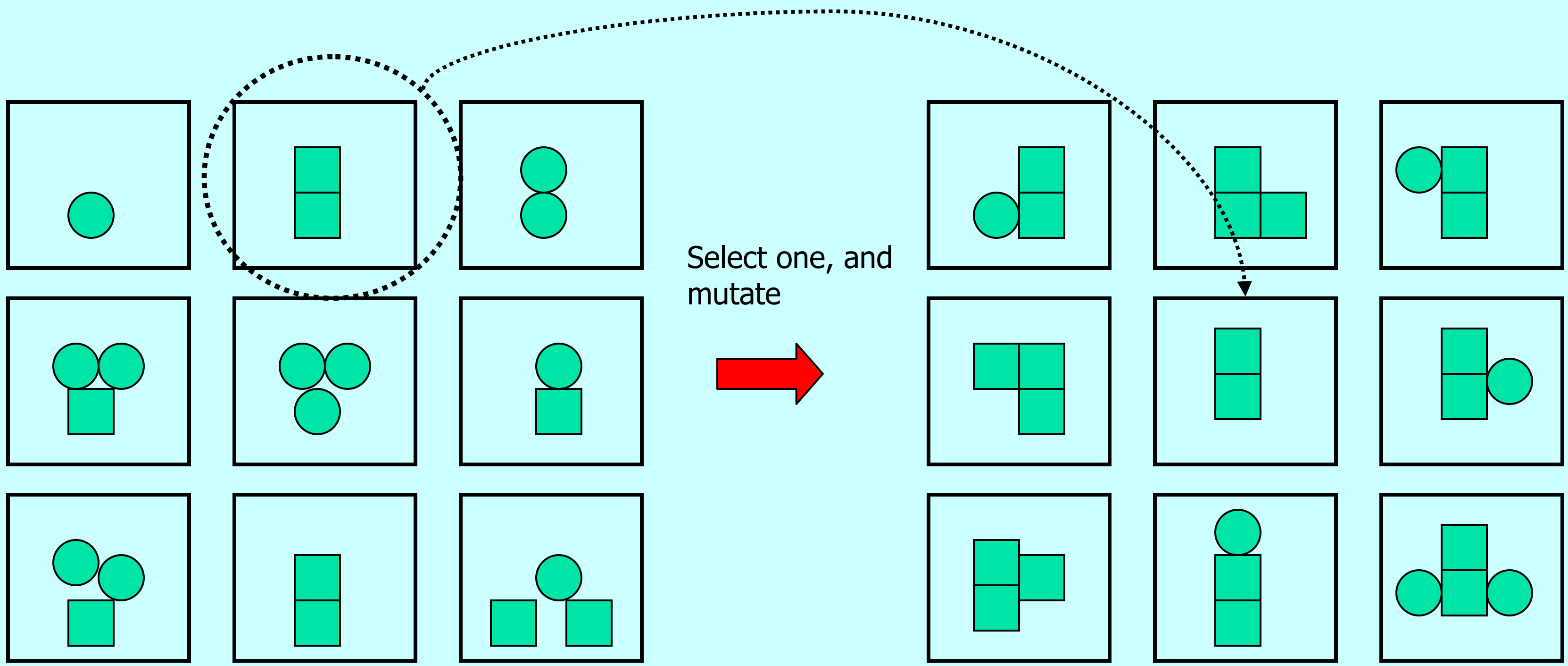
Subjective evolution (interactive genetic algorithms)



Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

Subjective evolution

- Used e.g.

  (1) As an illustration of the power of gradual, hereditary change (Dawkins)

  (2) Design problems

  (3) Evolutionary art

  (4) Face recognition (in e.g. crime fighting – chasing criminals through ”face space”

Biomorphs

- Introduced by Richard Dawkins (see e.g. *The blind watchmaker*)

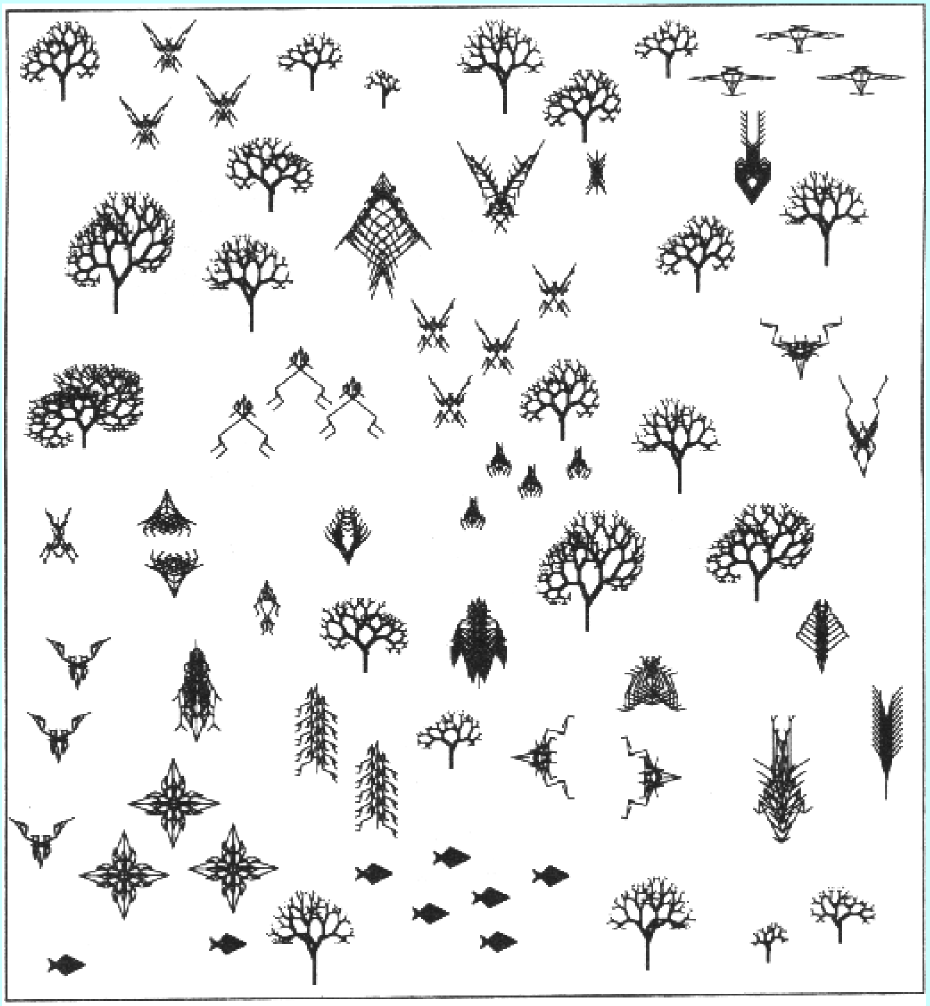
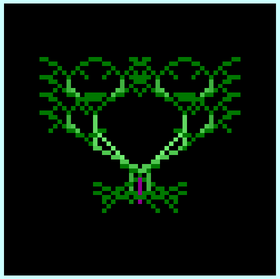


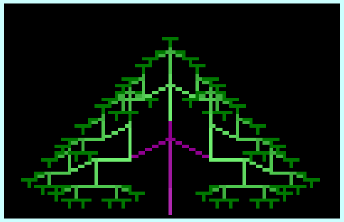
Figure 1.16 Safari park of black-and-white biomorphs, bred with the 'Blind Watchmaker' computer program.

Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

Biomorphs

- See also http://www.permadi.com/java/biomorph/



Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

Evolutionary design

- Example:

  Bentley and Wakefield:

  Evolution of table design

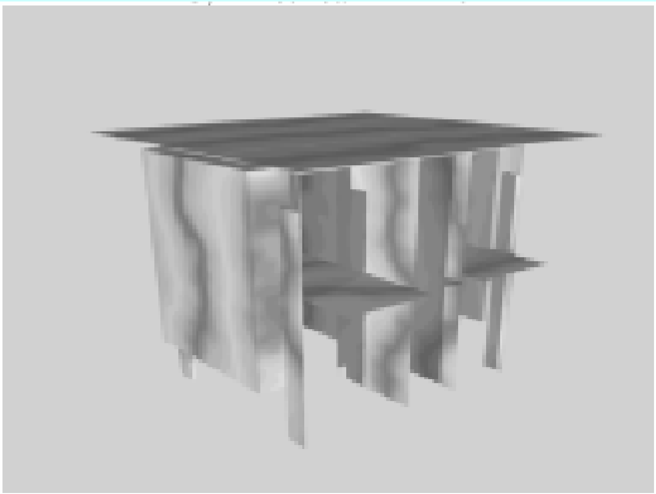


Fig. 13 Experiment 6: evolved design symmetrical in x = 0 and z = 0

See

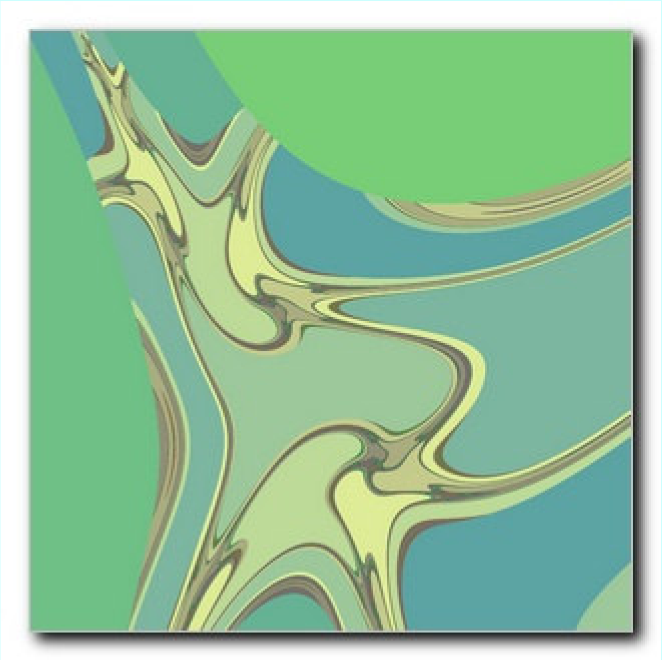
http://www.cs.ucl.ac.uk/staff/P.Bentley/BEWAC1.pdf

Evolutionary art

See

http://home.comcast.net/~davemc0/GenArt/About.html

http://web.genarts.com/galapagos/index.html



Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

Evolution of faces:

- http://dipaola.org/facespace/facespace.pdf

CHALMERS

Determination of orbital parameters for interacting galaxies

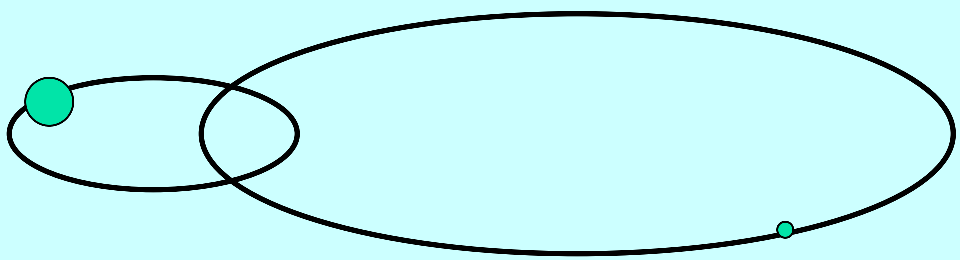
Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

Determination of Orbits

Given the positions and velocities of two (point) particles at any given time, the orbit can be determined.

$$\ddot{x}_1 = -G\frac{m_2(x_1 - x_2)}{\left\| x_1 - x_2 \right\|^3}$$

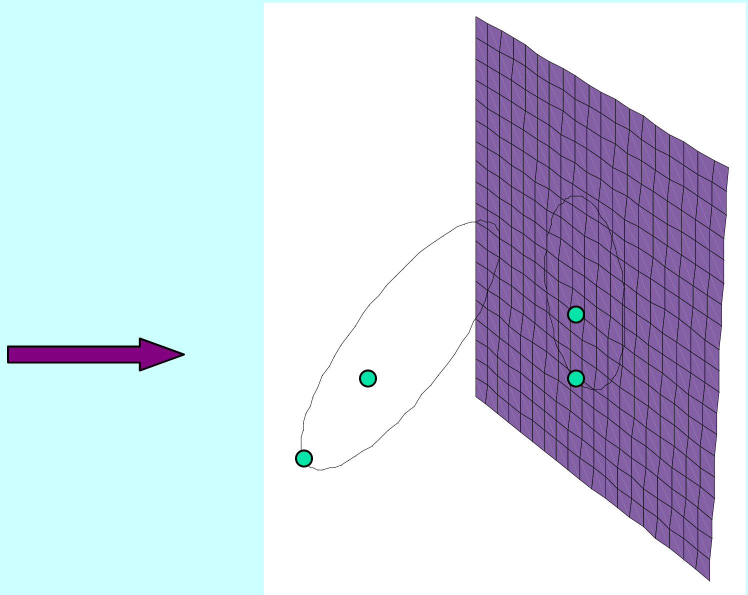
$$\ddot{x}_2 = -G\frac{m_1(x_2 - x_1)}{\left\| x_2 - x_1 \right\|^3}$$



Problems:

(1) One can only observe the projection of an orbit on the plane of the sky (ΔX, ΔY).

(2) One can only determine velocity components along the line-of-sight (ΔV_z)



Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

Galaxies are not point particles! For each galaxy, two angles (inclincation and position angle) are needed to determine the orientation of the (rotationally symmetric) disc.

Also, the masses of the two galaxies need to be estimated.

Thus, there are (at least) 9 unknown variables: ΔZ, ΔVx, ΔVy, I1, PA1, I2, PA2, M1, M2.

The values of these 9 variables can be determined using a genetic algorithm.

Testing a Candidate Solution

Assuming that we have a set of values for the unknown parameters, how does one test these values?

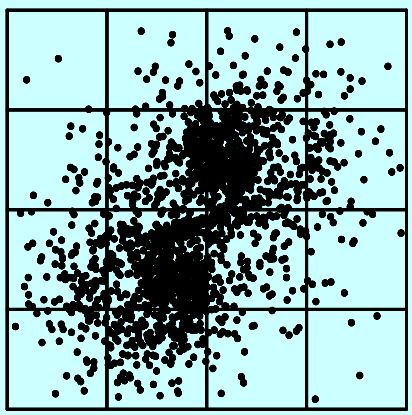
Method:

(1) Given the observed parameters and the values assigned to the unknown parameters, replace the galactic discs by point particles, and integrate backwards in time until the two point particles are far from each other.

Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

(2) Add a circularly symmetric disc of particles to each of the point particles (or, perhaps, only to the main galaxy).

(3) Integrate *forward* in time, until the present time is reached. At this point, compare the (light) density of the spiral structure (and, possibly, the velocity field) of the simulated galaxies with that of the observed galaxies



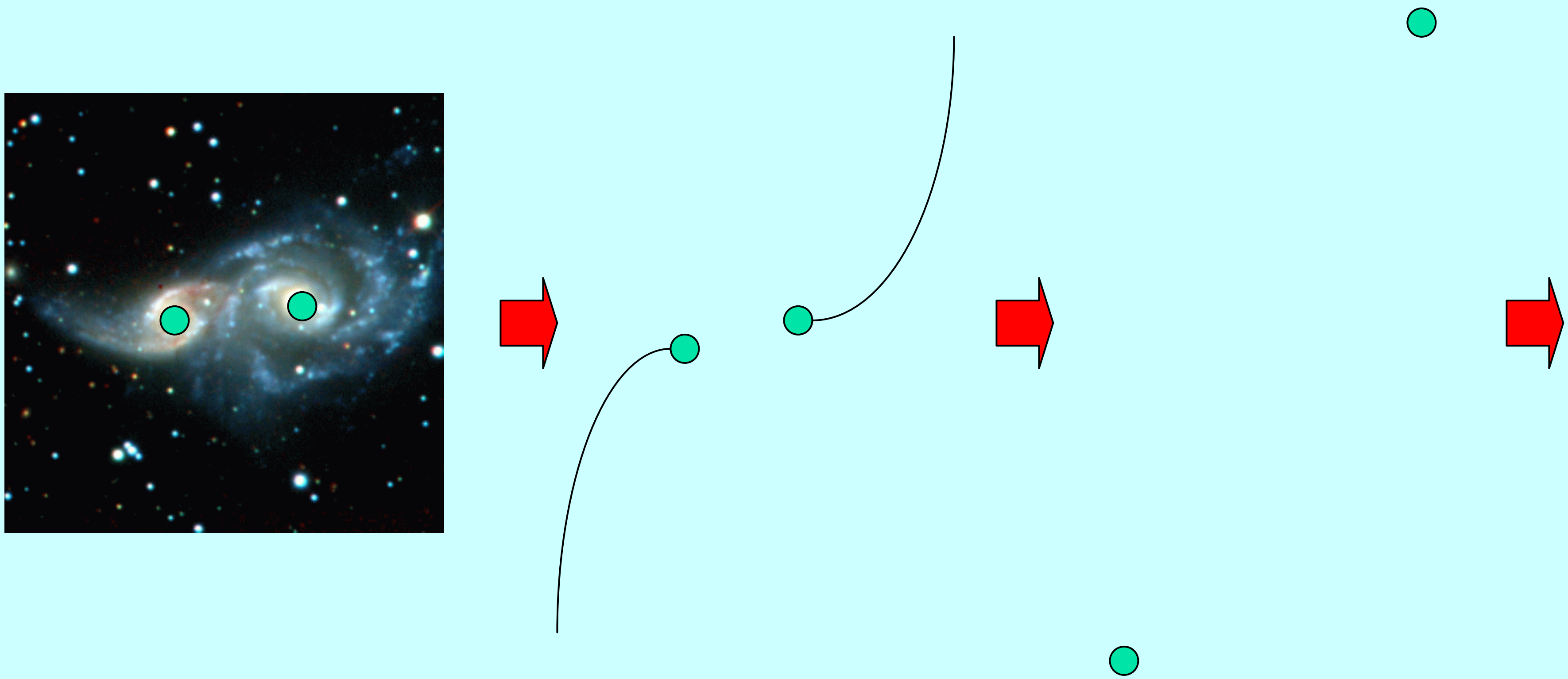
0.003	0.023	0.050	0.014
0.023	0.216	0.465	0.066
0.100	0.580	0.157	0.020
0.036	0.174	0.026	0.004

Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

The forward integration is carried out using a *restricted three-body simulator*.

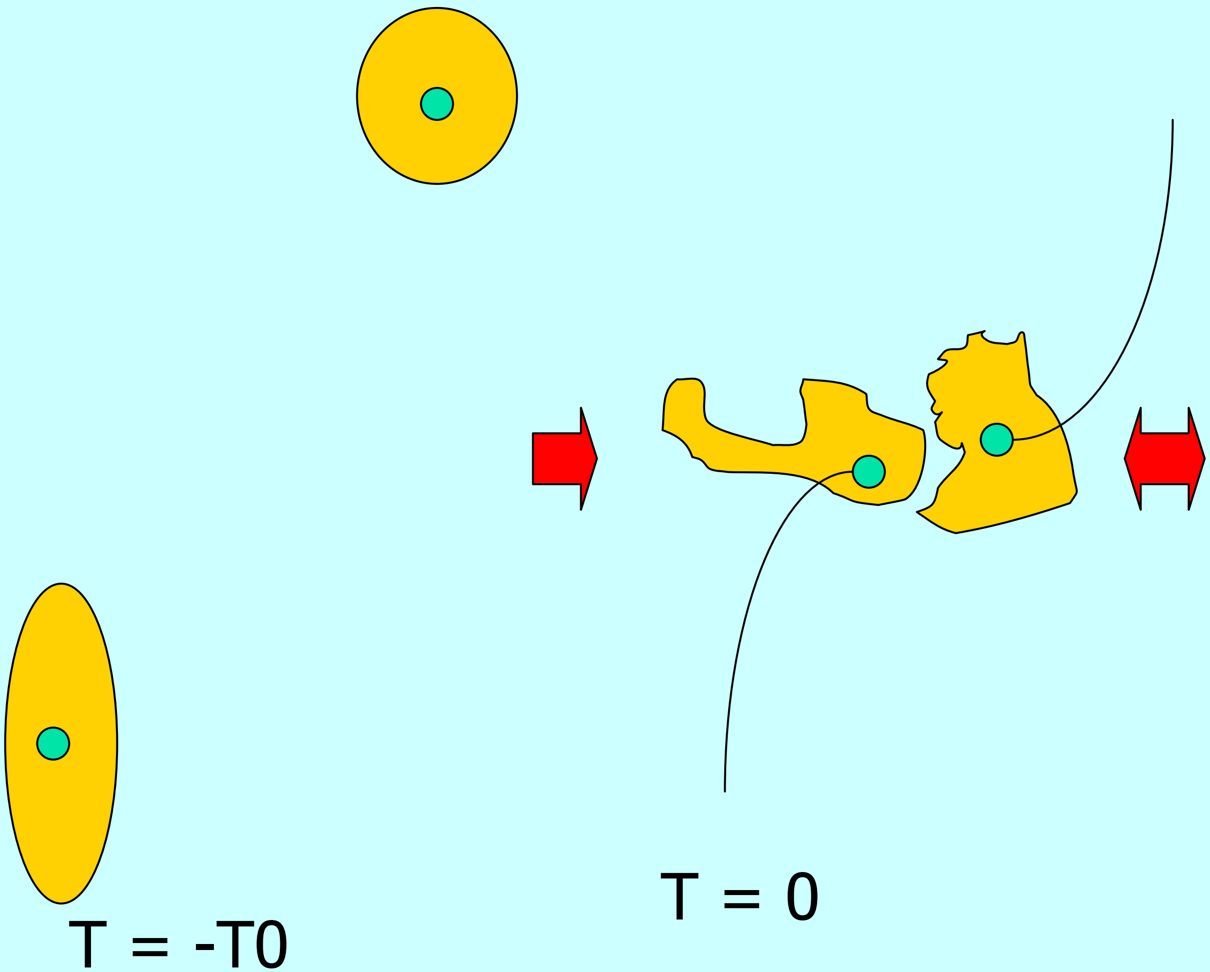
(4) Assign a score depending on the deviation between the observed and simulated galaxies, e.g.

$$f = \frac{1}{1+\delta} \qquad \text{where} \qquad \delta = \sum_{i,j} (m_{i,j}^{\text{obs}} + m_{\varepsilon}) \left| \ln \frac{m_{i,j} + m_{\varepsilon}}{m_{i,j}^{\text{obs}} + m_{\varepsilon}} \right|$$



T = 0 T = -T0

Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde



Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

Many candidate solutions (~10,000 – 100,000) are tested by the genetic algorithm until a satisfactory solution is found.

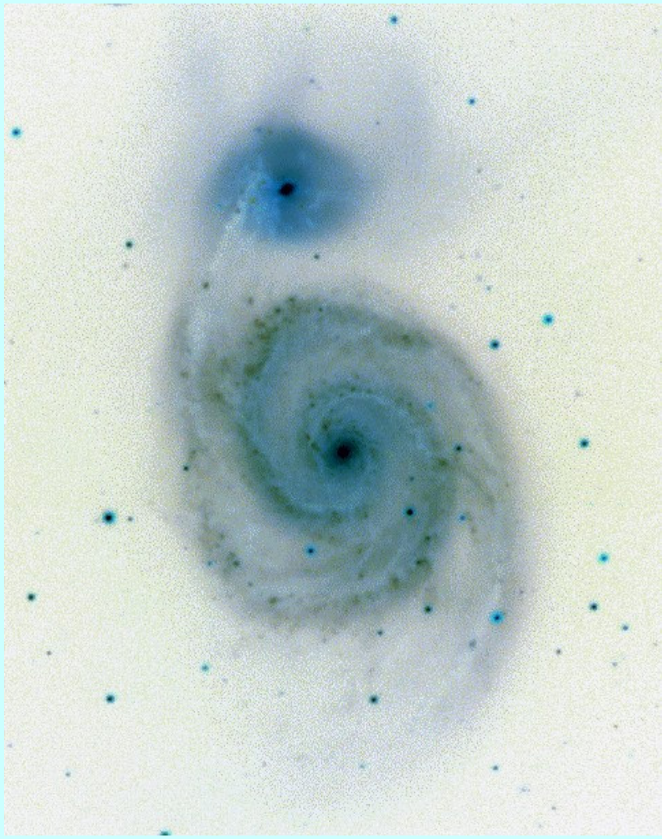
The method is *much* more efficient than trying to cover the whole multidimensional space with an undirected search.

Furthermore, the genetic algorithm does *not* get stuck in local minima, like gradient-based methods often do.

Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

The Case of NGC 5194/5195

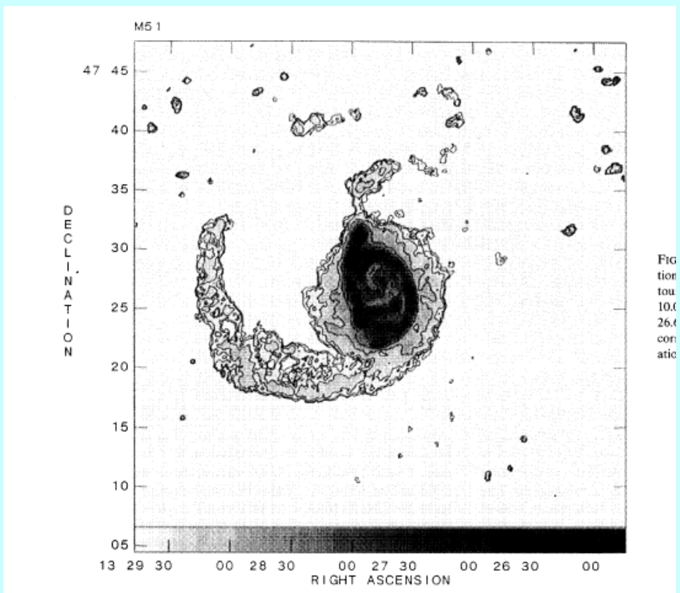
The M51 system, consisting of the main galaxy (NGC 5194) and its interaction partner (NGC 5195), is one of the most thoroughly studied galaxy pairs.

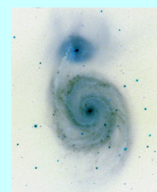


A reasonable guess would be that NGC 5195 is located in the same plane as the disc of NGC 5194, and that it has recently generated magnificent spiral arms.

Such a model (which gave rise to the term ”M51-like systems”) was made by Toomre and Toomre 1972.

However, in 1991, a gigantic arm of atomic hydrogen (HI) was discovered by Rots *et al*.





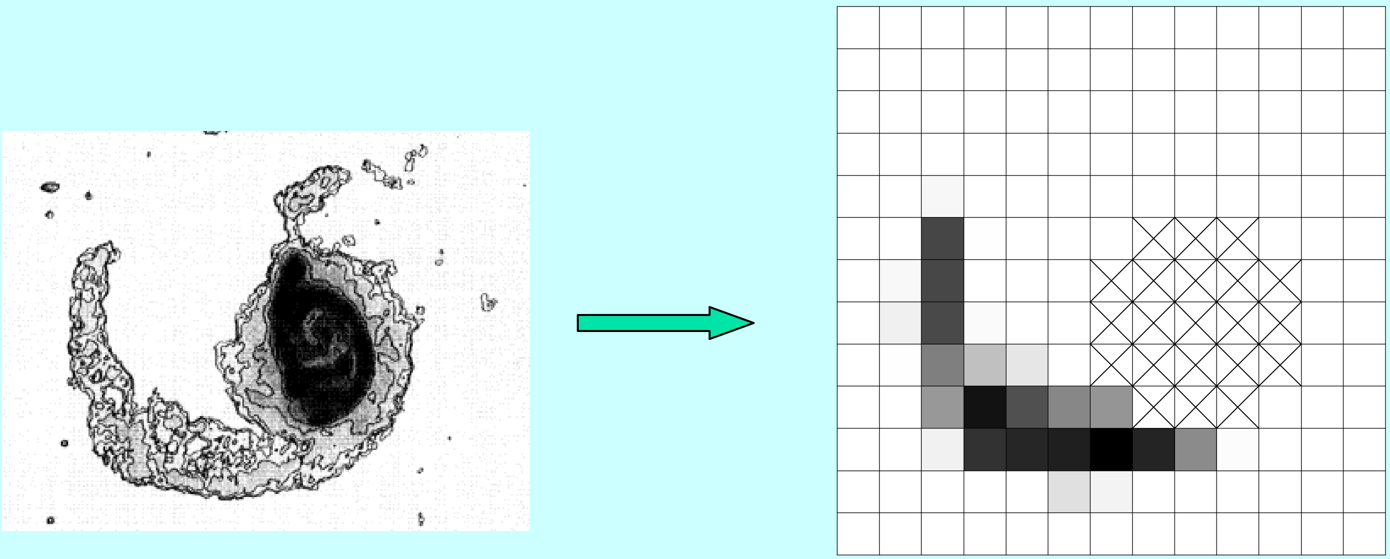
Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

The HI-arm is truly enormous, and cannot have formed in a recent passage by NGC 5195. Thus, M51 is *not* an M51-type system!

Engström and Athanassoula attempted to model the NGC5194/5194 interaction, but did not manage to fit the velocity field.

Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

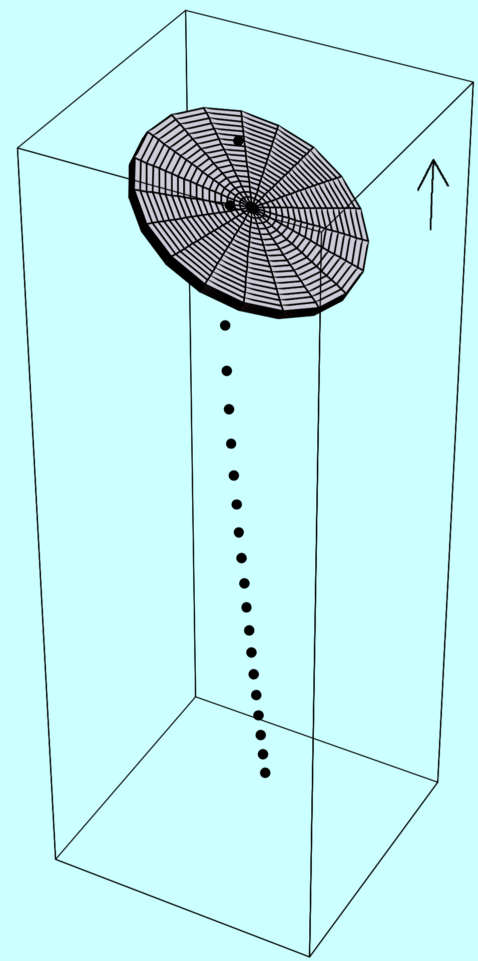
Karl Johan Donner and I applied the method to the NGC 5194/5195 system. Our aim was to fit the positions and velocities within the HI-arm. Thus, we blocked out the central region of the NGC 5194 galaxy.



Results for NGC 5194/5195

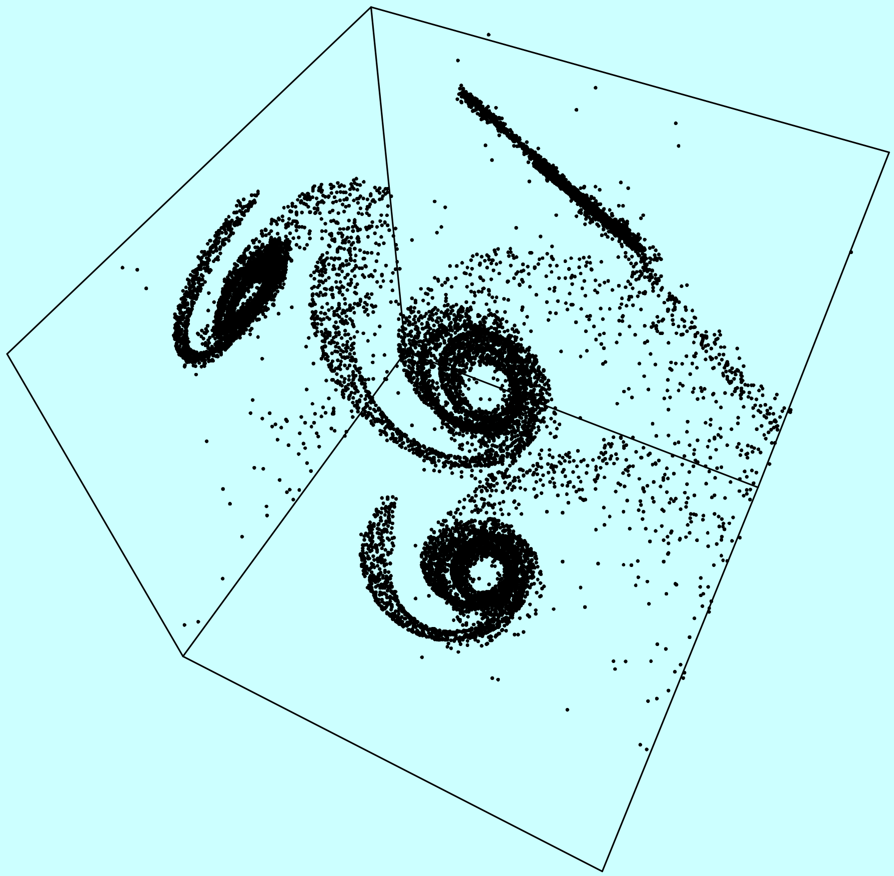
We found that, rather than being in the same plane as NGC 5194, the orbit of NGC 5195 is almost perpendicular to the plane of the sky.

On our best-fit orbit, NGC 5195 is presently located 147 kpc behind NGC 5194, and moving on a hyperbolic orbit. The pericentre passage occured 908 Myr ago, and the disc (of NGC 5194) was passed at a distance of 17 kpc from the centre, 849 Myr ago.



Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

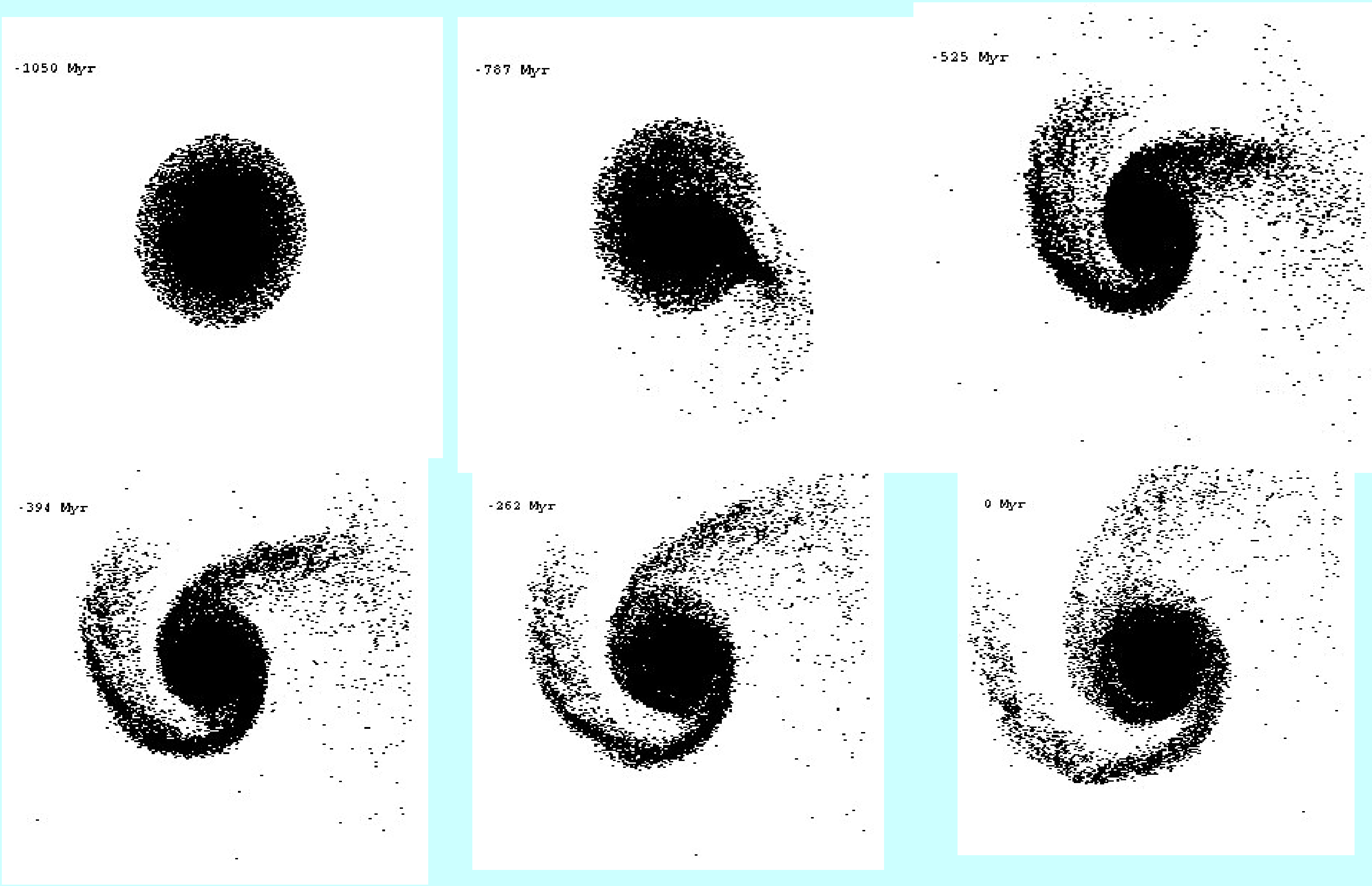
A view of the results obtained using the best-fit parameters (three different projections shown)



M51 is not an M51-type system!

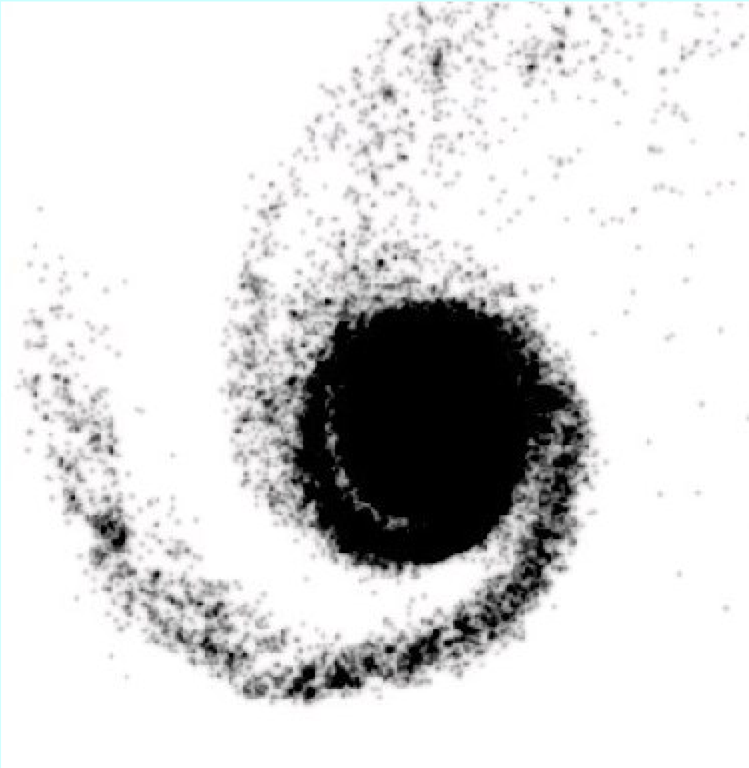
Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

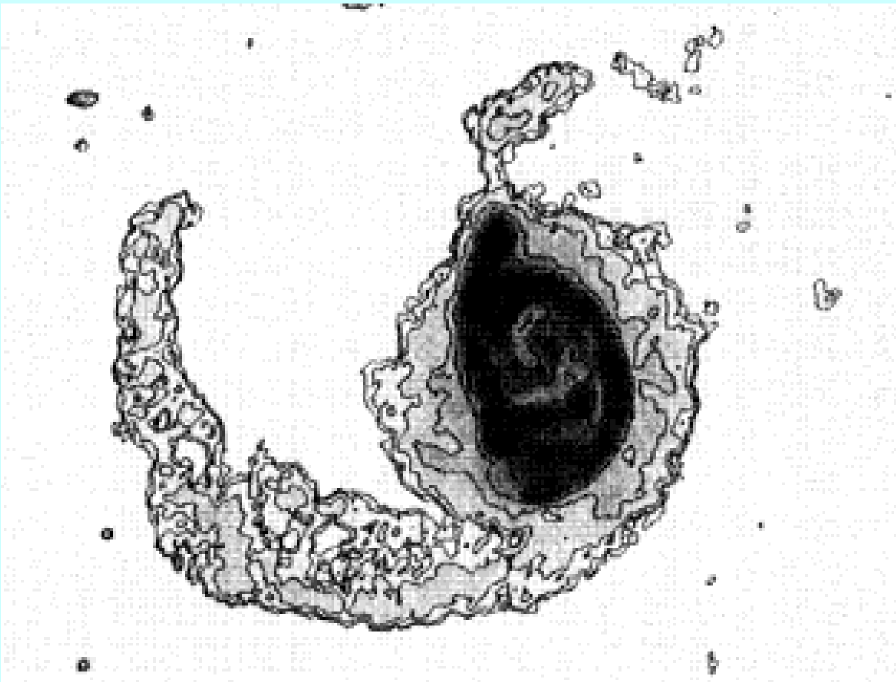
A self-gravitating simulation was performed as well, using the best-fit (initial) orbital parameters.



Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

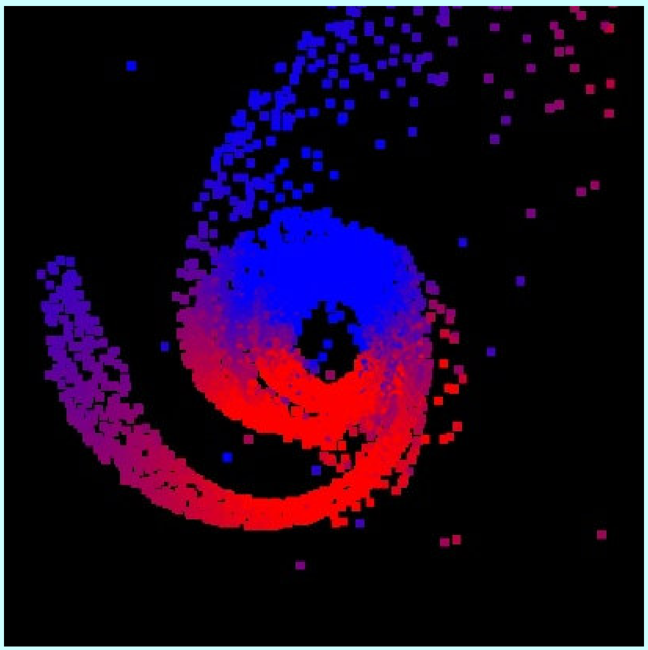
A comparison between our best-fit model and the observed galaxy:

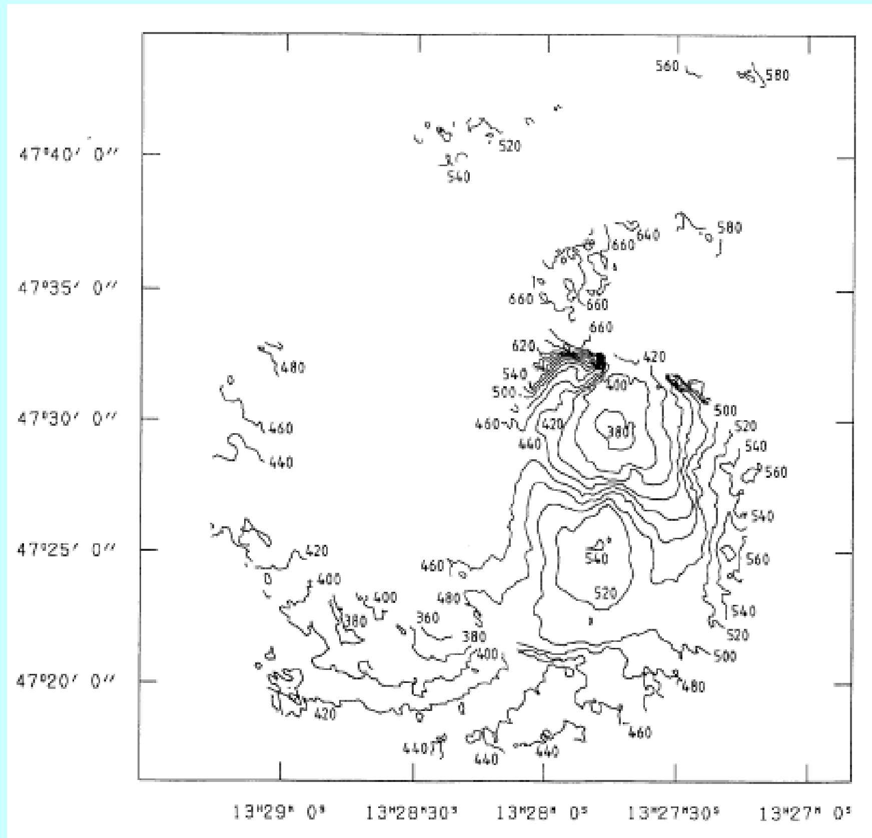




Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

Even though the position of the arm in our best-fit simulation corresponds well to observations, the velocity field is not reproduced perfectly:





Summary

- A method for determining the orbits of interacting galaxies has been developed.
- The method uses genetic algorithms, a powerful search and optimization method rooted in evolutionary biology.
- The method has been applied to the NGC 5194/5195 system.
- On the best-fit orbit, NGC 5195 passed the disc of NGC 5194 almost a billion years ago, and is now located far behind the disc of that galaxy.
- The main discrepancy between our model and the observed galaxy lies in the velocity field.

Home problem 1

- **Problem 1.1**

  Mostly correctly solved. Some common errors:

  -- Omitting fitness ranking

  -- Omitting the crossover probability (setting it to 1)

  -- Not checking for the exact solution (never give answers of the kind x1 = 0.999999977 in an engineering problem...)

  -- Providing an unsatisfactory report: **READ** the problem formulation!

Home problem 1

- **Problem 1.2**

  Also mostly correctly solved. Some common errors:

  -- Providing an unsatisfactory report: **READ** the problem formulation!

  -- Messy code => mistakes that are difficult to trace. **USE** separate functions, proper variable names etc.

  -- Searching for parameters in an incorrect range

  Correct values: tau1 = 2, tau2 = 3, w11 = 2, w12 = -5, w21 = 5, w22 = -3, b1 = b2 = 0.

Home problem 1

- **Problem 1.3**

  Also mostly correctly solved. Some common errors:

  -- Minor computational errors. However, in some cases those errors resulted in $\Sigma p(k) > 1$!!

Solution_EC2006_HP1.3.txt - Notepad

File Edit Format View Help

```
First pair        Second pair       Probability       Resulting strings               # Copies of S

Ind. 1   Ind. 2   Ind. 1   Ind. 2
101101   010101   101101   101101   16/100            010101,101101,101101,101101     3
101101   010101   101101   100101   12/100            010101,101101,100101,101101     2
101101   010101   101101   110000    8/100            010101,101101,110101,101000     1
101101   010101   101101   010101    4/100            010101,101101,010101,101101     2

101101   010101   100101   101101   12/100            010101,101101,101101,100101     2
101101   010101   100101   100101    9/100            010101,101101,100101,100101     1
101101   010101   100101   110000    6/100            010101,101101,110101,100000     1
101101   010101   100101   010101    3/100            010101,101101,010101,100101     1

101101   010101   110000   101101    8/100            010101,101101,101000,110101     1
101101   010101   110000   100101    6/100            010101,101101,100000,110101     1
101101   010101   110000   110000    4/100            010101,101101,110000,110000     1
101101   010101   110000   010101    2/100            010101,101101,010000,110101     1

101101   010101   010101   101101    4/100            010101,101101,101101,010101     2
101101   010101   010101   100101    3/100            010101,101101,100101,010101     1
101101   010101   010101   110000    2/100            010101,101101,110101,010000     1
101101   010101   010101   010101    1/100            010101,101101,010101,010101     1

p(0) = 0.00
p(1) = 0.52
p(2) = 0.32
p(3) = 0.16
p(4) = 0.00

n_avg = 1*0.52 + 2*0.32 + 3*0.16 = 1.58
```

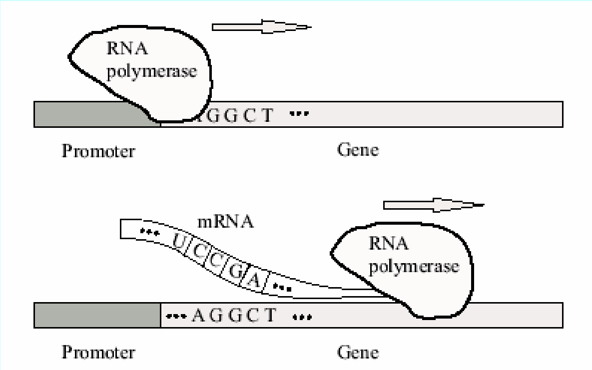
Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

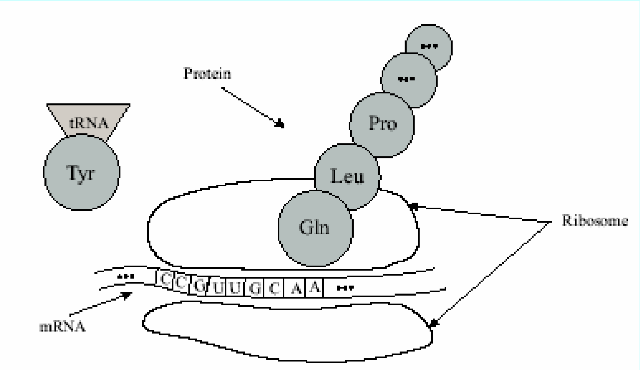
Brief course summary

- Handout 1: Biological basis
- Handout 2: Basics of EAs
- Handout 3: Using EAs
- Handout 4: Properties of EAs
- Handout 5: Advanced topics
- Handout 6: Version of EAs

Biological basis

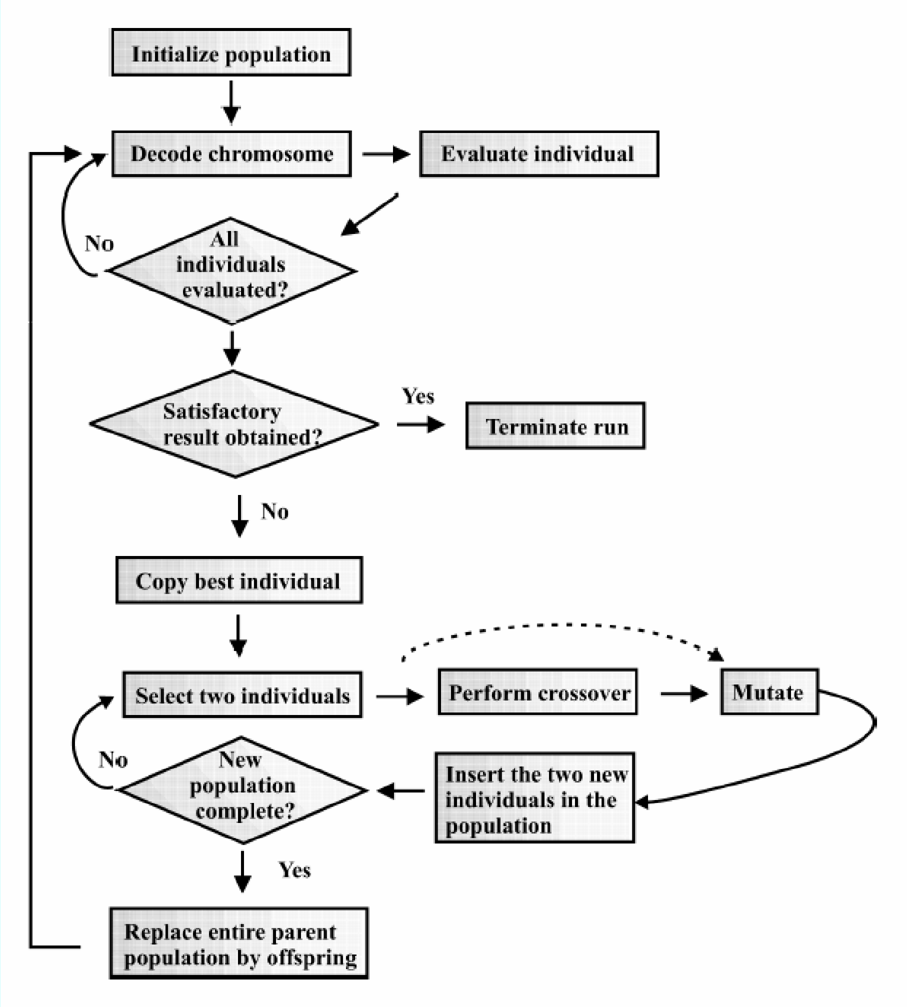
- EAs inspired by biological evolution
- Encoding (and use) of genetif information in biological systems: transcription, translation.
- Simplifications in EAs (compared to biological evolution) – gene interaction etc.





Basics of EAs

- General flow of an evolutionary algorithm
- Encoding schemes (binary, real-number etc.)
- Fitness assignment, fitness ranking
- Selection methods (roulette-wheel, tournament)
- Replacement methods (generational, steady-state)
- Crossover, mutation
- Elitism



Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se   www: www.me.chalmers.se/~mwahde

Using EAs

- Writing Matlab program code for a basic EA
- Selection of suitable parameters
  (by evaluation of benchmark functions)

```
function new_individuals = crossover(population,i1,i2,ngenes);

cp = 1 + fix(rand*(ngenes-1));

for j = 1:ngenes
  if (j < cp)
     new_individuals(1,j) = population(i1,j);
     new_individuals(2,j) = population(i2,j);
  else
     new_individuals(1,j) = population(i2,j);
     new_individuals(2,j) = population(i1,j);
  end
end
```

Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se   www: www.me.chalmers.se/~mwahde

Properties of EAs

- The concept of schemata, the schema theorem

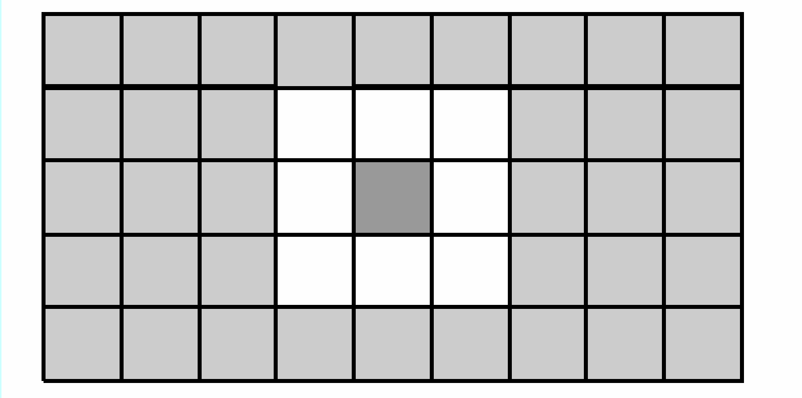
$$\Gamma(S, g+1) \geq \frac{\overline{f}(S)}{\overline{f}} \Gamma(S, g) \left(1 - p_c \frac{D(S)}{n-1} - O(S) p_{\text{mut}}\right),$$

- Premature convergence (+ methods for prevention)
- Analytical properties of EAs – calculations of distributions, average fitness values etc. for the infinite-population case

$$p_s(k) = \frac{f(k) p_{s-1}(k)}{\sum_{k=0}^{n} f(k) p_{s-1}(k)},$$

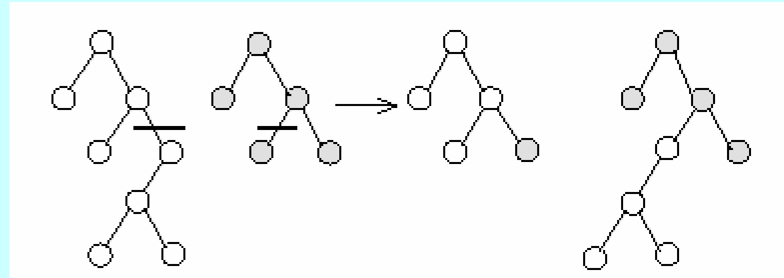
Advanced topics

- Representations (e.g. messy encoding schemes, variable-length encoding schemes (for ANNs etc.))
- Grammatical encoding
- Selection (Boltzmann)
- Choice of fitness measures
- Pareto-optimality
- Constrained optimization
- EAs with mating restrictions
- Experiment design – training, validation etc.



Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

Versions of EAs



- GA, GP, ES, (EP)
- Difference between GAs and GP
- Versions of GP: tree-based + LGP
- Operators in GP
- Encoding/decoding in LGP

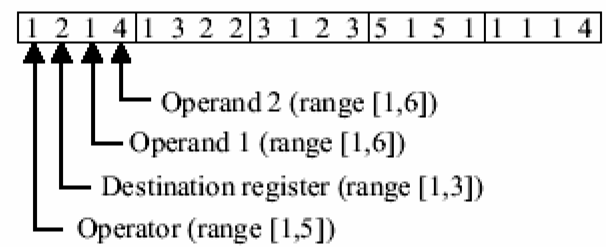


Figure 6.3: An example of an LGP chromosome.

Genes	Instruction	Result
1, 2, 1, 4	$r_2 := r_1 + c_1$	$r_1 = 1, r_2 = 2, r_3 = 0$
1, 3, 2, 2	$r_3 := r_2 + r_2$	$r_1 = 1, r_2 = 2, r_3 = 4$
3, 1, 2, 3	$r_1 := r_2 \times r_3$	$r_1 = 8, r_2 = 2, r_3 = 4$
5, 1, 5, 1	if $(r_1 > c_2)$	$r_1 = 8, r_2 = 2, r_3 = 4$
1, 1, 1, 4	$r_1 := r_1 + c_1$	$r_1 = 9, r_2 = 2, r_3 = 4$

Mattias Wahde, PhD, associate professor, Chalmers University of Technology
e-mail: mattias.wahde@me.chalmers.se www: www.me.chalmers.se/~mwahde

About the exam...

- Be ON TIME!!! (Starts at 08.30!)
- Make sure to bring a calculator: one that cannot store text. NO exceptions! (Simple calculators available at Cremona).
- Write solutions on the official Chalmers papers (will be provided)

Exam in FFR 105 (Evolutionary computation), 2006-10-21, 08.30-12.30, M.
It is allowed to use a calculator, as long as it cannot store any text. Furthermore, mathematical tables (such as Beta, Standard Math etc.) are allowed, provided that no notes have been added. However, it is *not* allowed to use the handouts (including the problems) from the course during the exam.

Note! In problems involving computation, show *clearly* how you arrived at your answer, i.e. include intermediate steps etc. Only giving the answer will result in zero points on the problem in question.