f(4)	f(N-2) f(N-1) f(N)
Stochastic Universal Sampling			©Dennis Bricker, U. of Iowa, 1997

5. Let n_k be the number of values in the set $\{s_i\}$ which fall within interval k.

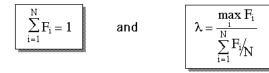


6. For each k (k=1,2,...N), create n_k copies of individual k and put into the mating pool.

©Dennis Bricker, U. of Iowa, 1997

A way to prevent this is to scale the fitness values so as to maintain a constant "best fitness to average fitness" ratio, the scaling factor.

Given scaling factor λ and a set of fitness values $\{f_1, f_2, \dots, f_N\}$, we scale them to $F_i = a f_i + b$ where a & b are chosen so that



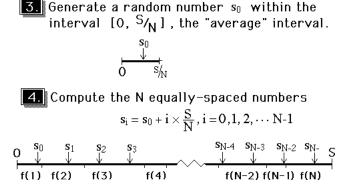
©Dennis Bricker, U. of Iowa, 1997

In sampling with replacement, the more fit individuals are likely to be selected, and may be selected several times. However, there is not quarantee that an individual whose fitness is above the population average will be selected

"Stochastic universal sampling" is a scheme which will guarantee this!

J.E. Baker, "Reducing bias and inefficiency in the selection algorithm", in Genetic Algorithms & Their Applications: Proceedings of the 2nd Int'l Conf. on Genetic Algorithms, J.J. Grefenstette (ed.), 1987.

©Dennis Bricker, U. of Iowa, 1997



©Dennis Bricker, U. of Iowa, 1997

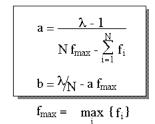
SCALING The distribution of fitness values may strongly affect premature convergence of a genetic algorithm.

If a few individuals have much larger fitness values than the others, they will generate many more offspring and quickly dominate the generations which follow, and the algorithm will prematurely converge to a population of identical "clones" which are not optimal.



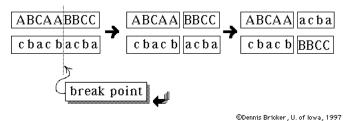
Solution: Choose a & b so that

where



CROSSOVER

Two individuals can mate to generate two offspring, by randomly selecting a break point and exchanging substrings:



The GA is initiated with a random population of N individuals



1. Evaluates the fitness of each individual in the population

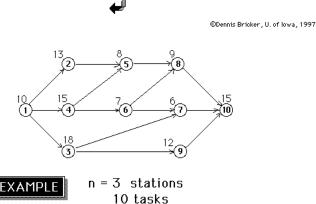
2. Based on fitness, chooses individuals from the population based on fitness, to form a *mating pool*, and pairs these individuals randomly.

©Dennis Bricker, U. of Iowa, 1997

There are numerous decisions which much be made in the design of a genetic algorithm:

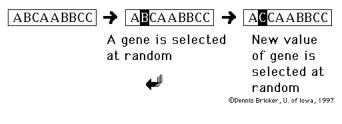
- population size
- probability of crossover
- probability of mutation
- whether to "seed" the initial population with "good" solutions
- scaling factor (best to average ratio) for fitness

See sample Pascal code in Chapter 3: "Computer Implementation of a Genetic Algorithm", in book by D. Goldberg.



MUTATION

With relatively low probability, a gene on a chromosome might change its value from one allele to another.



3. For each pair of individuals in the mating pool, *either* mates them & puts the two offspring into the new population *or* copies them directly into the new

copies them directly into the new population

4. Carries out mutation on the new population

5. Checks whether to terminate.

Usually a genetic algorithm will terminate after a given number of iterations, or when nearly all of the individuals in the population are identical.

©Dennis Bricker, U. of Iowa, 1997

GENETIC ALGORITHM FOR LINE BALANCING

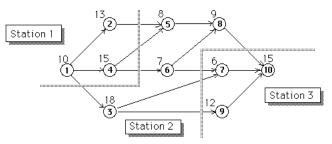
Consider the Assembly Line Balancing (ALB) Problem in which the number of stations n is fixed, and the tasks are to be assigned to the stations so as to minimize the cycle time.

4

©Dennis Bricker, U. of Iowa, 1997



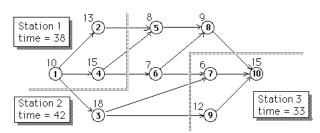
For the ALB problem, a natural encoding would be, using the station numbers as alleles, to identify the station to which task i is assigned by the number in the i^{th} position on the string.



The assignment shown above would be coded by the string: 11212333

ENCODING

©Dennis Bricker, U. of Iowa, 1997



The assignment of the ten tasks to the three stations shown above results in a cycle time equal to $T_{max} = max\{38, 42, 33\} = 42$.

©Dennis Bricker, U. of Iowa, 1997

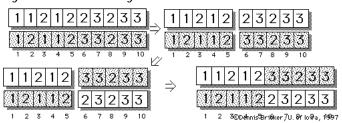
However, because individuals which violate one or more precedence restrictions might appear in the population, we will include a penalty P times the number V of such violations:

$$T$$
 (i) = T_{max} + $P \times V$

©Dennis Bricker, U. of Iowa, 1997

CROSSOVER

Suppose that the two individuals below have been paired for "mating", and the "break point" has been randomly chosen to be after the 5 "gene" in the string:



FITNESS

Our objective is to

- minimize the cycle time, while
- satisfy the precedence restrictions
- The "fitness" measure of a solution should be a positive value which reflects both of these considerations,
- with a larger value for those solutions which better satisfy the objective & restrictions.

©Dennis Bricker, U. of Iowa, 1997

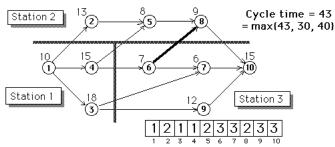
The cycle time is to be minimized, and so is an indication of "unfitness", i.e., the larger the cycle time, the more unfit the solution is.

If T(i) is the cycle time of individual i, then the fitness could be measured by

where
$$\label{eq:f(i) = k \times M - T(i)} \begin{bmatrix} f(i) = k \times M - T(i) \end{bmatrix}$$

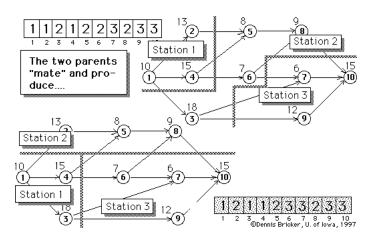
and k>1. With this definition, f(i)>0 for all i.

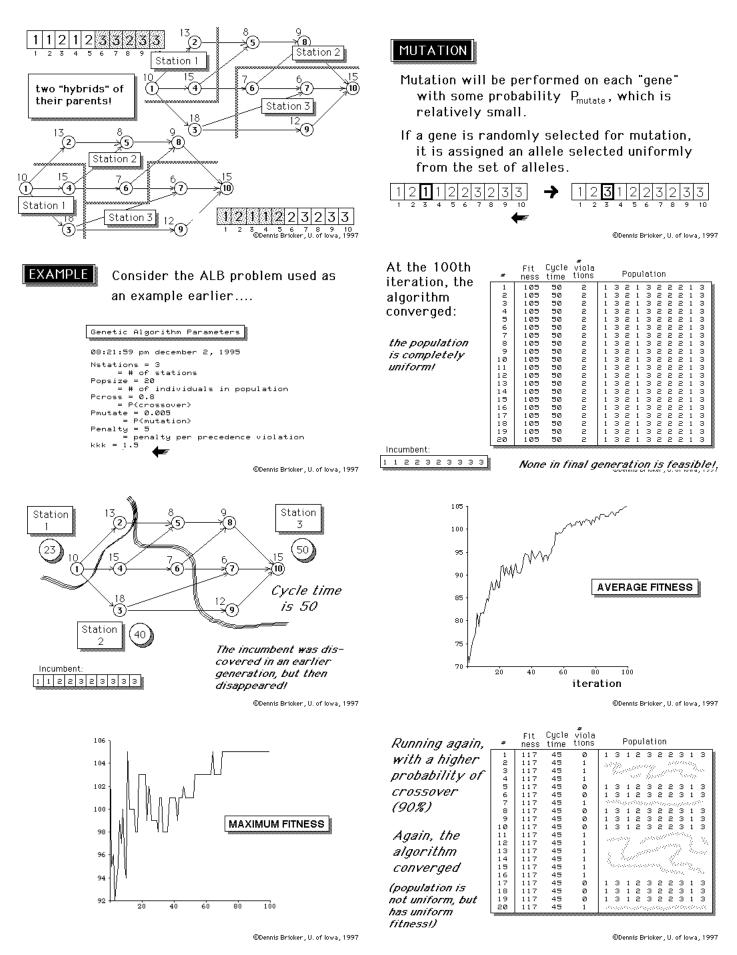
©Dennis Bricker, U. of Iowa, 1997

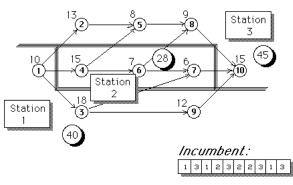


In the assignment shown, task 6 does not precede task 8 as required, and so we reduce the fitness by a penalty P=10:

$$\Gamma(i) = 43 + 10 \times 1 = 53$$

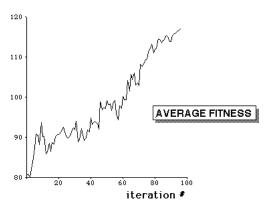




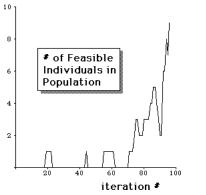


Cycle time is 45

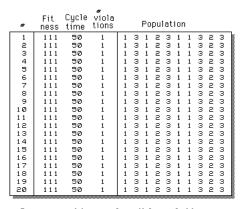
©Dennis Bricker, U. of Iowa, 1997



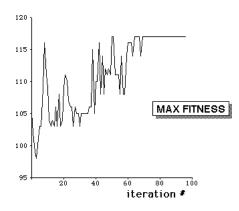
©Dennis Bricker, U. of Iowa, 1997



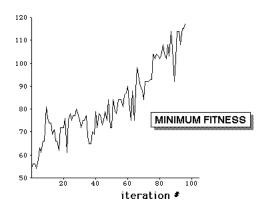
©Dennis Bricker, U. of Iowa, 1997



Converged to nonfeasible solution!

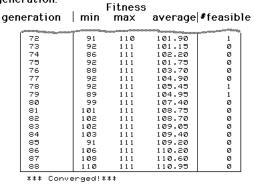


©Dennis Bricker, U. of Iowa, 1997

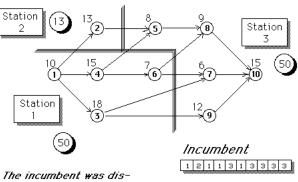


©Dennis Bricker, U. of Iowa, 1997

Running the GA again with Pcross=90%, it converged in 88th generation:

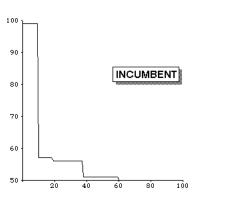


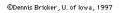
©Dennis Bricker, U. of Iowa, 1997

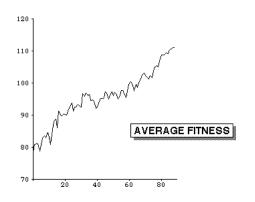


covered in an earlier generation, and then disappeared!

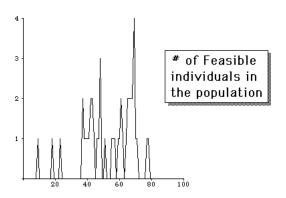
Cycle time is 50







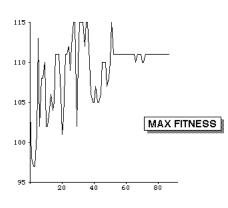
©Dennis Bricker, U. of Iowa, 1997

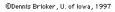


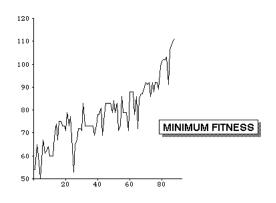
©Dennis Bricker, U. of Iowa, 1997

#	fitness	Cycle time	<pre># violations</pre>	F	°OP	UL	. A	ΤI	40	1		_,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	800			-		××****	000000				~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	\$xxxxx		
1	120	44	Ø	12	21	З	2	2	3 :	3 :	з 🛛	16	1	120	44	ø	1	2	2	1	з	1	2	з	з:	зТ
2	120	44	ø	12	21	з	2	2	з :	з :	з 🎚	17		120	44	ø	1	2	2	1	з	1	2	з	з:	з 📗
3	120	44	ø	12	21	з	2	2	з :	з :	з 🛛	18		120	44	ø	1	2	2	1	з	2	2	з	з:	з 🛛
4	120	44	ø	12	21	з	2	2	з :	з :	з 🛛	19		120	44	ø	1	2	2	1	з	2	2	з	з:	3
5	120	44	ø	12	21	з	2	2	з :	з :	з 🎚	20		120	44	ø	1	2	2	1	з	2	2	з	з:	з 📱
6	120	44	ø	12	21	з	2	2	з :	з :	з 🎚	21		120	44	ø	1	2	2	1	з	2	2	з	з:	3 3
7	120	44	ø	12	21	з	1	2	з :	3 :	з 🎚	22		120	44	ø	1	2	2	1	з	1	2	з		з 🛛
8	120	44	ø	12	21	з	2	2	з :	3 :	з 🎚	23		120	44	ø	1	2	2	1	з	2	2	з	з :	з 🛛
9	120	44	ø	12	21	з	2	2	3 :	3 :	з 🎚	24	1	120	44	ø	1	2	2	1	з	2	2	з	3 :	з 🛛
10	120	44	ø	12	21	з	2	2	3 :	-	з 🎚	25		120	44	ø	1	2	2	1	з	2	2	з	-	з 🛛
ι1	120	44	ø	12	21	з	2	2	3 :	-	з 📱	26		120	44	ø	1	2	2	1	з	2	2	з	-	3
12	120	44	ø	12	21	з	2	2	3 :	3 :	з 🛿	27		120	44	ø	1	2	2	1	з	2	2	з	3 :	з 📱
13	120	44	ø	12	21	з	2	2	3 :	3 :	з 📱	28		120	44	ø	1	2	2	1	з	2	2	з	-	
ι4	120	44	ø	12	21	з	1	2	-		з 🎚	29		120	44	ø	1	2	2	1	з	2	_	-		3
15	120	44	ø	12	21	з	1	2	з :	3 :	з 🎚	30		120	44	ø	1	2	2	1	з	1	2	з	з :	з 🛙
						_			_	_			-				_	_	_	_	_	_			_	

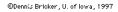
©Dennis Bricker, U. of Iowa, 1997



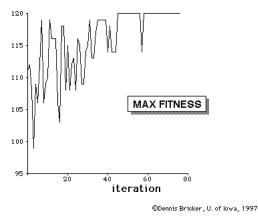


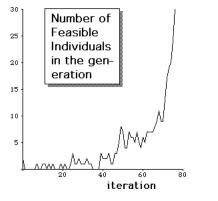


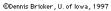
gen	eration	min	Fitness max	average	#feasible
- · · · ·		and the second			
Running the	60	83	120	106.63	4
-	61	83	120	105.60	6
GA again,but	62	83	120	104.47	5
e e	63	90	120	105.63	7
with popula-	64	81	120	106.20	7
	65	75	120	104.50	7
tion size =	66	84	120	106.03	7
70 1.00	67	83	120	106.10	8
30, not 20	68	86	120	109.17	9
-	69	97	120	111.10	11
	70	93	120	113.90	9
	71	98	120	114.47	9
	72	98	120	114.83	13
Converse in	73	98	120	116.63	17
Converges in	74	102	120	117.17	19
7644	75	115	120	119.50	20
76th genera-	76	119	120	119.80	24
tion	Converged	!***			

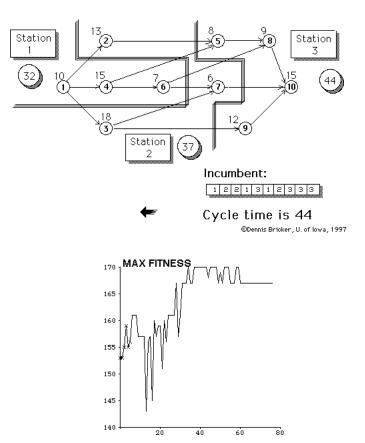


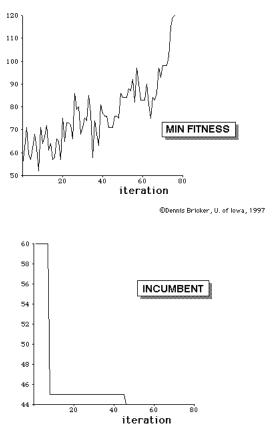








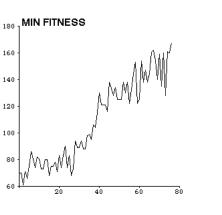




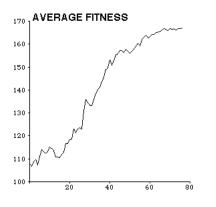
©Dennis Bricker, U. of Iowa, 1997

Finally, the GA was run with a population size of 50, and a probability of crossover equal to 75%.

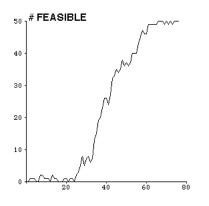
Convergence occurred on the 77th iteration.



©Dennis Bricker, U. of Iowa, 1997







©Dennis Bricker, U. of Iowa, 1997

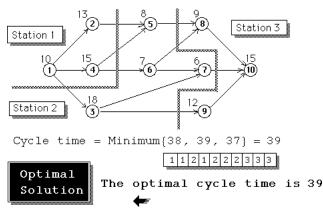
All solutions in the final population are feasible, with fitness = 167 and cycle time = 43

There	are	fοι	ır			
distinct						
soluti	ions	in	the			
population:						

	1	2	1	1	З	2	З	З	2	З
:	1	2	1	1	2	2	З	З	2	З
[1	2	1	1	2	2	З	З	З	З
	1	2	1	1	2	З	З	З	2	З

The best solution was found in the 35th generation, with a cycle time equal to 40.

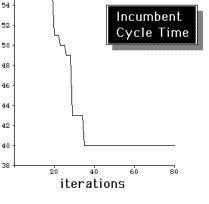
©Dennis Bricker, U. of Iowa, 1997



©Dennis Bricker, U. of Iowa, 1997

▽GA;detail;toprinter;Max_iterations;I;detail;sho

[1]	A	
[2]	A	Genetic Algorithm
C33	A	for Assembly Line Balancing Problem
C43	A	namely, minimize Cycle Time, s.t. #stations
[5]	A	
C63	A	Global variables used:
E73	A	Nstations = # of stations to be used
[8]	A	P = vector of task times
[9]	A	A = adjacency matrix of precedenc
C10]	A	Popsize = population size
C113	A	kkk = scalar > 1
[12]	A	Pcross = probability of crossover
C133	A	Pmutate = probability of mutation
C143	A	



©Dennis Bricker, U. of Iowa, 1997

APL Code

> The following screens contain listings of some APL functions written to implement GA for the Assembly Line Balancing problem.



- C15] show_pop+detail+ASKYN 'Show details?'
- C16J C17J \$ detail/'show_pop+ASKYN ''Show pools?''
- Max_iterationsRequest_Box 'max # iterations' toprinter+AskPrint & PrSelect
- C18]
- C193 Prologue
- C20J A
- . POP+Initial_Pool(Popsize,Nstations,ρP) CycleTime+Evaluate_Pool POP C21]
- F221
- Fitness+(Fkkk×UB+F/CycleTime)-CycleTime [23]
- Incumbent+10 <> MinCycleTime+BIG [24]
- [25]
- [26] Record_Status
- C273 'Random Seed: ', Set_Random_Seed

Genetic Algorithm for Line Balancing

7/15/98

[28]	A	[39]	8
[29]	\$(~detail)/'OL ◊ UNDERLINE ''-GenerationMinimu	C403	POP+(
	easible-'''	C413	A
	Next:→Stop ∆if Max_iterations <iteration←iteration< td=""><td>E423</td><td>A</td></iteration←iteration<>	E423	A
[31]	A	E433	POP←N
[35]	→Converged Δif(F/Fitness)=L/Fitness	E443	A
[33]	A	C453	Fitne
C343	A Select pool of individuals who will survive	E463	Recor
[35]	A & be candidates for mating (& mutation)	E473	A
C36]	I←Select_Pool Fitness	E483	→Ne×t
E373	Report_GA	E493	'***
		[50]	→Stop

- E381 e
- Mate pairs in the selected pool -Crossover POP[I;]
- Mutate individuals in the pool
- Mutate POPEI;]
- ess←(Fkkk×UB)-CycleTime←Evaluate_Pool POP
 - rd_Status
 - End of iteration
 - Maximum # of iterations performed! ***'
 - top
 - [51] Converged: '*** Converged!***'
 - E521 A
- [53] Stop: '*** terminated ***'

©Dennis Bricker, U. of Iowa, 1997

©Dennis Bricker, U. of Iowa, 1997

POOL←Initial_Pool X;msg;rl

- C13 A [2] Get initial population for genetic algorith A. C33 A C43 X[1]= population size A X[2]= # of stations X[3]= # of tasks [5] A E63 A C73 A →Random △if ASKYN 'Random initial genetic pool?' [8] [9] Fnter:PODL+±30 Alpha_Request_Box 'name of initial [10] →0K Δif^/(X[1]=1↑pPODL),(X[3]=⁻¹↑pPODL),,PODLE:X [11] Message_Box '.Error!.Please try again' [12] →Enter [13] Random:rl+Set_Random_Seed [14] POOL↔?X[1 3]pX[2]
- C153 OK:

©Dennis Bricker, U. of Iowa, 1997

▼Report_GA;table;M;output

- [1] A [2] A
- Report results in iteration of Genetic Algorithm for ALB problem
- [3] A [4] A [5] . ≰detail/'SHIFT≏CTR UNDERLINE ''Generation #'',₹iteration
- M+1↑pPOP C6 3
- [6] M+1↑pPOP [7] +(*show_pop)/Summary [8] table+'I2,< | >,I6,< | >' DFMT(\q(2,M)p(\1M),Fitness) [9] table+table,'I2' DFMT POP [10] DL,' # Fitness ---mating pool----' [11] table ◊ PAUSE ◊ →End

- [11] table o FHOSE o 7=00 [12] Summary: [13] output+(iteration,(~1↑MinVals),(~1↑MaxVals),(~1↑AvgVals) [14] 'I7,2110,F10.2,I10' DFMT 1 5poutput
- [15] End:

©Dennis Bricker, U. of Iowa, 1997

⊽Record_Status;b;i

- С1Ј А C2J A Record status of the population
- [3] at the end of each generation A
- Е4Ј А
- MinVals+MinVals,L/Fitness [5]
- MaxVals+MaxVals,F/Fitness AvgVals+AvgVals,AVERAGE Fitness E63
- C73
- [8] NumFeas+NumFeas,+/0=nv+Number_Violations POP
- [9] \rightarrow End $\triangle if^{\vee}/b \leftarrow (n \vee = 0)^{CycleTime}(-1^{MinCycleTime})$
- [10] MinCycleTime+MinCycleTime,CycleTimeEi+1↑b/lPopsize
- Incumbent←POP[i;] C113
- C120
- [13] End:MinCycleTime←MinCycleTime, T1↑MinCycleTime

©Dennis Bricker, U. of Iowa, 1997

⊽t←Ctime pop

- C1J A [2] Compute Cycle time for solution "pop" A C33 where pop[i]= station to which task i is A C43 assigned. Include penalty times number A [5] A of precedence violations. E63 А C73 Global variable: Penalty A [8] P = vector of task times A C93 A C103 $t {\leftarrow} \Gamma / (({\tt lNstations}) {\bullet} . {=} {\tt pop}) {+} . {\tt xP} \qquad {\tt A Cycle time}$
- [11] t+t++/Penaltyx,A>pop•. \leq pop A Violation penalty

- 623 A Compute number of violations of
- E3] A precedence restrictions
- E43 e
- C53 A Global variable: A = precedence matrix
- →Single ∆if 1=pppop E63
- E73 z+(1↑ppop)p0 ◊ i+0
- [8] Next:→End △if(ppop)[1]<i+i+1
- [9] z[i]++/,A>pop[i;]•.≤pop[i;] [10] →Ne×t
- [11] Single:z++/,A>pop•.≤pop
- 0

©Dennis Bricker, U. of Iowa, 1997

- ⊽z←E∨aluate_Pool POP;i;M C1J A
- C2] A Evaluate the individuals in the pool
- C31 A (cycle time of assignment)
- [4] А
- i+0 ♦ M+(pPOP)E13 ♦ z+Mp0 [5] [6] Next:→End △if M<i←i+1
- z[i]←Ctime POP[i;] E73
- [8] →Ne×t
- E93 End:

©Dennis Bricker, U. of Iowa, 1997

▽z←Number_Violations pop;i

- [12] End:

E1J A

⊽I←Select_Pool F;S;N;s0;s;f;n;a;b

C 1]	A
[2]	A Select individuals from pool according
C33	A to their fitness values F
C43	A Fitness values are scaled so as
[5]	A to sum to 1.0 and
C63	A Max Fitness = scalefactor÷N
C73	A Global variable: GA_Scale_Factor
[8]	a←(GA_Scale_Factor-1)÷((PopsizexΓ/F)-+/F)
[9]	b+(-axL/F)F(GA_Scale_Factor÷Popsize)-axF/F
C10]	F+b+a×F
C11]	A
[12]	S←+/F ◊ f←+∖F
C133	s0←(S÷N←ρF)×0.001×?1000
C143	s←s0+(S÷N)×0,ıN−1
C153	n←n−0,¯1↓n←+/(f∘.≥s)
E161	$\rightarrow 0$ Aif Popsizes((p)0)/p)\(p)0)/1N

[16] $\rightarrow 0 \text{ } Aif Popsize=pI \leftarrow ((n>0)/n) \setminus (n>0)/1N$

⊽Q←Sample Fitness;N;S;sØ;s;F

- C1J A Select individuals from the population, C2J A
- [3] A
- who will survive to the next generation. Argument: Fitness = vector of fitness values C4] A
- C53 A of the individuals in the population,
- Result: Q = vector of number of copies of individuals to be included in the next ge E6J A
- С73 А E8] A

Method used is "stochastic universal samplin C93 A

- C103 A due to J.E. Baker
- С11Ј А [12] S++/Fitness \diamond N+ ρ Fitness \diamond F++Fitness
- [13] s0+(?1000)×N÷1000×S
- s←sØ+(S÷N)×0,ıN-1 Г147
- Q←(+/F•.2≤)-(+/(0,-1↓F)•.2≤) C153

©Dennis Bricker, U. of Iowa, 1997

⊽Q←Crossover Pop;M;N;PAIR;I [1] A For each pair of individuals [2] A in the mating pool, either ma in the mating pool, either mate [2] A in the method [3] A them & put the 2 offspring into [4] A the pool, or else copy them into [5] A the new pool directly [6] N+(ρPop)[1] ◊ M+(ρPop)[2] [7] Q+(0,-1↑ρPop)ρ0 [8] A Randomly shuffle the pool [7] P--FN2N+]

- [10] Next:→End △if 2>(pPop)[1]
- С11Ј А Remove pair from the pool
- [12] PAIR+(2,M)↑Pop ♦ Pop+2 Ø↓Pop
- E133 A

- Liss n [14] →Copy Aif Peross((?1000)+1000 [15] A Choose crossover point [16] Q+Q,[1]((2,I)↑PAIR),¢[1](0,I+1+?M-2)↓PAIR [17] →Next [17] →Next
- [18] Copy:Q←Q,[1]PAIR ◇ →Ne×t

[19] End:Q←Q,[1]Pop

©Dennis Bricker, U. of Iowa, 1997

©Dennis Bricker, U. of Iowa, 1997

⊽Q←Mutate Pop;I

- С1Ј А C2] A Perform mutation of the individuals in pool
- C3] A
- [4] Q←,Pop [5] I←(Pmutate≥(?(ρQ)ρ10000)+10000)/ιρQ
- Q[I]←?(pI)pNstations [6]
- C73 Q+(pPop)pQ

┢┲╴