A Multi-objective Genetic Algorithm

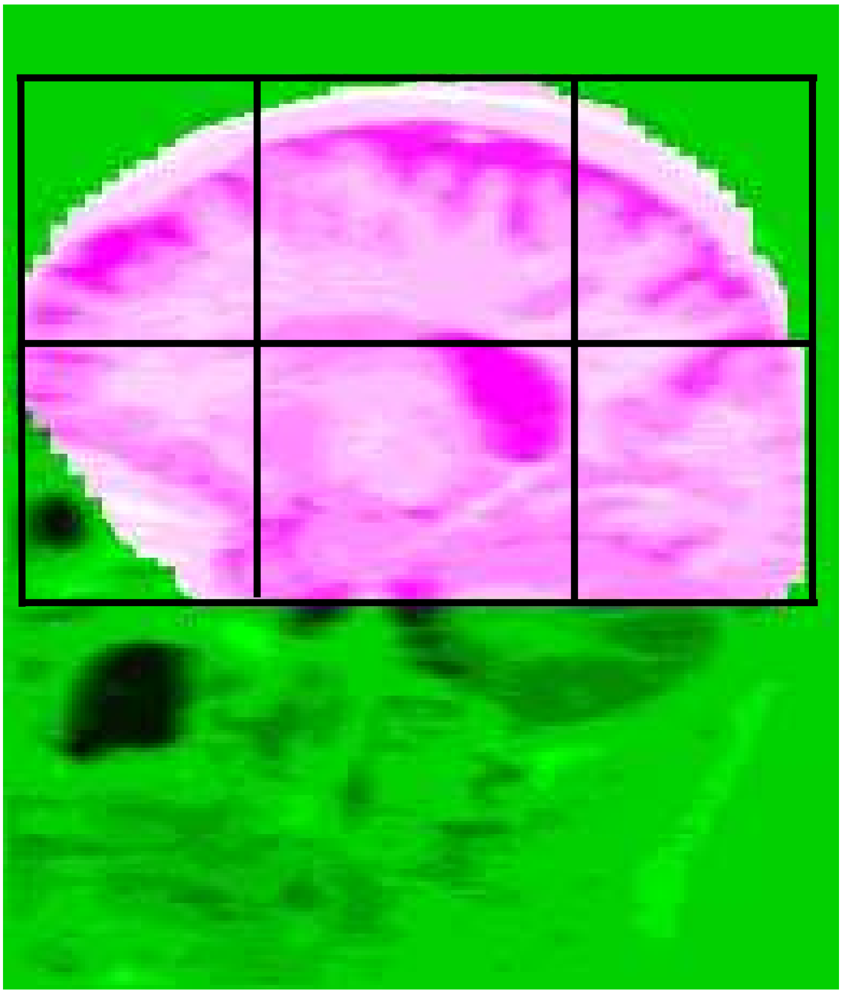
Dr. Paul A. Bromiley
Imaging Science and Biomedical Engineering

Overview

- Diagnosis of dementing diseases using the distribution of cerebral atrophy
- Optimisation of the technique using a genetic algorithm
  - preselection and static extraction
  - novel mating restriction
  - self adaptive mutation parameters
  - a multi-objective cost function test suite
- Current state of the project

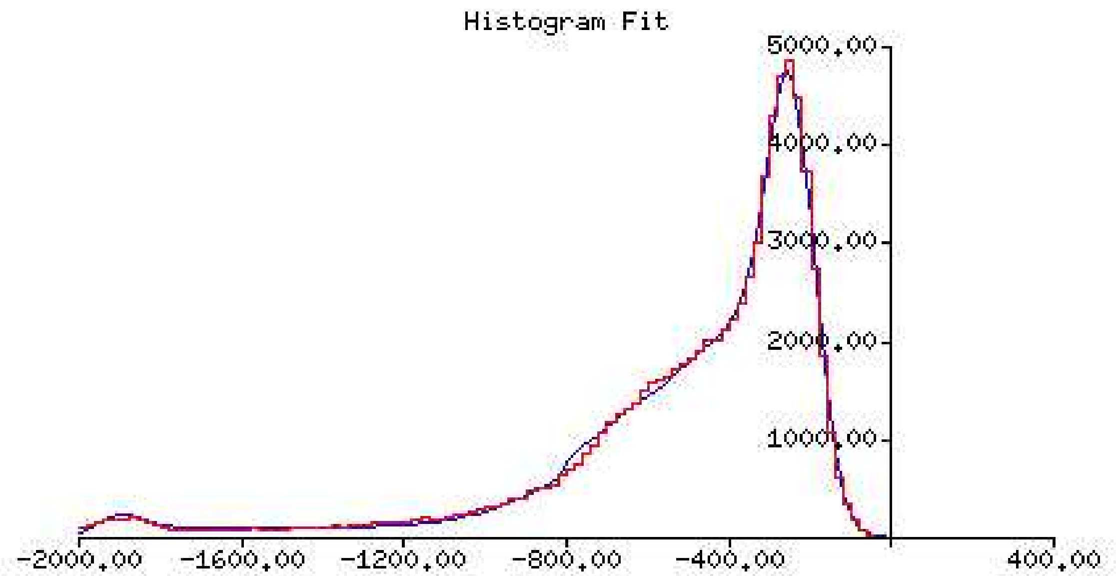
Atrophy Analysis: Introduction

- Original technique developed by Thacker et. al. (2002) Radiology 224 p278-285
  - hypothesis: dementing diseases result in specific patterns of cerebral atrophy
  - use minimal parameter atrophy measurement
  - co-register to a standard coordinate system
  - divide head into twelve equi-sized boxes



Atrophy Analysis: CSF Segmentation

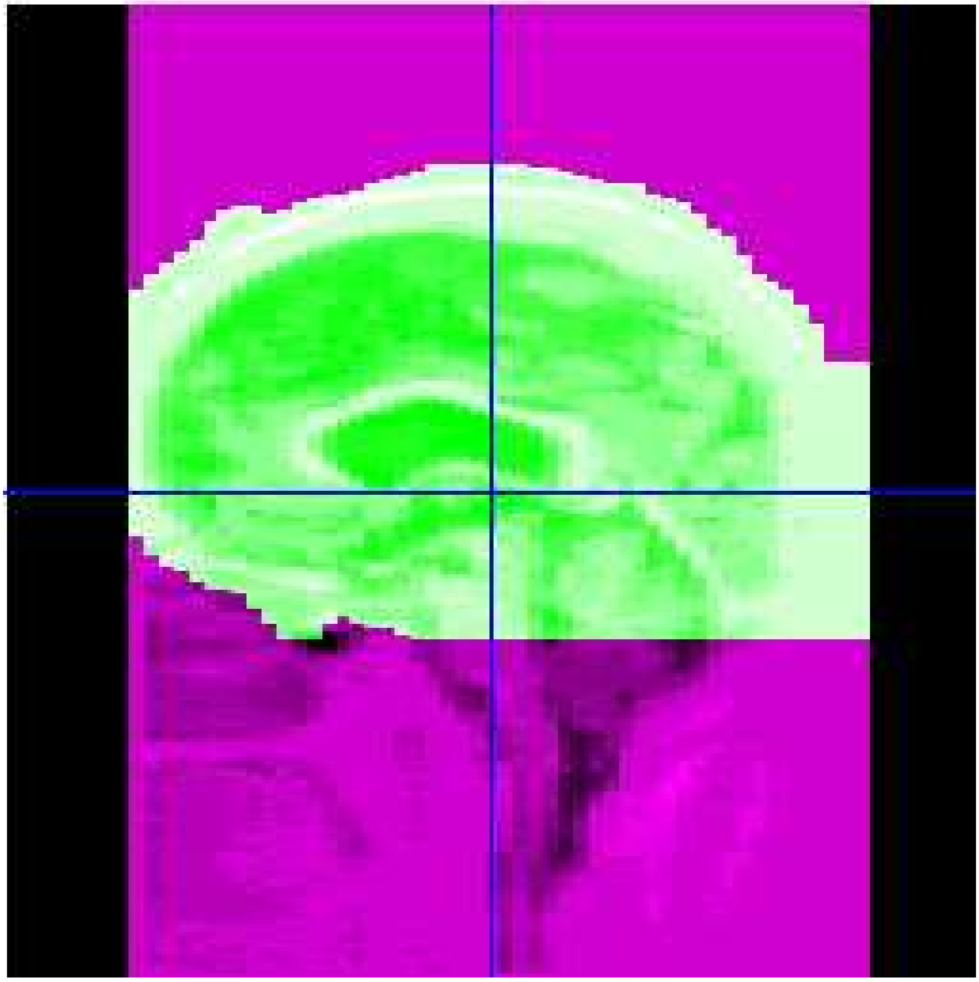
- Segment the CSF
  - use T2 IR scans
  - plot histogram of interior of head
  - fit with Gaussians + triangular distributions convolved with Gaussians



  - place threshold at CSF probability = 0.5
  - produce binary CSF maps

Atrophy Analysis: Masking

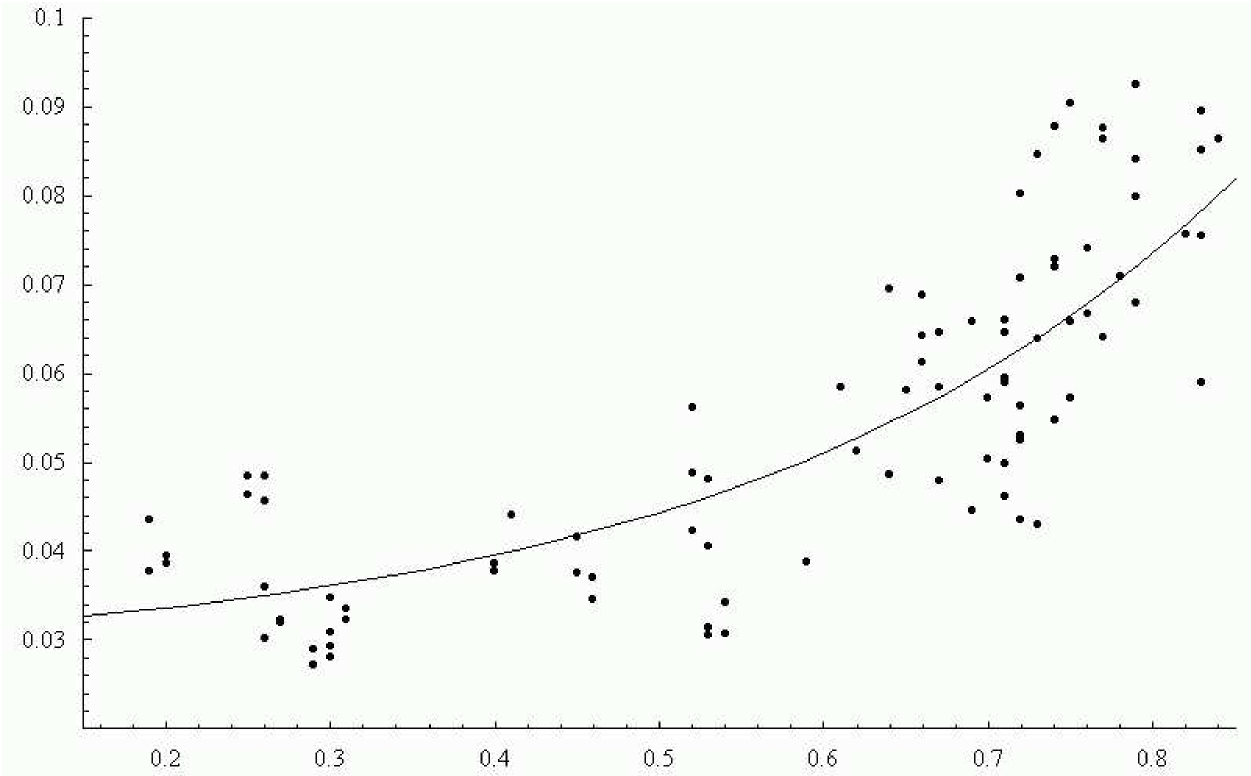
- Binary images include all fluid spaces
  - CSF, eyes, sinuses
- Delete eyes, sinuses etc. using binary masks



- Masks also define lower extent of the space: all other limits defined by outside of head

Atrophy Analysis: Age Correction

- Count CSF in the twelve boxes
- Normalise for head size (divide by total volume)
- Normalise for age-related atrophy



- Age normalisation
  - 94 normals: age range 19 to 85
  - fit curves

$$volume = a + b\ exp^{\ c\ age}$$

  - normalise to mean age point

Atrophy Analysis: Variable Reduction

- Calculate reduced variables
  - F = sum front four volumes
  - M = sum middle four volumes
  - B = sum back four volumes
  - P = sum left six volumes
  - S = sum right six volumes
  - U = sum top six volumes
  - L = sum bottom six volumes
- Calculate relative variables

$$W_1 = \frac{\sqrt{M} - \sqrt{F}}{\sqrt{2}}$$

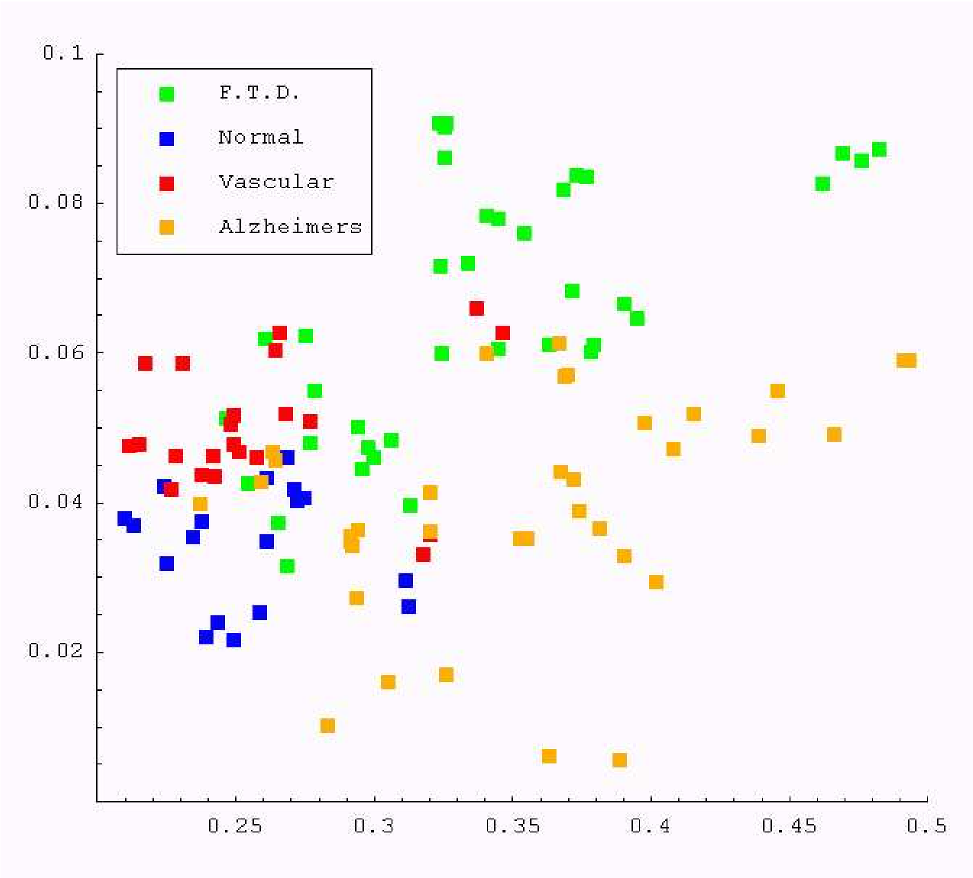
$$W_2 = \frac{\sqrt{M} - \sqrt{B}}{\sqrt{2}}$$

$$W_3 = \frac{\sqrt{F} + \sqrt{M} + \sqrt{B}}{\sqrt{3}}$$

$$W_4 = \frac{\sqrt{P} - \sqrt{S}}{\sqrt{2}}$$

$$W_5 = \frac{\sqrt{U} - \sqrt{L}}{\sqrt{2}}$$

Atrophy Analysis: Results



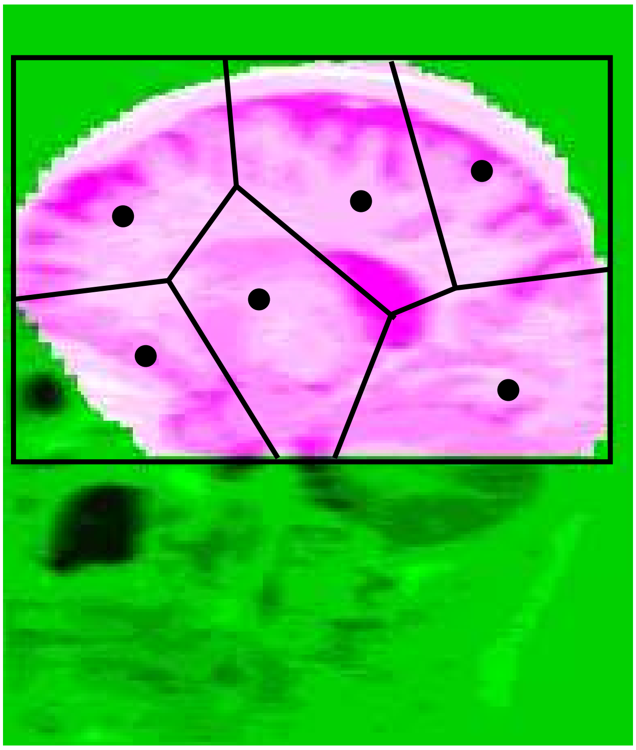
W_3 vs. W_2

- Cross-validated Parzen classifier results

Diagnosis	Normal	FTD	Vascular	Alz.
Normal	7	2	8	1
FTD	5	21	3	7
Vascular	3	2	13	4
Alz.	1	3	6	28

Motivation for GA Optimisation

- CSF boxes not related to brain structure
  - sub-optimal e.g. FTD often produces degeneration of frontal lobes
- Optimise boundary positions
  - replace boundaries with centre points
  - tessellate space with Voronoi cells



  - point pairs form natural building blocks
  - cost function noisy, form unknown
  - use GA

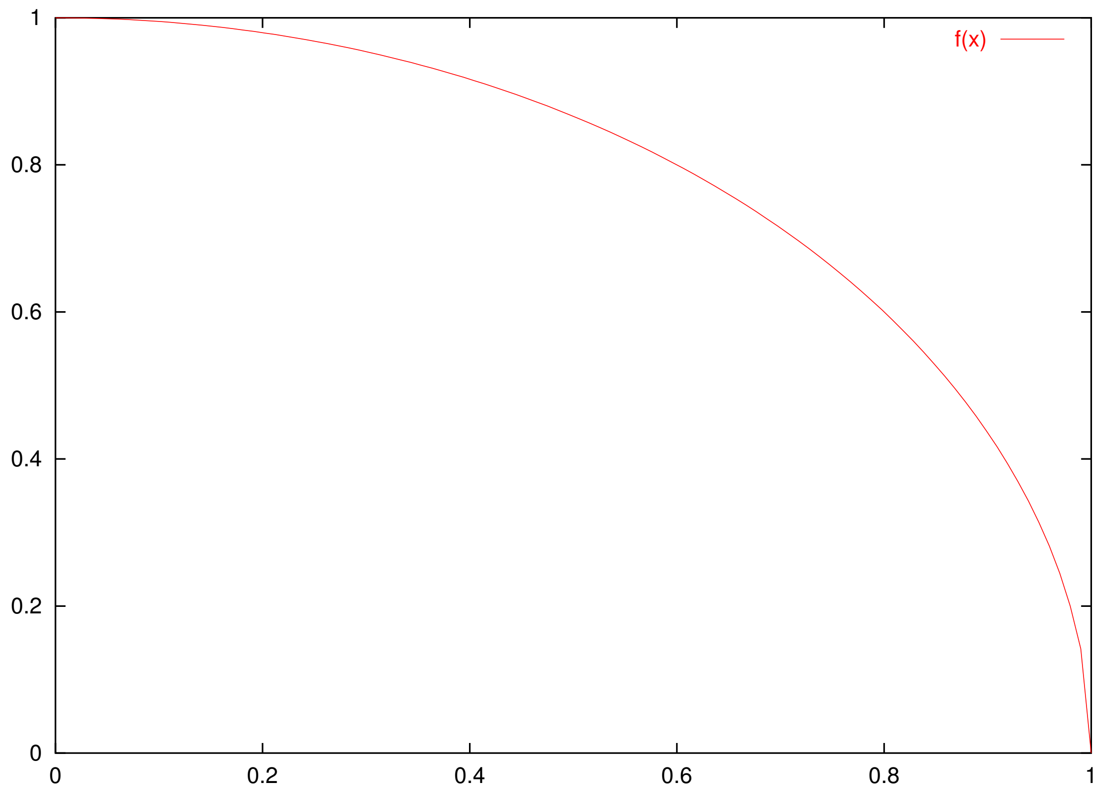
GA Goals

- Multi-objective cost function
  - use diagonal elements of confusion matrix
  - diversity preservation essential
- Diversity preservation
  - global convergence will occur if any pair of solutions can interact
  - nature: strict mating restrictions (inter-species) + dynamic cost function (intra-species)
  - convergence to occur locally, but not globally
- Minimise free parameters
  - population size, mutation probabilities
  - diversity preservation strategy parameters (e.g. crowding factor)

Multi-objective Fitness Evaluation

- Non-Pareto approaches
  - e.g. Schaffer (1985) VEGA
  - linear weighting of objectives
  - does not cover concave regions of cost surface
- E.g. minimise x and y for

$$y \geq \sqrt{1 - x^2}$$



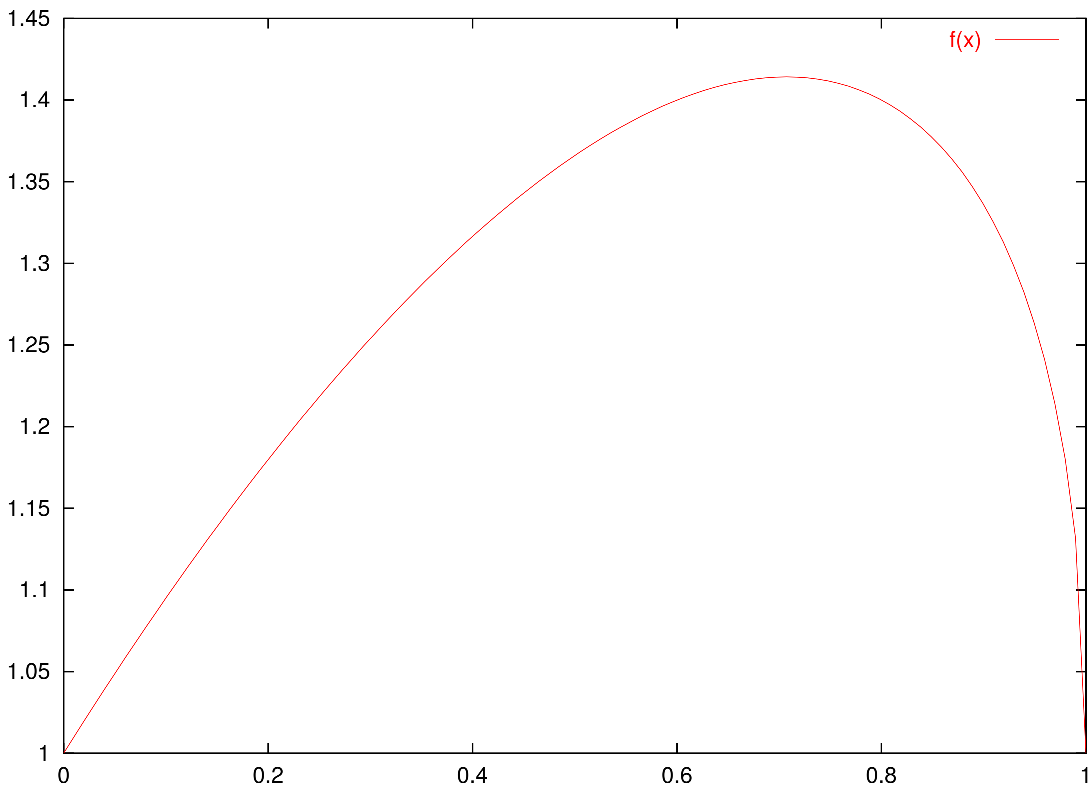
  - cost is linear weighting of x and y

$$f = \alpha x + \beta y$$

Multi-objective Fitness Evaluation 2

- For solutions on the Pareto front

$$f = \alpha x + \beta\sqrt{1 - x^2}$$



f vs. x for $\alpha = 1$ and $\beta = 1$

- Only the ends of concave regions of the cost surface will be found

Multi-objective Fitness Evaluation 3

- Pareto-based approaches
  - e.g. Goldberg (1989) Pareto ranking
  - find non-dominated solutions → rank 1
  - remove rank 1 solutions
  - iterate → rank 2, 3...
  - probability of reproduction based on rank
- Fitness evaluated purely on dominance basis

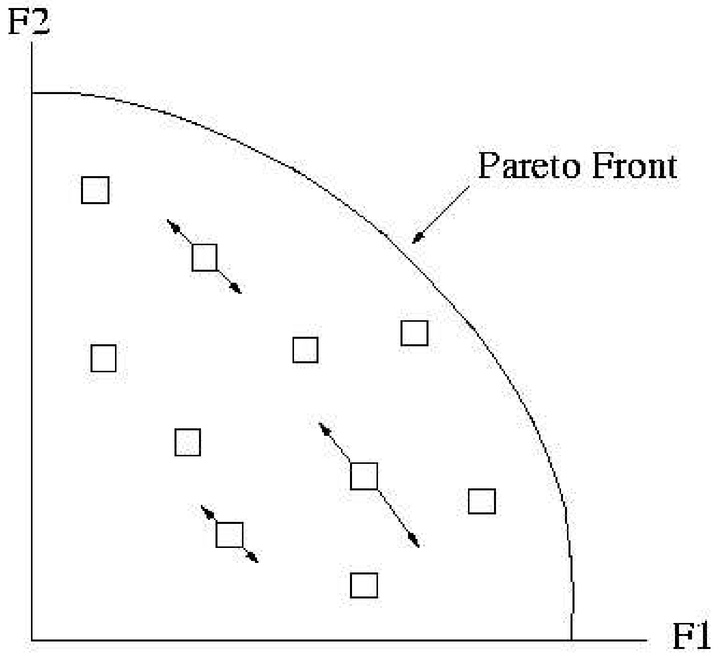
The Larcombe GA: Reproduction

- Larcombe (1996) Ph.D. thesis: GA optimisation of chip layer partitioning problem
- Avoid fitness proportional selection
  - select single pair of solutions per iteration
  - fitter offspring replace parents in a first-come, first-served competition
  - c.f. Cavicchio's preselection
- What about the Fundamental Theorem?

$$m(H, t+1) \geq m(H, t)\frac{f(H)}{\bar{f}}...$$

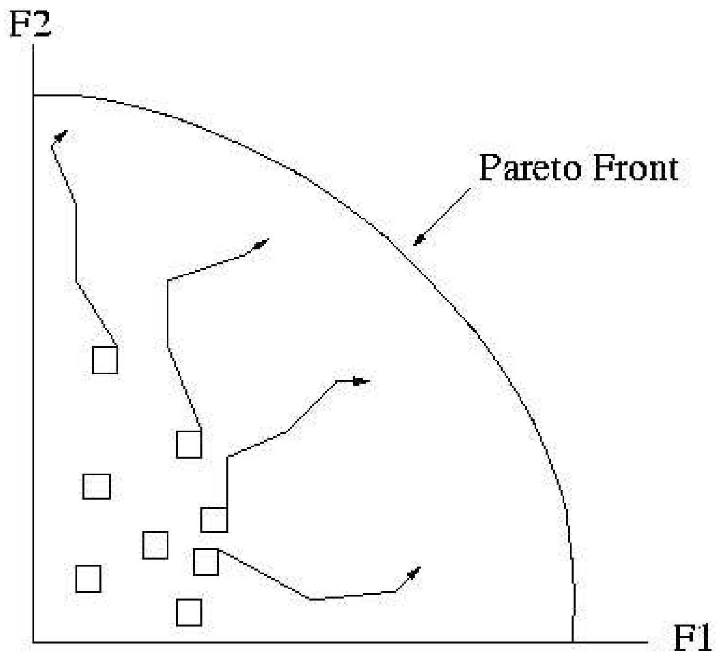
  - schemata can double per iteration $\Rightarrow$ quadratic growth
  - c.f. fitness scaling with scaling factor $= 2$

Fitness and Direction Vectors



- Parent-offspring pairs can continually replace each other
  - oscillatory behaviour
- Add direction flags to genome
  - one flag for each objective: initialised at 0
  - child replaces parent if it wins on any one objective
  - children of that child must win on the same objective
  - more directions can be set to 1 at any time

Static Extraction



- Solutions can become stuck at local minima
  - super-fit individuals: *statics*
- Remove these after n failed reproduction attempts
  - place in a secondary population: *static extraction*
  - replace with *random immigrants*
  - population size effectively infinite
- Diversity preservation + free parameter reduction

Coding and Mutation

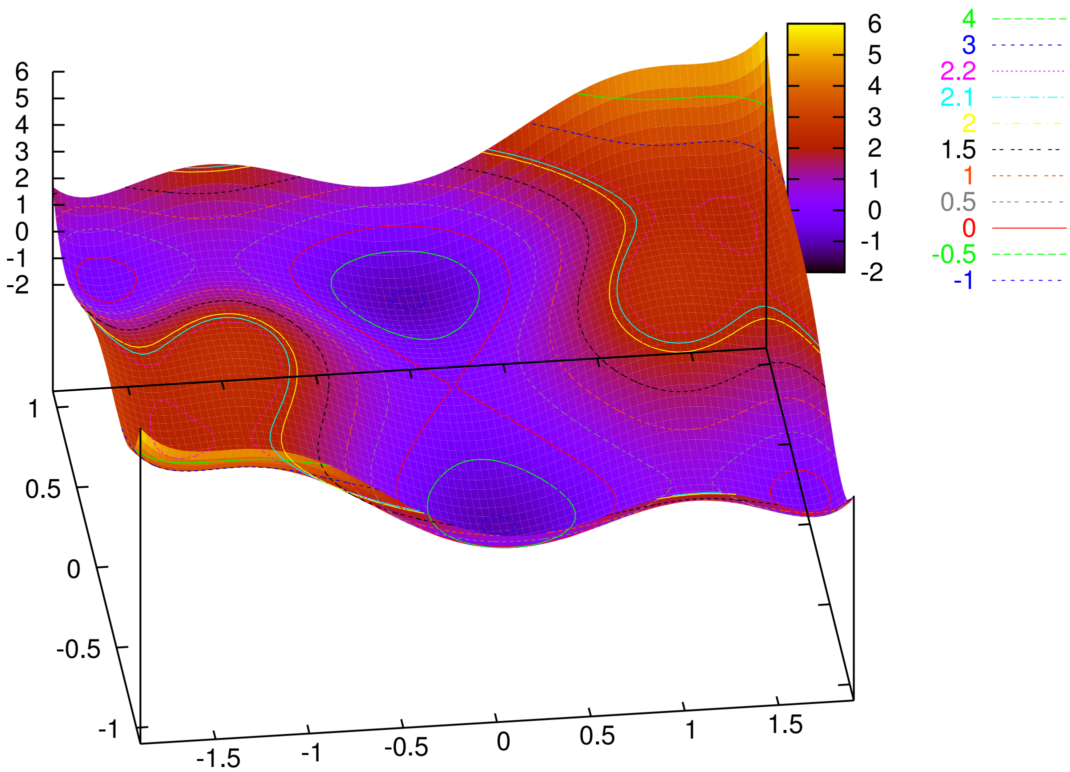
- Larcombe: binary coding
- Atrophy analysis: list of 6 real-valued points
  - point pairs form natural building blocks
  - distance parameter required for mutation
- Self-adaptive mutation
  - add mutation parameters to genome as a header
  - points: probability and fractional distance
  - header: probability and probability change
  - allow mutation probabilities to mutate
- Previous work (e.g. Cavicchio, 1970)
  - centralised control

Crossover

- For atrophy analysis
  - treat each point as a single gene
  - implement single point crossover
- For testing on simple functions
  - code a single point on a test function
  - new crossover routine required
- Previous work
  - Ono, Kita and Koboyashi, GECCO 1999
  - many techniques e.g. place children randomly on line between parents
  - adopt simple method: place children randomly in hypercube defined by parents

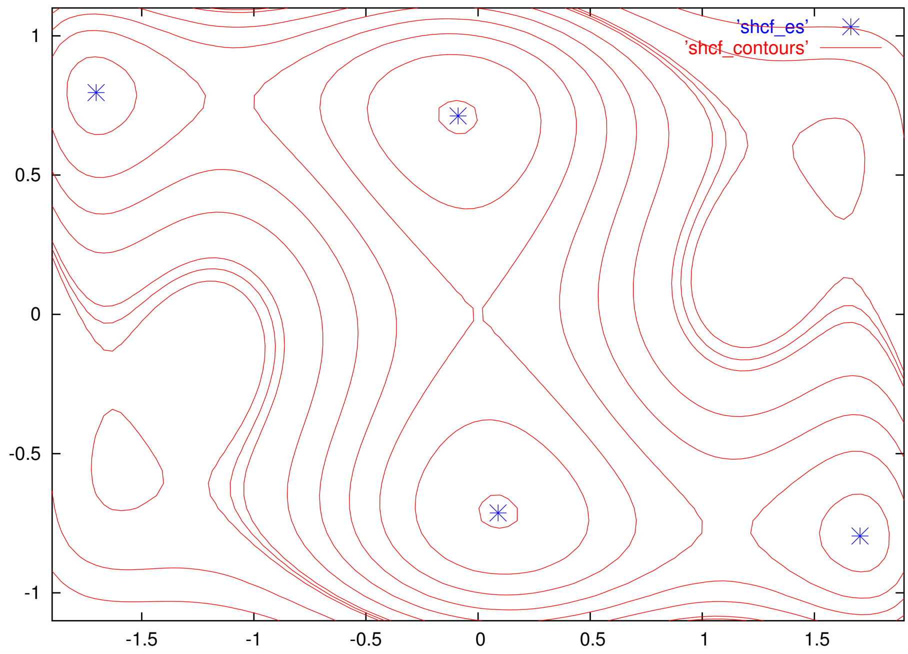
Initial Testing

- Implementation:
  - Larcombe GA: static extraction, random immigration
  - additions: real-valued representation, self-adaptive mutation
  - one free parameter left: extraction age
- Initial testing: six-hump camelback function

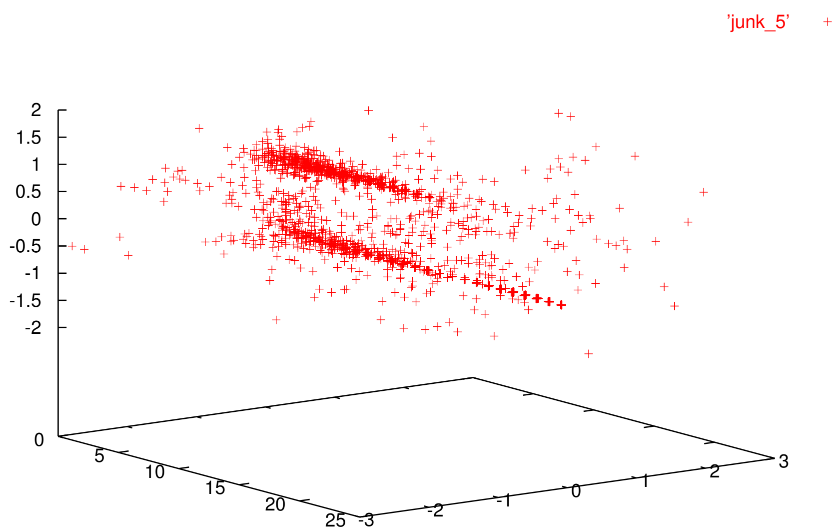


Initial Testing 2

- Extracted set at generation 25:



- General population:



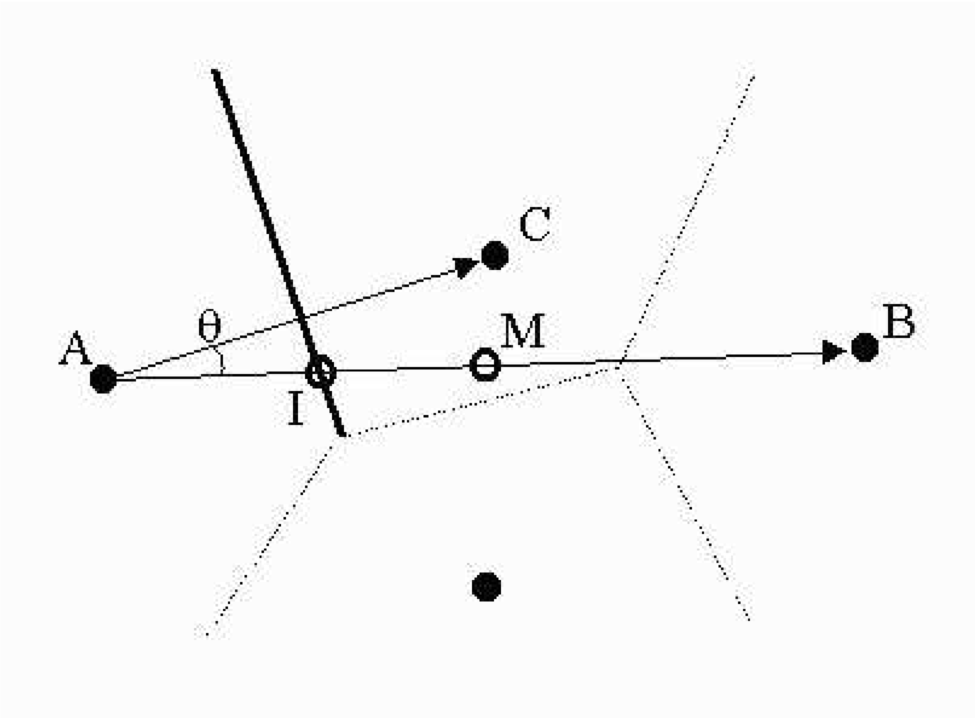
– diversity loss due to random selection

Diversity Preservation

- c.f. fitness scaling
  - start of run: super-fit individuals
  - end of run: lack of fitness differentiation
  - multi-modal function: loss of global optima
- Larcombe GA
  - static extraction limits super-fit individuals
  - global optima still lost at end of run
- Additional diversity preservation required
  - add mating restriction
  - aim: population convergence should be local but not global
  - convergence in extracted set instead of population

Voronoi Mating Restriction

- Avoid additional parameters: use algorithm's definition of local vs. global
  - restrict mating to Voronoi neighbours



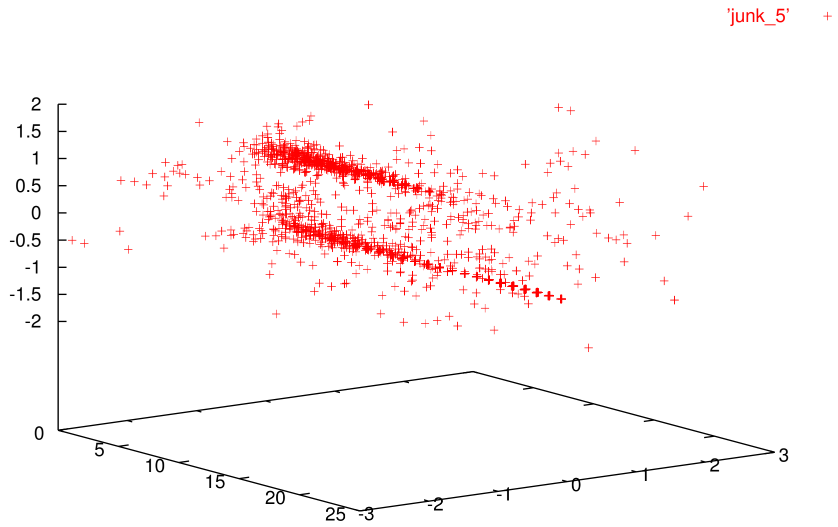
  - rejection criterion

$$|\mathbf{AM}| > |\mathbf{AI}| \Rightarrow \frac{|\mathbf{AC}|^2}{\mathbf{AB}.\mathbf{AC}}$$

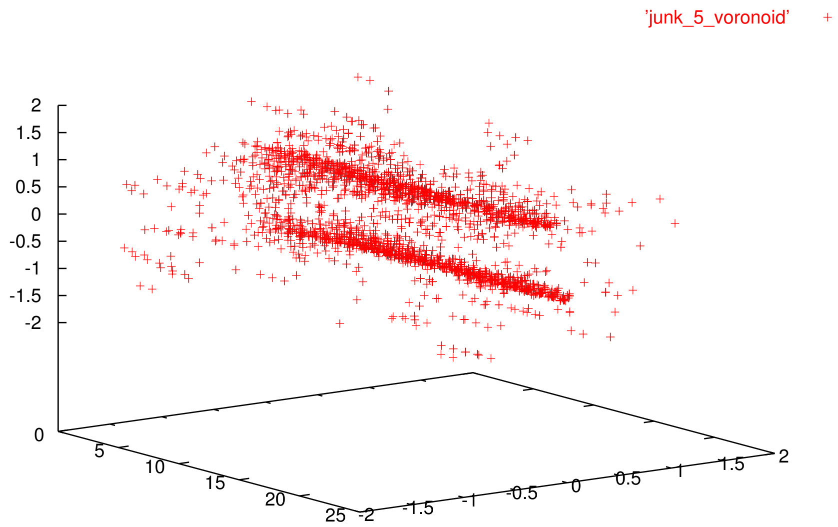
- During selection
  - randomly select parent 1
  - randomly select parent 2 from the Voronoi neighbours

Influence of Mating Restriction

- General pop. without mating restriction



- General pop. with mating restriction



- Diversity preserved indefinitely

Multi-Objective Testing

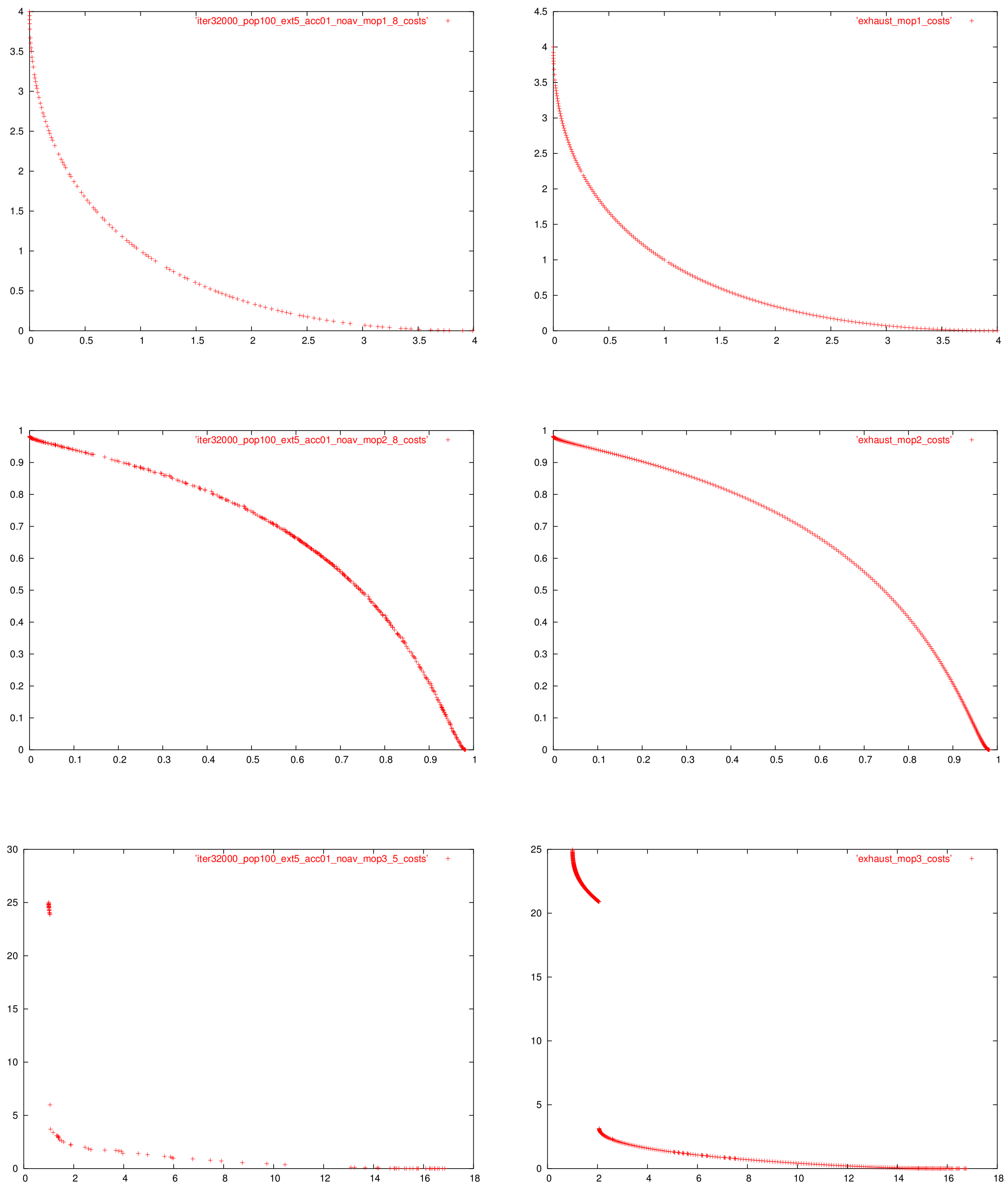
- Use seven function test-function set from Van Veldhuizen (1999): Pareto fronts
  - connected/unconnected in search space
  - connected/unconnected in cost space
  - convex/concave
  - single/multiple Pareto curves
- Use test procedure/metrics from Van Veldhuizen (2000)
  - find true Pareto front using exhaustive search
  - run GA for 2^{16} cost function evaluations
  - 0.39% of space covered
  - generational distance

$$G = \frac{\sqrt{(\Sigma_{i=1}^{n} d_i^2)}}{n}$$

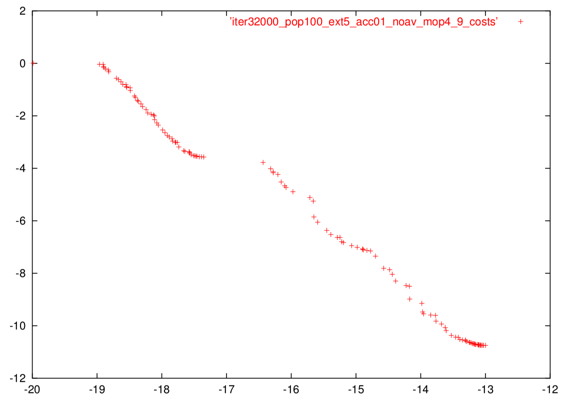
    d_i = distance to true Pareto front
  - ONVG = no. of solutions

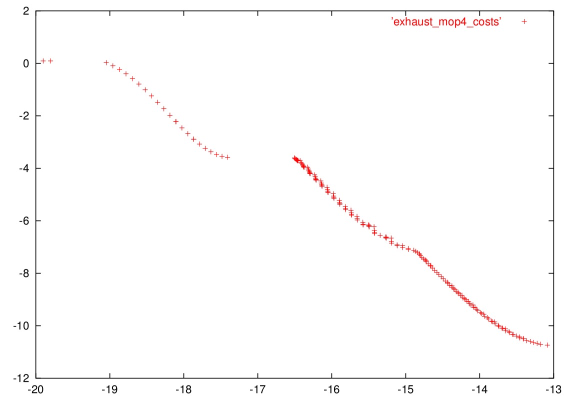
Multi-Objective Results

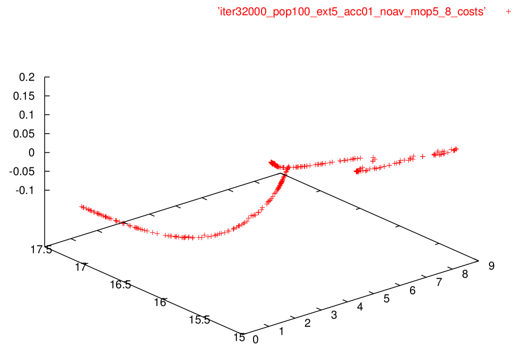
- Plots of GA results and exhaustive search in cost space

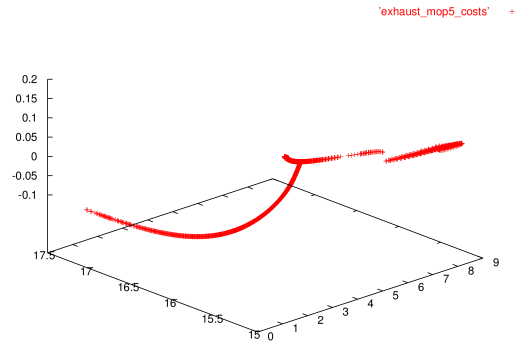


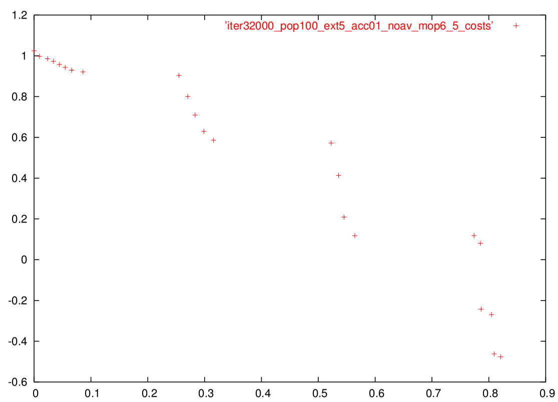
Multi-Objective Results

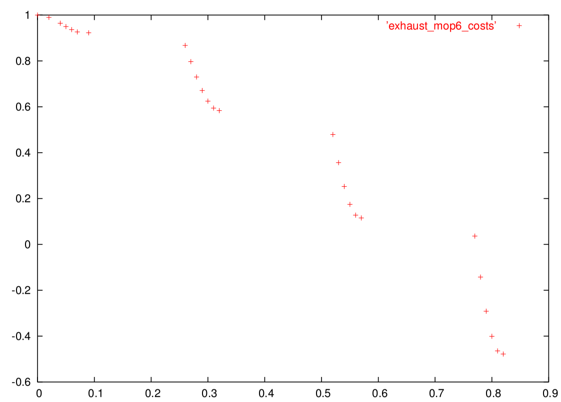




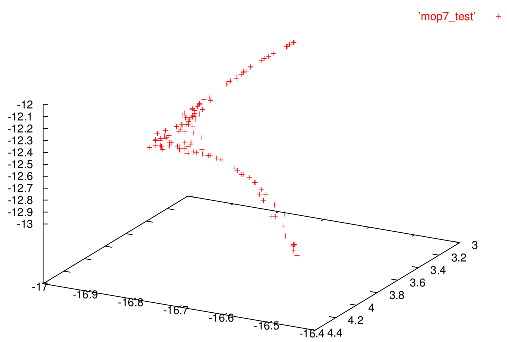


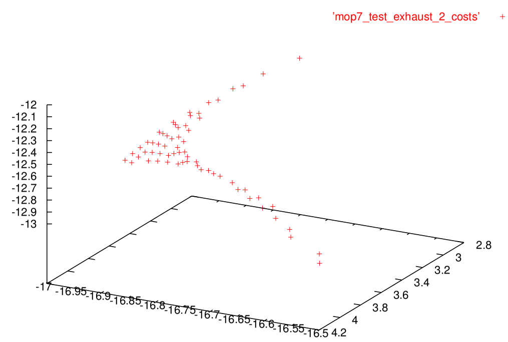






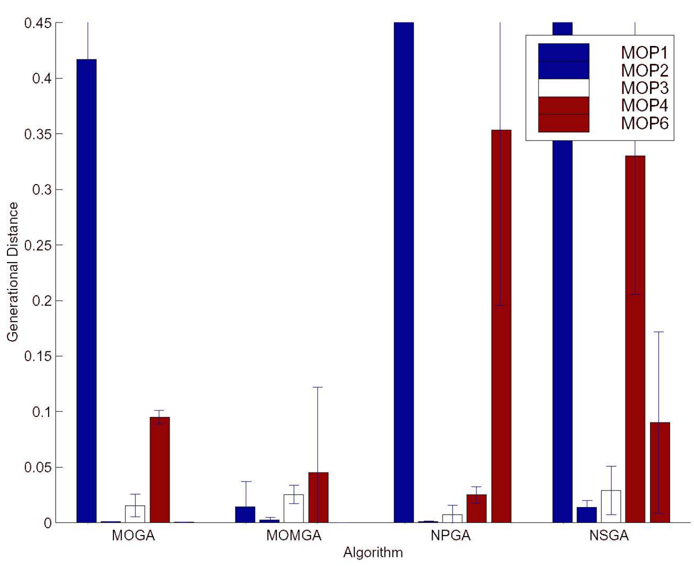
Multi-Objective Results

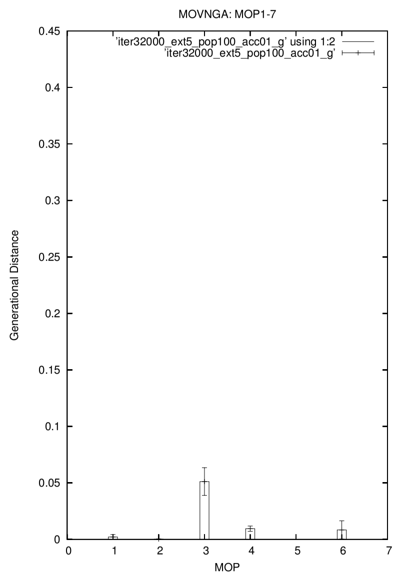




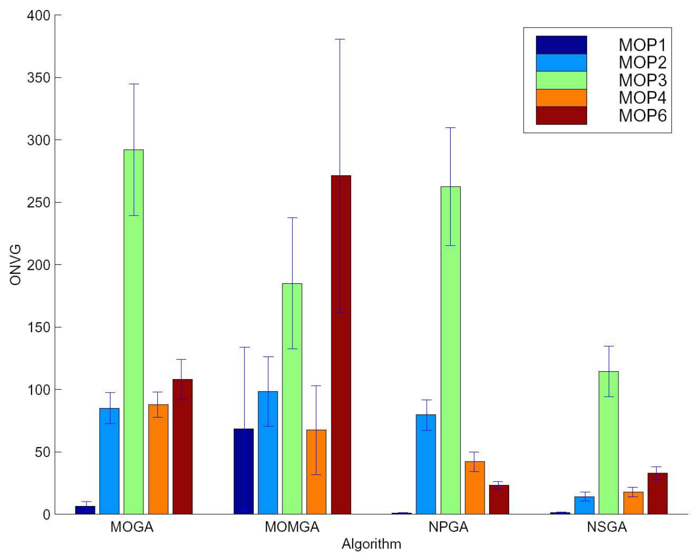
- Pareto front found for all 7 test functions
- Now compute G and ONVG
  - average over 10 GA runs for each test function
  - compare to GAs tested by Van Veldhuizen
    MOGA (Fonaseca and Fleming, 1998)
    MOMGA (Van Veldhuizen and Lamont, 2000)
    NPGA (Horn et al, 1994)
    NSGA (Srinivas and Deb, 1994)
  - all written to test aspects of GA theory
  - all highly cited
  - good competition

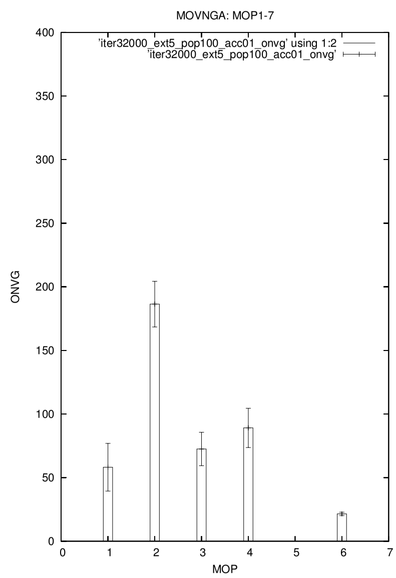
Multi-Objective Results





- Generational distance (lower is better)





- ONVG (higher is better)

Conclusions

- Novel GA
  - no fitness proportional reproduction
  - static extraction
  - random immigration
  - real-valued representation
  - self-adaptive mutation parameters
  - novel mating restriction (Voronoi neighbours)
- Testing
  - Van Veldhuizen's test function collection and metrics
  - comparable performance to MOMGA

Further Work

- Progress
  - GA written / tested on simple cost functions
  - atrophy analysis software (almost) finished
  - patient data collected and quality controlled
  - 200 sets / 6 groups
- Future work
  - finish age correction
  - more self-adaptation testing
  - test GA on complex cost functions (TSP?)
  - relate back to GA theory
  - parallelise GA