Genetic algorithms

Martin Netolický

November 6, 2008

The outline

- Introduction: The interferometric observations of υ Sgr
- The description of the genetic algoriths
- The results
- Conclusion

Introduction

υ Sgr

- single-line spectroscopic binary
- peculiar spectrum with emission lines
- infrared excess, 2 BBs approximation of the SED
- $\rightarrow$ presence of the dust shell
- VLTI/MIDI interferometric observations
    - summer 2007 + May 2008
    - 12 visibility measurements (UTs 2, ATs 10)

MC3D

The code

- continuum RT, based on the Monte-Carlo method
- emitting, scattering, absorbing and reemitting photons from the central source on the spherical dust grains
- input: geometry, parameters of the model, dust catalogue
- output: model, SED, spatial & spectral brightness (polarization maps)

MC3D

The geometry

$$\varrho(r,z) = \varrho_{100}\left(\frac{100}{r}\right)^{\alpha} \exp\left[-\frac{1}{2}\left(\frac{z}{h(r)}\right)^2\right]$$

$$h(r) = h_{100}\left(\frac{r}{100}\right)^{\beta}$$

Input parameters

- the source: $\underline{d}$, $\underline{T}$, $\underline{L}$ (assuming BB approximation)
- the geometry of the disk: R_{in}, $\underline{R_{\mathrm{out}}}$, i, α, β, h_{100}
- the dust properties: chemical compostion, size distribution, total mass of the dust

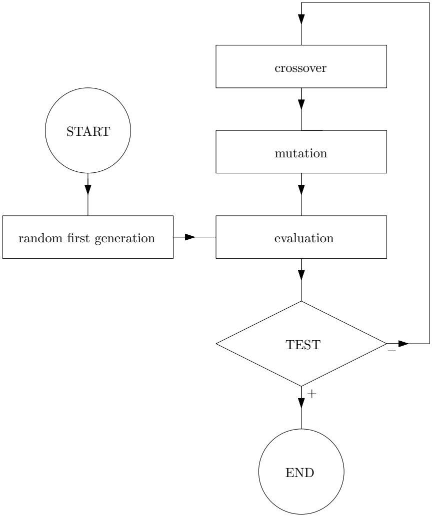
Parameter space

parameter	min	max	step	n(step)
α	1.80	2.50	0.10	8
β	0.70	1.50	0.10	9
h	2.0	7.5	0.5	12
$R_{\rm in}$	2.0	7.5	0.5	12
i	30	75	5	10
$\log M_d$	-7.0	-3.5	0.5	8
silicate / am. C.				6

The parameter space of the MC3D models:

- $5 \cdot 10^6$ possible combinations of parameters
- $\sim$ 800 years of computation on 3GHz 1CPU PC.

The scheme of the GA



The first generation of the models

First generation of the models

- each generation has n models
- random selection of the $n \cdot k$ parameters that we want to find
- $\rightarrow$ n models of 1st generation

$$M_{1,j} = (p_{1,j,1}, \ldots, p_{1,j,k}),\ j \in (1, n)$$

Evaluation

- weight: *the ability to survive* (fitness function)

$$w_{1,j} = \left(\chi^2_{1,j}\right)^{-1}$$

Crossover

We have models and the corresponding weights now.

$$M_{i,j} \qquad w_{i,j} \qquad j \in (1, n)$$

(i ... the number of the generation)
Selection of the n pairs of the models for the crossover

$$M_{i,a}, M_{i,b} \qquad a, b \in (1, n)$$

$$P(M_{i,(a,b)} = M_{i,j}) = \frac{w_{i,j}}{\sum_j w_{i,j}}$$

Crossover probability $p_c \sim 0.95 - 0.99$

$$x_{i,j} \in (0, 1) \qquad j \in (1, n), P(x_{i,j} = 1) = p_c$$

$x_{i,j} = 1$: j-th pair of the models undergo the crossover
$x_{i,j} = 0$: one of the models of the j-th pair pass directly to the mutation

Crossover

In the case of the crossover ...

$M_{i,a}$	$p_{i,a,1}$	$p_{i,a,2}$	$p_{i,a,3}$	$p_{i,a,4}$	$\cdots$	$p_{i,a,k}$
$M_{i,b}$	$p_{i,b,1}$	$p_{i,b,2}$	$p_{i,b,3}$	$p_{i,b,4}$	$\cdots$	$p_{i,b,k}$
$c_{i,j}$	1	1	2	1	$\cdots$	2
$M_{i+1,j}$	$p_{i,a,1}$	$p_{i,a,2}$	$p_{i,b,3}$	$p_{i,a,4}$	$\cdots$	$p_{i,b,k}$

$M_{i,a}$, $M_{i,b}$ – the "parents", 2 models selected from the i-th generation
$C = c_{i,j}$ – crossover matrix for the i-th generation, j-th pair of "parents"
$M_{i+1,j}$ – the "child"; will become the member of $i+1$-th generation after mutation

Mutation

From the previous steps we have:

- n models $M_{i+1,j}$ with k parameters $p_{i+1,j,l}$

Mutation

- probability of the mutation $p_m \sim 0.01 - 0.05$

$$m_{i+1,j,l} \in (0,1) \qquad j \in (1,n), l \in (1,k), P(m_{i+1,j,l} = 1) = p_m$$

If...

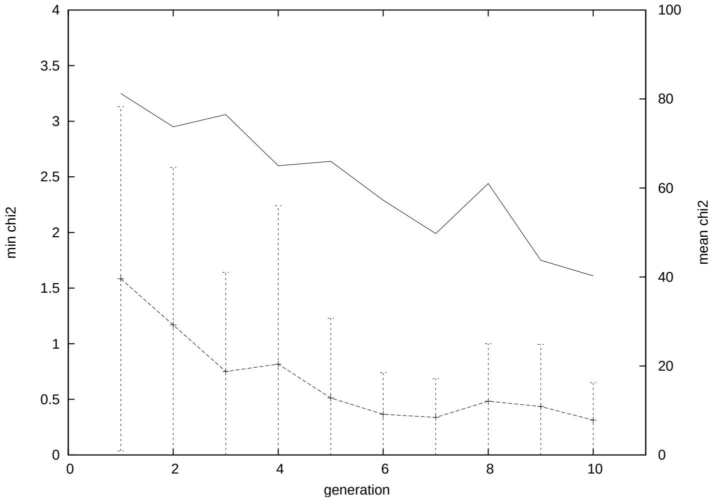
- $m_{i+1,j,l} = 1$
  $\rightarrow$ the parameter $p_{i+1,j,l}$ is replaced with new, random value
- $m_{i+1,j,l} = 0$
  $\rightarrow$ nothing happens

We have the new generation models now! $i + 1 \rightarrow i$

New generation

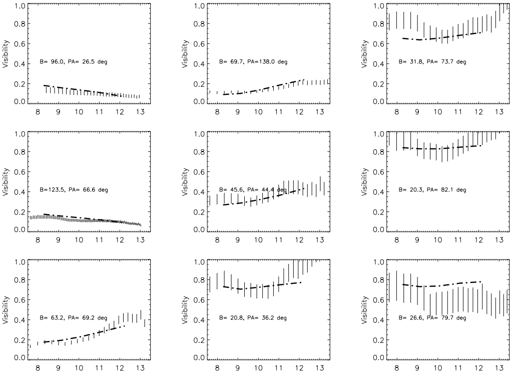
- computation of the new models
- evaluation of the new models ($w_{i,j}$)
- A) do the models fit our criteria?
- B) do the models still evolve? (aren't the models degenerated?)
- A) & B) = NO $\rightarrow$ proceed to the next loop of the scheme

The results



The evolution of the mean and minimal χ^2.
$n = 96,\ p_c = 0.975,\ p_m = 0.05$

The result



B= 96.0, PA= 26.5 deg
B= 69.7, PA=138.0 deg
B= 31.8, PA= 73.7 deg
B=123.5, PA= 66.6 deg
B= 45.6, PA= 44.4 deg
B= 20.3, PA= 82.1 deg
B= 63.2, PA= 69.2 deg
B= 20.8, PA= 36.2 deg
B= 26.6, PA= 79.7 deg
Visibility

The comparison of the results

parameter	new	old
d [pc]	595	513
$R_{\rm in}$ [AU]	$6.0^{+0.5}_{-1.5}$	$4.0^{+2.0}_{-1.0}$
i	$50^\circ{}^{+10^\circ}_{-20^\circ}$	$40^\circ \pm -15^\circ$
α	$2.0^{+0.5}_{-0.3}$	$2.4^{+0.1}_{-0.4}$
β	$0.7^{+0.3}$	$0.95^{+0.05}_{-0.3}$
h_{100} [AU]	$3.5^{+2.0}_{-1.5}$	3.0 ± 2.0
$\log(M_{\rm d}/M_\odot)$	$-3.5_{-3.0}$	-6.0
$M_{\rm am.C}/M_{\rm d}$	$0.6^{+0.2}_{-0.4}$	0.6
χ^2	1.51	2.92

Conclusion

GA works! But ...

- ...they are efficient just for searching in huge parameter space
- ...they have problems with searching the precise solution
- ...they need large number of evaluated models

However ...

- ...they do not require much knowledge about the system
- ...you will find at least some solution
- ...it's unlikely to find just local minima
- ...their results are quite good for huge parameter spaces
- ...they can be adopted for large number of problems

It can be shown that the number of models that are ...

- ...better than average increases exponentially with time
- ...worse than average decreases exponentially with time

Reference

- Šíma J., Neruda R.: Theoretical Issues of Neural Networks, 1996
  http://www2.cs.cas.cz/~sima/kniha.html (in Czech)