Application of Genetic Algorithms:

Sample Applications and Methodologies

Tom Kiehl (kiehl@crd.ge.com)
http://www.cs.rpi.edu/~kiehlt

Crypto GA

A CryptoQuote is a simple substitution code where each letter that appears may stand for a different letter.

The substitutions are consistent throughout the puzzle. Punctuation is not translated. For example: POLYYONTOP = RENSSELAER

So, how do we approach this?

[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Encoding

- 26 different possible letters => 26 Integer Alleles
- No two letters can convert to the same letter => 26 Unique allele values
- The allele positions tell us what each letter must convert to.

[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Fitness

Many different approaches:

- Total Number of real words found in solution.
  - The GA had a tendency to solve the short words in the problem and leave larger words untouched.
- For Each real word, calculate the percentage (lengthwise) of the entire string which this word represents. Sum all of the real words' percentages.
  - This fitness function worked somewhat better the previous one, but still provided little pressure to attain large words.

[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Fitness

Many different approaches:

- A very complicated scheme in which each word in the solution contributes to the overall fitness based on how connected the word is to other words. This measure also took into account whether or not the GA had solved some of the words that the word was connected to.
  - We hard-coded some values for one particular problem and than ran just that one. It was obvious that the added value (almost nil) was not worth the programming effort to make this work.
- Sum of the squares of the lengths of all of the real words.
  - Amazingly enough, this seems to work best, against all of our intuitions about how much information we were providing the GA.

[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Fitness

Many different approaches:

- Extra points for single letter words. This was used in conjunction with the sum of squares method.
  - This accelerated the GA through some of the early stages of deciding what the single character words should be. Moved the population into a state where most had a's and i's in place fo the single character words

[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Fitness

The Dictionary's Role:

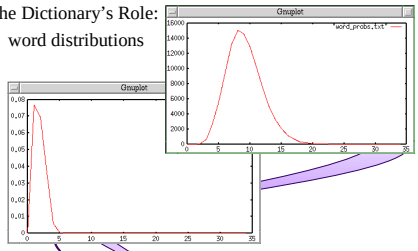
- Tested with two different dictionaries
  - /usr/dict/words 25143 words
    - contained individual letters
    - contained abbreviations
    - contained common names
    - contained only root words (ed, s, ing, etc were missing)
  - an NL dictionary 104217 words
    - addressed the above problems and provided more words
- Significantly better performance with NL dictionary

[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Fitness

The Dictionary's Role:

- word distributions



[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Selection

Worked great even with a uniform random selection mechanism.

We found that this problem was just filled with local maxima, so the random nature of the selection mechanism gave much better coverage of the search space.

[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Population

- Size/Generations:
  - We ran with a large population size of 400, which was necessary to maintain diversity. Much smaller when using speciation.
- Initialization:
  - Each individual had to adhere to the constraints of the problem, having unique allele values.
  - Later we seeded the initial population with individuals which already had solutions for large words. This gave us a significant boost in performance

[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Mutation

- Swap Mutator
  - Had to maintain constraints
  - Real simple, just caused a swap of values between two positions on the chromosome

[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Crossover

- Partial Match Crossover
  - Had to maintain constraints

[encoding][fitness][selection][population][mutation][crossover]

Crypto GA: Vs. Hill Climbing

- Hill Climbing Observations
  - Most of the time the hill climbing methods couldn't move from one point to another in the space
  - even if the hill climbing method got a "valid" answer, it wasn't ever correct.
- GA Comparisons
  - GA maintained a set of valid individuals, final result actually correct.
  - Searched the space much more effectively

Crypto GA: Results

- Iuisbh supab wsxosib rt wsrwfs lar ens nbybunbuon fus. --Eizirli

Gen #50 Fitness:109 Best:

** Biblty liwqt rlmclbt je rljrnl vqj xsl statistics nil. --Xbdbjvb

Gen #100 Fitness:145 Best:
** Ninety eiwbt lenzent be lebile kbb xse statistics fie. --Xeabdkn
Gen #150 Fitness:185 Best:
** Ninety eifbt percent mq pempze jbm zse statistics sie. --Zekzmtjn
Gen #200 Fitness:194 Best:
** Ninety eiwrt percent wh pewple zvw qse statistics lie. --Qebmwzn
Gen #250 Fitness:207 Best:
** Ninety eimnt percent of people vxn use statistics die. --Unznnwn
Gen #300 Fitness:230 Best:
** Ninety eigft percent ox people qfo kse statistics lie. --Kobnnqn
Gen #350 Fitness:239 Best:
** Ninety eimt percent ok people quo vse statistics lie. --Vobnnqn
Gen #400 Fitness:243 Best:
** Ninety eimnt percent ob people kno use statistics lie. --Unznnokn
Gen #450 Fitness:255 Best:
** Ninety eight percent ob people who for statistics lie. --Fnznnoxn

Gen #1100 Fitness:301 Best:

** Ninety eiblt percent ox people who use statistics lie. --Unknown

Gen #1300 Fitness:326 Best:

** Ninety eight percent of people who use statistics lie. --Unknown

Crypto GA: Results

- Xul eohp bp kpwnlip O'b n iwulagvpe. Mfpvp nvpa'm paulzf iwulagvpei oa xulv eoyp. -Fnvvoiua Yuvg, Mfp Pbcovp Imvohpi Knwh

Gen #50 Diversity:0.861545 Fitness:81 Convergence:1 Best:

** Cob fkzs hs lsdibqs K'h i qdobrwesf. Muses iesr'm srobyu qdobrwesfq ke esbe fkgs. -Uieekqre Goew, Mus Shakes Qmekzsq Lidz

Gen #100 Diversity:0.813886 Fitness:102 Convergence:0.882353 Best:
** Qlgt rioe de pomqpe I'd a zoikgcomer. Therne aren't enlglth zokgcomerz in qlgm rive. -Hannnizkio Vkmx, The Edwine Zomieo Ruin
Gen #150 Diversity:0.797175 Fitness:106 Convergence:1 Best:
** Lam rots cv vsatinlu O'r i kazmcyere. Buses iesr'h xsarmpu kzamcyerk ox lazne rewx. -Uieeokax Wary, Bus Scopes Kbesnk Vxst
Gen #200 Diversity:0.797005 Fitness:122 Convergence:0.92623 Best:
** Qlfg kipe me aeyagpe I'm a syfgnlesk. There aren't enlglth xylgnlesks in qlgr kibe. -Harrisdn Blfrd, The Empire Strypes Zayp
Gen #300 Diversity:0.779562 Fitness:163 Convergence:1 Best:
** Qlfg kipe me aeyagpe I'm a syfgnlesk. There aren't enlglth xylgnlesks in qlgr kibe. -Harrisdn Blfrd, The Empire Strypes Zayp
Gen #350 Diversity:0.808347 Fitness:163 Convergence:1 Best:

** Qlfg kipe me aeyagpe I'm a syfgnlesk. There aren't enlglth xylgnlesks in qlgr kibe. -Harrisdn Blfrd, The Empire Strypes Zayp

Gen #2550 Diversity:0.34339 Fitness:532 Convergence:1 Best:

** You like me because I'm a scoundrel. Tgere aren't enouzg scoundrels in your life. -Garrison Ford, The Empire Strikes Back

Gen #2600 Diversity:0.345111 Fitness:538 Convergence:1 Best:

** You like me because I'm a scoundrel. There aren't enough scoundrels in your life. -Harrison Ford, The Empire Strikes Back

The End

Genetic Algorithms for Autonomous Satellite Control

- The Problem
- Representation
- Constraints: static/dynamic
- Operators
- Results
- Comparison w/ Linear Programming

Genetic Algorithms for Autonomous Satellite Control

- Coverages

Sat.	City	Instance	Start	End	Priority
1	Beijing	1	27.01	55.06	2
1	Beijing	2	157.86	186.3	2
1	Beijing	3	289.76	311.11	3
1	Beijing	4	828.41	893.21	3
1	Beijing	5	999.66	1027.08	2
1	Beijing	6	1129.25	1157.81	2
1	Beijing	7	1263.34	1289.86	3
1	Beijing	8	1397.66	1424.56	2
1	Bombay	1	159.03	173.91	3
1	Bombay	2	287.46	314.55	2
1	Bombay	3	417.50	445.26	1
1	Bombay	4	555.91	559.96	2
1	Bombay	5	1116.15	1143.41	2
1	Bombay	6	1246.2	1273.75	2
1	Bombay	7	1385.6	1401.85	3
1	Buenos Aires	1	95.01	123.75	3
1	Buenos Aires	2	226.01	251.86	2
1	Buenos Aires	3	934.91	960.91	2
1	Buenos Aires	4	1063.16	1091.86	3
1	Buenos Aires	5	1197.4	1223.05	3
1	Buenos Aires	6	1356.76	1357.93	3
1	Calcutta	1	28.31	40.01	3
1	Calcutta	2	139.05	183.15	2
1	Calcutta	3	286.51	337.25	2
1	Calcutta	4	420.8	443.61	2
1	Calcutta	5	993.3	1014.69	2
1	Calcutta	6	1108.9	1147.56	2
1	Calcutta	7	1250.16	1277.42	2
1	Calcutta	8	1394.61	1407.4	3

Genetic Algorithms for Autonomous Satellite Control

Task Types

- **Battery Offline** : Executed when a battery must be put offline (1)
- **Battery Recondition** : Forces a full reconditioning of a battery. This event must be scheduled to occur on a somewhat periodic basis. Batteries require full reconditioning after spending a period of time being intermittently charged and discharged (2)
- **Momentum Dump** : A fairly short task which allows a satellite to maintain it's orbit (3)
- **Keep Station** : this task is also related to the mechanics of maintaining orbit (3)
- **Verify Link** : This test the satellite's ability to communicate with ground stations.(3)
- **Calibrate** : General maintenance (3)
- **State of Health** : send a measure of the satellite's state to a ground station (4)
- **Self Destruct**: Burn up satellite in atmosphere

GE Corporate Research & Development
Information Technology Laboratory

Genetic Algorithms for Autonomous Satellite Control

Typical Task Set

param : Tasks : dur start_after finish_by :=

1	BatteryRecond	1 GS_Any	50	0	1440
1	Calibrate	1 GS_Any	0.34	0	1440
1	KeepStation	1 GS_Any	16	0	1440
1	StateOfHealth	1 GS_Any	10	0	720
1	StateOfHealth	2 GS_Any	10	360	1080
1	StateOfHealth	3 GS_Any	10	720	1440

GE Corporate Research & Development
Information Technology Laboratory

Genetic Algorithms for Autonomous Satellite Control

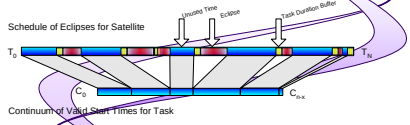
Task Interactions

Type	City	Ground Station	Eclipse	I/w types 1,2,3	I/w type 4
1	Shrinks	Shrinks	No Interaction	No Overlap	No Interaction
2	Shrinks	No Interaction	No Overlap	No Overlap	No Interaction
3	No Interaction	Within	No Interaction	No Overlap	No Interaction
4	No Interaction	Within	No Interaction	No Interaction	No Overlap

GE Corporate Research & Development
Information Technology Laboratory

Genetic Algorithms for Autonomous Satellite Control

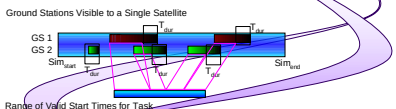
- Interaction of tasks with Eclipses



GE Corporate Research & Development
Information Technology Laboratory

Genetic Algorithms for Autonomous Satellite Control

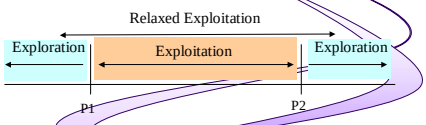
- Ground Station Tasks



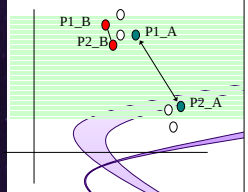
GE Corporate Research & Development
Information Technology Laboratory

Genetic Algorithms for Autonomous Satellite Control

- Crossover of Real Numbers



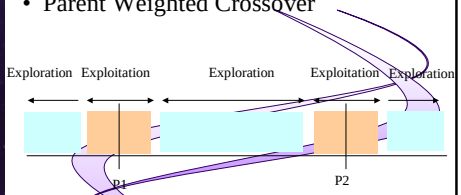
Genetic Algorithms for Autonomous Satellite Control



Assume that P1_A and P2_A are a large distance apart, while P1_B and P2_B are quite close to one another.
Choosing a point between P1_A and P2_A might be viewed as being more exploratory and less as exploitation.
Meanwhile, a point chosen between P1_B and P2_B could be viewed as a great amount of exploitation and very little exploration.

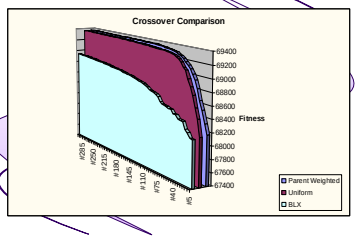
Genetic Algorithms for Autonomous Satellite Control

- Parent Weighted Crossover

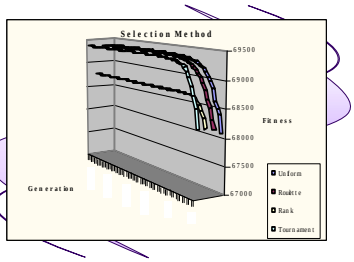


Genetic Algorithms for Autonomous Satellite Control

- Crossover of Real Numbers



Genetic Algorithms for Autonomous Satellite Control

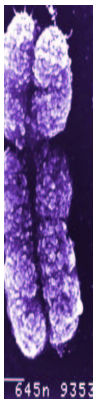


Genetic Algorithms for Autonomous Satellite Control

- Fitness Function
- Penalty Function

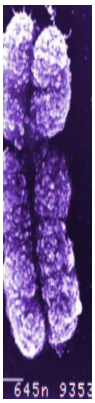
Genetic Algorithms for Autonomous Satellite Control

- The Problem
- Representation
- Constraints: static/dynamic
- Operators
- Results
- Comparison w/ Linear Programming



Chem GA

- Explored Uncharted Chemical Space
- 8 Alleles positions
  - Four chemicals (of a particular set) and their concentrations
- Fitness Function was a physical experiment.
  - A generation of 50 individuals took two weeks



Satellite Gen 2: Circuit Allocation

- Giant Decision Tree
- Determines the allocation of circuits for a satellite/city interaction
- Has been very difficult for the GA