UCD CASL

Complex & Adaptive Systems Laboratory

Dr. Michael O'Neill
Dr. Miguel Nicolau

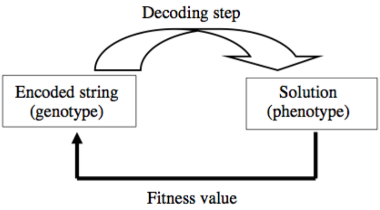
Genetic Algorithms

COMP30290 Natural Computing
COMP41190 Natural Computing and Applications

Genetic Algorithm

Overview

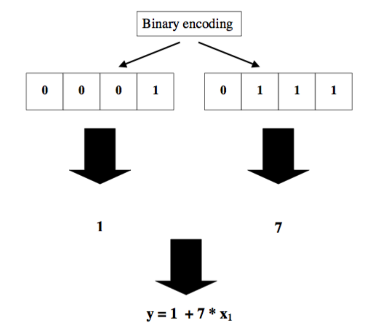
- Holland (1975), Goldberg (1989);
- Binary String individuals;
- Evolutionary search operates on encoding of solution (genotype);
- Decoding: genotype-phenotype map.



Genetic Algorithm

Representation

- Fixed-length chromosome;
- Each locus 1 bit;
- Fixed-size genes;
- Encode reals, ints.



Genetic Algorithm

Encoding

- Integers (binary vs. gray):
  - n bits encode 2^{n-1} ints.
- Reals:

$$x = \frac{\text{decoded integer}}{2^{n-1}}$$

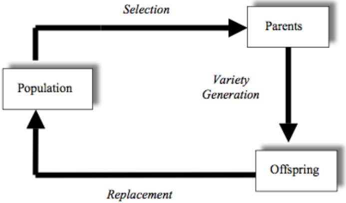
- Interval $a \longrightarrow b$:

$$z = a + x(b - a)$$

Integer Value	Binary Code	Gray Code
0	000	000
1	001	001
2	010	011
3	011	010
4	100	110
5	101	111
6	110	101
7	111	100

- Beware of hamming cliffs!
- small phenotype change requires large genotype change
- 3 to 4 requires 3-bit changes
- Gray 1-bit change

Evolutionary Algorithm

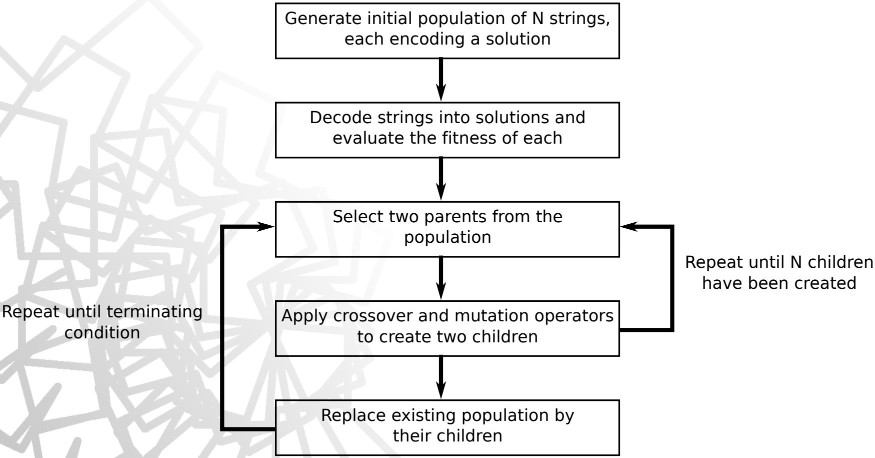


$$x[t+1] = r(v(s(x[t])))$$

Pseudo-code

```
Initialise Population;
While (termination condition FALSE):
   select Parents;
   create Offspring;
   Update Population;
EndWhile
```

Genetic Algorithm



Example Problem: ONEMAX

Definition

- **Maximise numbers of 1s in a bit string of length** n;
- Example: $n = 8$, popsize $= 4$, ...

1. Generate initial population:
   - Randomly assign 1 or 0 to each locus.
2. Calculate fitness:
   - Count number of 1s.

Candidate	String	Fitness
A	00000110	2
B	11101110	6
C	00100000	1
D	00110100	3

Example Problem: ONEMAX

Fitness

- Could normalise and standardise fitness:
  - Normalise between 0.0 and 1.0;
  - Standardise 1.0 is best, 0.0 is worst.

Candidate	String	Fitness	Normalised Fitness
A	00000110	2	0.25
B	11101110	6	0.75
C	00100000	1	0.125
D	00110100	3	0.375

Example Problem: ONEMAX

Selection

- Fitness proportionate selection:
  - Let f_i = fitness of individual i;
  - Average population fitness:

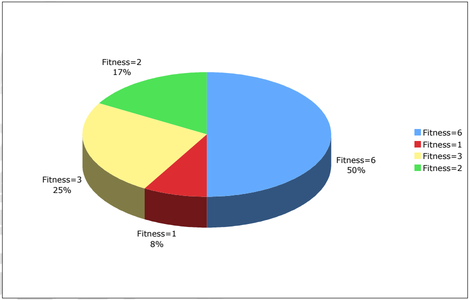
$$\bar{f} = \frac{1}{n}\sum_{i=1}^{N} f_i$$

  - Individual j selected with probability:

$$p_j = \frac{f_j}{\sum_{i=1}^{N} f_i}$$

Candidate	String	Fitness	Normalised Fitness	p_j
A	00000110	2	0.25	.17
B	11101110	6	0.75	.50
C	00100000	1	0.125	.08
D	00110100	3	0.375	.25

Example Problem: ONEMAX



$$r \in \left[0, \sum_{i=1}^{N} f_i\right) \qquad \sum_{i=1}^{j-1} f_i \leq r < \sum_{i=j}^{N} f_i$$

Example Problem: ONEMAX

Variation

3. Select two parents:

Candidate B	Candidate C
11101110	00100000

4. Crossover:

Candidate E	Candidate F
01101110	10100000

5. Mutation:

Candidate E	Candidate F
01001110	10100000

Example Problem: ONEMAX

Variation

3. Select two parents:

Candidate B	Candidate D
11101110	00110100

4. Crossover:

Candidate G	Candidate H
00111110	11100100

5. Mutation:

Candidate G	Candidate H
10111110	11000100

Example Problem: ONEMAX

Replacement

- Generational Replacement Strategy:

6. Replace parents with offspring:

Candidate	String	Fitness
A	00000110	2
B	11101110	6
C	00100000	1
D	00110100	3

Candidate	String	Fitness
E	01001110	4
F	10100000	2
G	10111110	6
H	11000100	3

7. Unless termination criteria met, go to step 3.

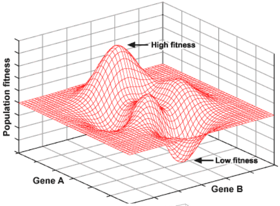
Exploration vs. Exploitation

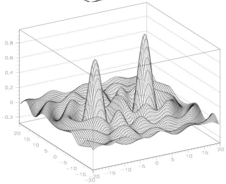
Black art of EC

- Careful choice of **Selection**, **Variation** and **Replacement** operators.
- Rate of convergence;
- Local Optima.

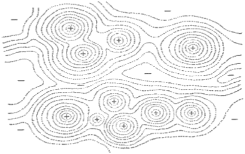
Exploration vs. Exploitation

Adaptive Landscape





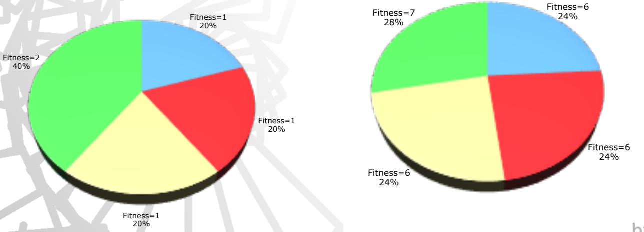
- Wright (1932)



Selection

Roulette Wheel Selection (Fitness Proportionate Selection)

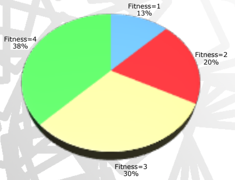
- High selection pressure earlier on;
- Premature convergence;
- Low selection pressure later:
  - Similar fitness values;
  - Uniform probability of selection;



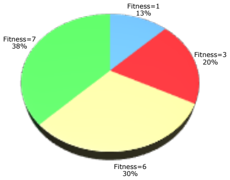
Selection

Rank Selection

- Rank from worst to best and calculate rescaled fitness;
- Linear ranking: $f_{\text{rank}} = 2 - P + 2 \times (P - 1) \times \frac{(\text{rank}-1)}{(n-1)}$
- Non-proportional selection.



Ranking	1	2	3	4
Fitness	1	2	3	4
f_{rank}	0.5	0.8	1.2	1.5
p_j	0.125	0.2	0.3	0.375



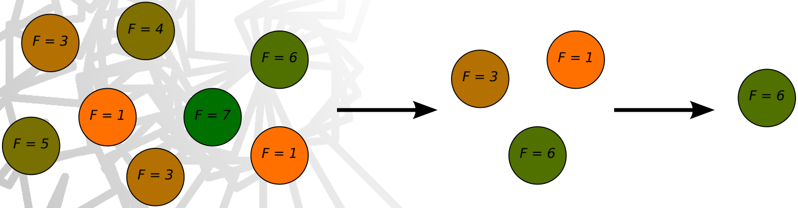
Ranking	1	2	3	4
Fitness	1	3	6	7
f_{rank}	0.5	0.8	1.2	1.5
p_j	0.125	0.2	0.3	0.375

P=1.5

Selection

Tournament Selection

- Select t individuals at random;
- Best of t individuals becomes parent.
- Selection pressure easily adjustable ($t \in [1..N]$);
- Can force fair tournament and unique parents.



Variation

Crossover and Mutation

- Mutation introduces novelty:
  - Too little. . . stuck in local optima;
  - Too much. . . random search!
- Crossover should exploit (share) good subsolutions;
- Exploration/exploitation balance;
- Adaptive mutation.

Variation

Alternative Crossover

- 2-point Crossover:

Parent 1	Parent 2
11101110	00110100

Offspring 1	Offspring 2
11110110	00101100

- Uniform Crossover:

Parent 1	Parent 2
11101110	00110100

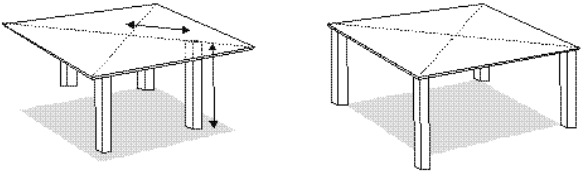
Offspring 1	Offspring 2
11100100	00111110

Replacement

Alternative Replacement

- Generational:
  - Children replace parents;
- Elitism:
  - Keep best fitness individual(s);
  - Generational for remainder.
- Steady-state:
  - Sort parents and offspring ($2N$), choose N best;
  - Can apply at variation operator level:
    - Two parents produce two children;
    - Best of four individuals make it to offspring population.

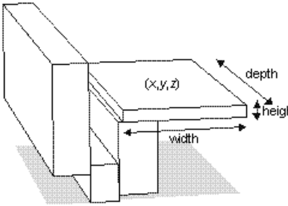
Example Encodings

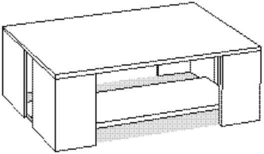


Table

- Consisting of fixed top and four legs defined by:
  - Length of leg 1, distance of leg 1 from centre;
  - Length of leg 2, distance of leg 2 from centre;
  - Length of leg 3, distance of leg 3 from centre;
  - Length of leg 4, distance of leg 4 from centre.
- Genotype: 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000 (l1, d1, l2, d2, l3, d3, l4, d4)

Example Encodings





Table

- Consisting of several 3-dimensional blocks:
  - x_1, y_1, z_1, width$_1$, height$_1$, depth$_1$;
  - x_2, y_2, z_2, width$_2$, height$_2$, depth$_2$;
  - ...
- Genotype: 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
  (x_1 y_1 z_1 width$_1$ height$_1$ depth$_1$ ...)

Example Encodings

Polar Bear

Colour	Legs	Soles	Forepaws	Length	Weight
White	4	Furred	30.4cm	2.2m	785.4kg
Category	Integer	Boolean	Float	Float	Float
000	0100	1	0101010011001101	11001010	1101011001011100

$$x = \frac{\text{decoded integer}}{2^{n-1}} \qquad a \rightarrow b \qquad z = a + x(b - a)$$

Example Encodings

Tokaido-Sanyo Shinkansen N7000

Assessing Performance

Monitoring

- Never draw conclusions from a single run;
- Use sufficient number of runs ($R > 30$);
- Use statistical measures (averages, medians, std. dev. , etc);
- Record as much data from your population as possible:
  - Mean, Best, Worst Fitness at each generation;
  - Diversity (genotypes, phenotypes);
  - Graph progress of these.
- Use controls:
  - Compare to equivalent Random Search.

Exercises

Examples to try

- Implement Onemax;
  - What are the best parameters to solve Onemax of size n?
  - With roulette selection, what is the error threshold for mutation, when using no crossover?

HINT: Goldberg's Simple GA in C (SGA) is available from the module website (+ app.onemax.c)

GA Literature

In Library

- Holland (1975): Adaptation in Natural and Artificial Systems;
- Goldberg (1989): Genetic Algorithms in Search, Optimization and Machine Learning;
- Mitchell (1996): An Introduction to Genetic Algorithms;
- Fogel (2000): Evolutionary Computation.

Next Classes

- Lecture Thursday 19th September 15h - 16h (GP with Miguel);