Artificial Intelligence 2, Lecture 1, 20101109

Course introduction and motivation

Biological basis of evolutionary algorithms

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

What will you learn? (Aims)

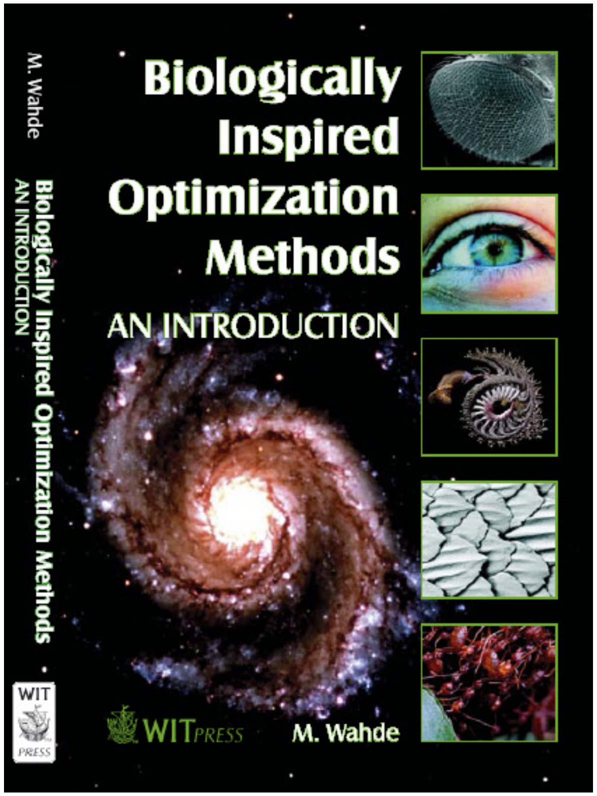
- A brief review of classical optimization methods.
- The basics of several *stochastic optimization methods* (evolutionary algorithms, particle swarm optimization, ant colony optimization).
- The basic biology and physics behind the methods.
- How to *implement* and *apply* the methods, both in simple, straightforward cases and in more complex problems.
- How to select which method to use for a given problem.

Why should you take this course?

- Stochastic optimization algorithms can be used for solving many problems where classical methods are insufficient.
- Stochastic optimization methods are used more and more frequently in industry, particularly in large problems.
- The number of application areas is steadily increasing.
- (Most important!) It is interesting and fun to work with stochastic optimization! ☺

...about the book:

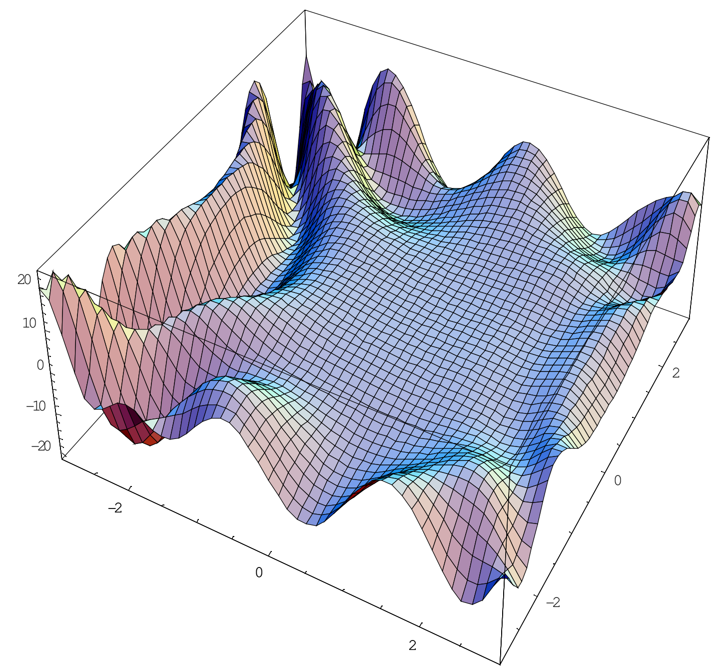
- Available at Chalmers' bookstore (Cremona)
- Unfortunately, it's a bit expensive...
- ..I have negotiated with the publisher to minimize the price that you will have to pay (meaning that I get no income at all from the books sold at Cremona, if that's any comfort.. ☺)
- You may wish to buy the book online.
- Note, however, that Cremona sells the book at a reduced price.



Optimization

- In general, optimization is the problem of finding the (global) minimum (or maximum) of an *objective function*.

- Sometimes (but not always!) the objective function is a specific, well-defined mathematical function, $f = f(x_1,x_2,...)$.



pp. 1-2

Optimization methods

- Optimization methods can be divided into two broad categories: *classical (deterministic) optimization methods* and *stochastic optimization methods*.
- Classical optimization methods include, for example, gradient descent (following the steepest slope), Newton's method, penalty methods, lagrange multiplier methods etc.

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

Limitations of classical optimization

- Classical methods are less useful in cases with
  - non-differentiable objective functions
  - objective functions whose values can only be obtained as a result of a (lengthy) simulation
  - varying number of variables (as in optimization of neural networks).

- For such problems, *stochastic optimization methods* are more suitable. This course mainly concerns such methods.

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

Stochastic optimization methods

- As the name implies, stochastic optimization methods contain an element of stochasticity (randomness).

- Many (but not all) stochastic optimization methods are inspired by biological phenomena.

- Thus, an important subset of stochastic optimization methods are *biologically inspired optimization methods*.

Biologically inspired optimization

- Why would one use inspiration from biological phenomena when defining an optimization method?

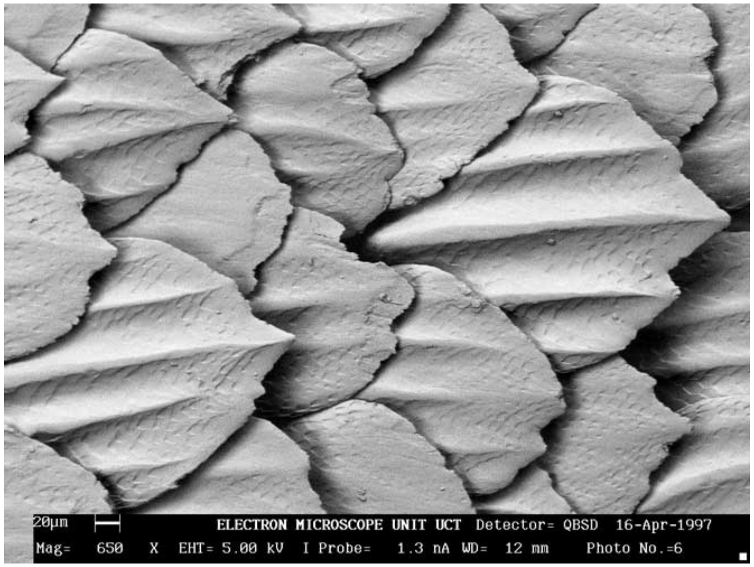
- ...Nature is all about *adaptation*, which can be seen as a kind of optimization.

- However, note that, in nature, the target is constantly moving – unlike the case in engineering problems, evolutionary adaptation takes place in a varying fitness landscape (more about this later).

pp. 2-5

Example 1: Shark skin

- A good example of adaptation (biological optimization).
- The skin of sharks has evolved to allow very fast swimming.



pp. 2-5

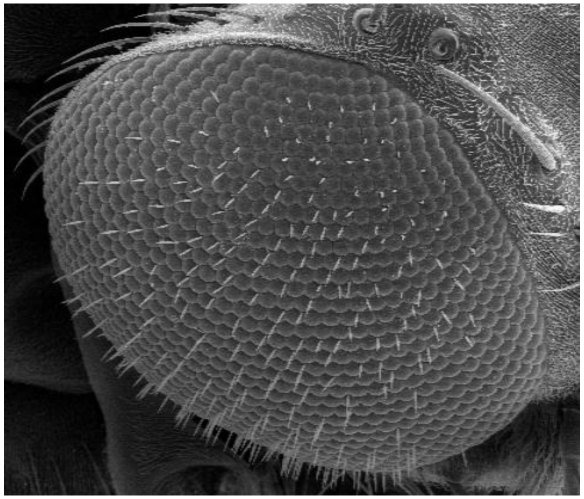
Example 1: Shark skin

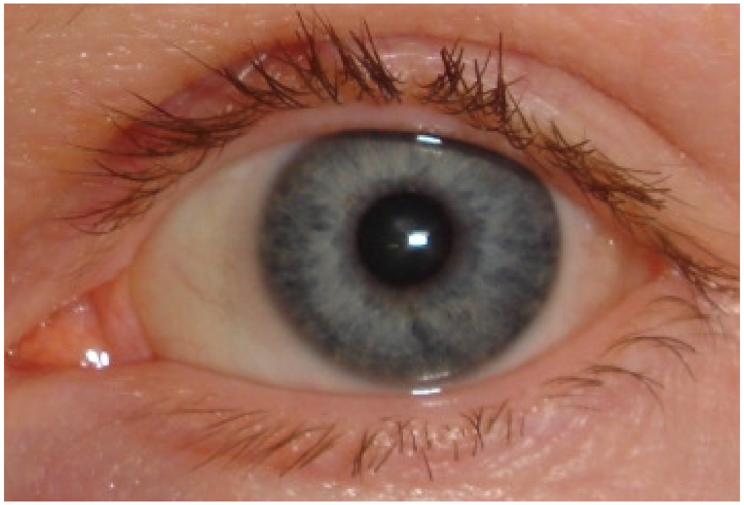
- The rib-like structures affect the interaction between the surface of the shark and the surrounding water, essentially reducing drag.
- Prime example of biological inspiration for engineering concepts: Similar ideas are being applied in order to reduce drag (and noise) of aircraft and other vehicles.
- Shark skin has also inspired the development of the fabric used in swimsuits worn by athletes.

pp. 2-5

Example 2: Evolution of the eye

- The evolution of eyes is another intersting example.
- Faced with the problem of generating light-gathering devices, evolution has come up with no less than 40 *completely independent* solutions to the problem.





pp. 2-5

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

Example 3: Swarming

- Some species (particularly ants, bees and termites) display very advanced forms of cooperation.
- Specific example: Weaver ants (Oecophylla)

pp. 2-5

Example 3: Swarming

- Many organisms (e.g. some bird and fish species) display swarming behavior (for protection against predators, efficient food gathering etc.)

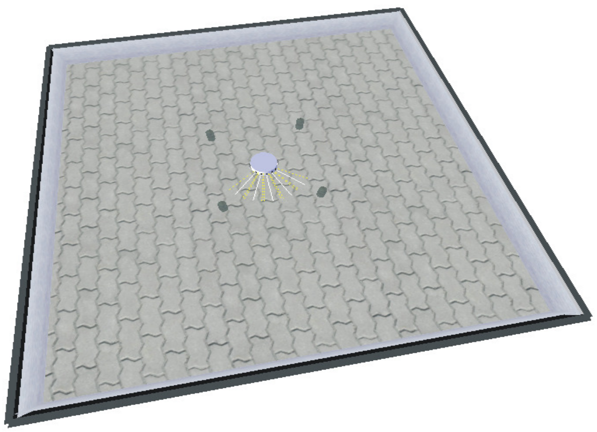
Biologically inspired optimization

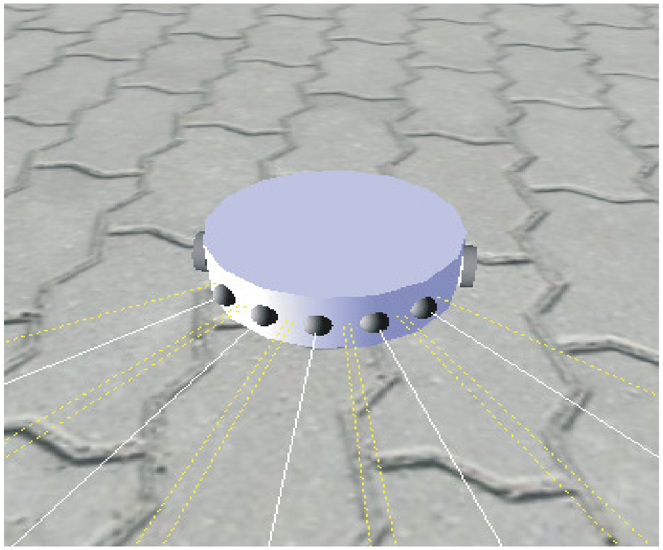
- Several different optimization methods have been developed, based on biological phenomena:
  - Evolutionary algorithms (inspired by evolution)
  - Ant colony optimization (inspired by cooperative behavior)
  - Particle swarm optimization (inspired by swarming)

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

Simple example: robot behavior

- Stochastic optimization methods can be used for the optimization of robot behaviors.
- Simple example: Evolution of a cleaning behavior:





pp. 5-8

Simple example: robot behavior

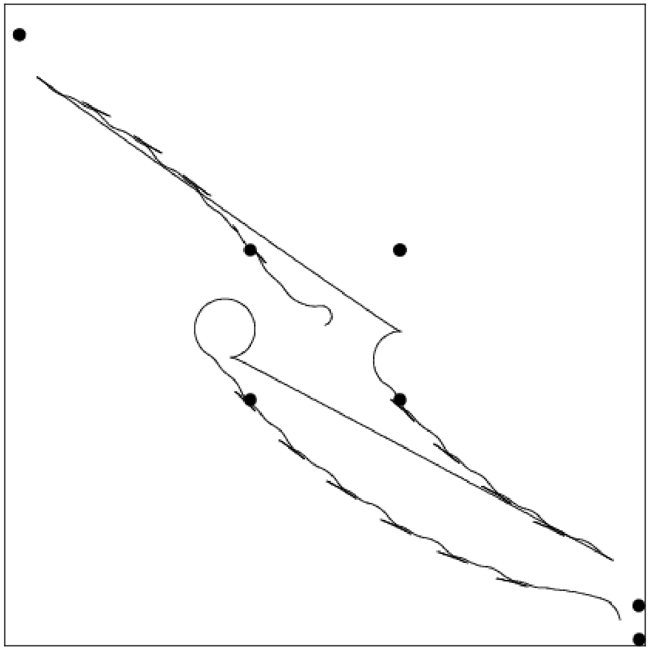
- The robot is required to move the small objects out to the edge of the arena.
- The brain of the robot consists of a sequence of *IF-THEN-ELSE rules*, using sensory readings as inputs.
- Only the final situation (at the end of an evaluation) is used as feedback to the optimization algorithm: (an evolutionary algorithm).

$$F = \sum_{j=1}^{n_o} \left(x_j^2(T) + y_j^2(T) \right),$$

pp. 5-8

Simple example: robot behavior

- The population of robots gradually evolves the capability to clean the arena, passing through several stages along the way.



pp. 5-8

Brief introduction to EAs

- Evolutionary algorithms = EAs.
- Based on darwinian evolution, which is a process involving *gradual*, *hereditary* changes of biological organisms, over long periods of time.
- Basic biological concepts: **Population**, **fitness**, **selection**, **mutation**.
- Individuals that are well adapted to their environment have a high probability of generating offspring.

pp. 35-98

Brief introduction to EAs

- Random changes to the *genotype*, (mutations) sometimes leading to large changes of the *phenotype.*
- Mutations provide new material for evolution to work with.
- Note that, while mutations are random, *selection* is not.

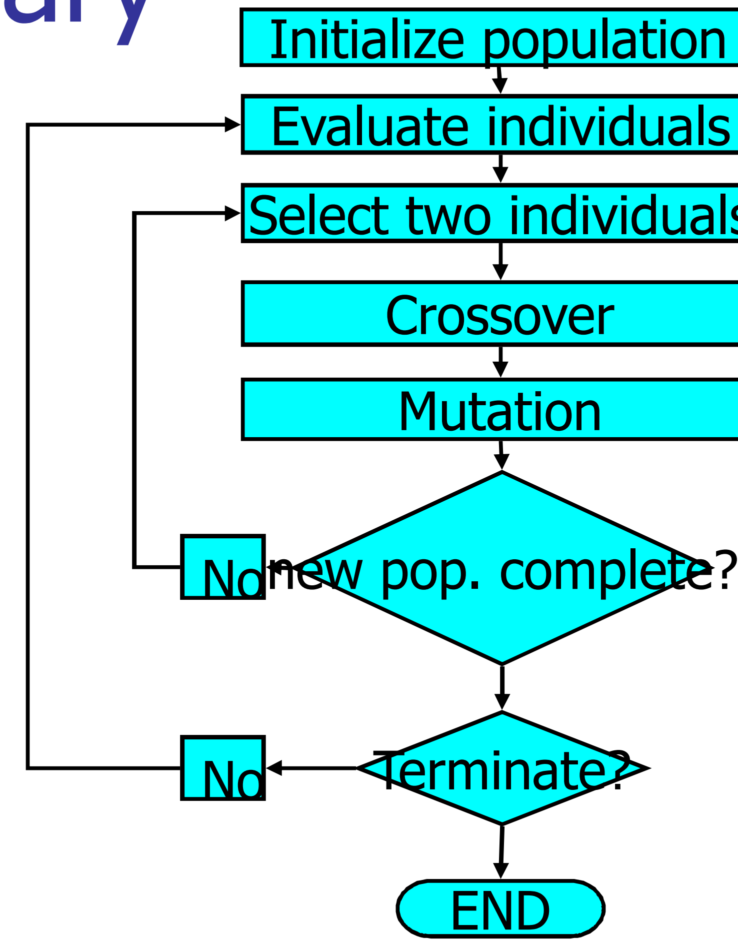
pp. 35-98

Basic mode of operation of EAs

- In an EA, a **population** of candidate solutions (to the problem at hand) is formed (random initialization)
- Each **individual** in the population is evaluated and assigned a **fitness score** based on its performance.
- New individuals are formed through the processes of **selection**, **crossover**, and **mutation**.
- The new individuals form the second **generation**, which is evaluated in the same way as the first generation etc.
- The process is repeated until a satisfactory solution has been found to the problem at hand.

pp. 35-98

Summary

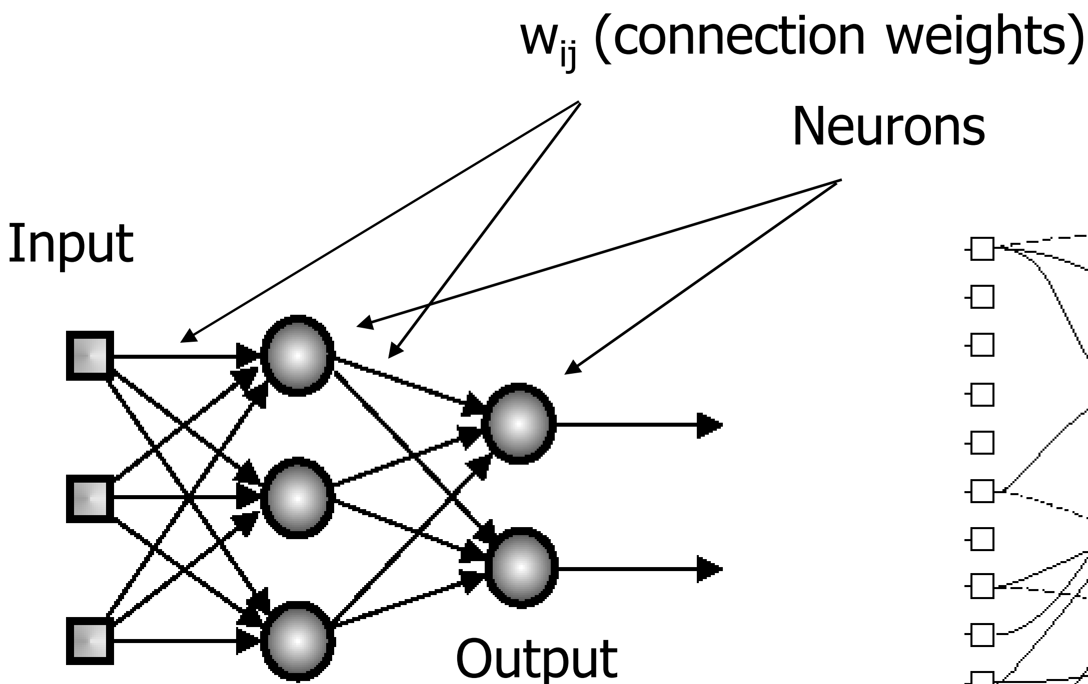


pp. 35-98

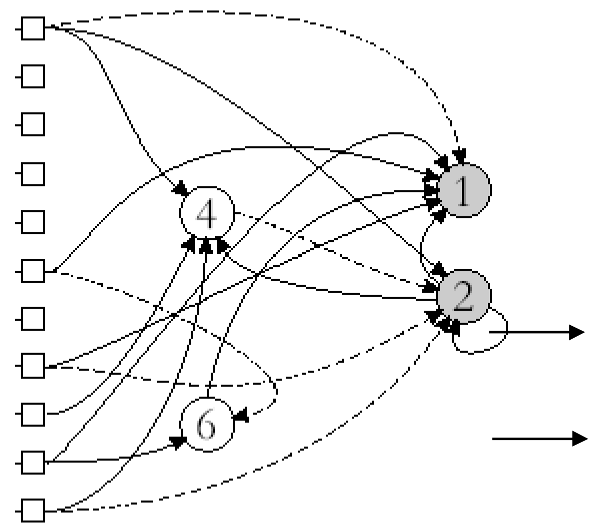
Brief introduction to artificial neural networks (ANNs)

- Based on the properties of the brain.
- Consist of a large number of simple computational elements (neurons) that are connected and together can perform complex calculations.
- The human brain consists of around 10^{11} neurons, and 10^{15} connections between the neurons.
- In ANNs, the computation is determined by the architecture of the network as well as the values of the connection weights connecting different neurons to each other.

Typical ANNs:



Feedforward neural network (FFNN)

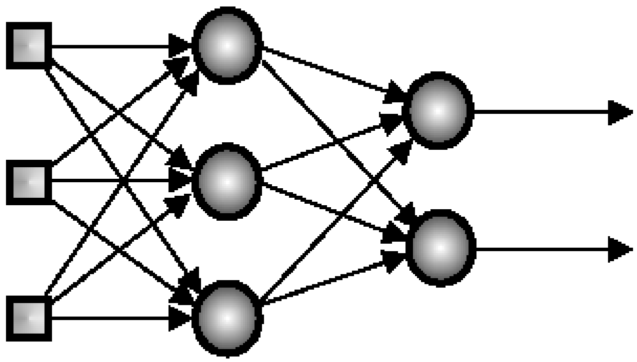


Recurrent neural network (RNN)

The computation in an ANN is integrated with its structure.

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

Training ANNs



- Information is stored in the connection weights.
- The procedure of setting the weights in an ANN is referred to as **training**.
- Several methods for training exist, e.g. the **delta rule** (for one-layer networks), **backpropagation** (for multi-layer networks) and **EAs** (for any type of network).

Properties of neural networks

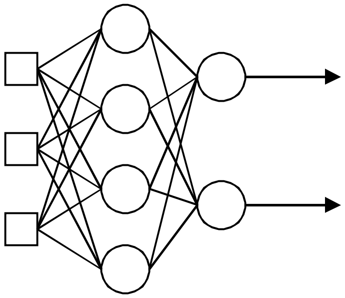
- **Graceful degradation**: Some neurons can be removed without significant alteration of the (distributed) computation performed by the network.

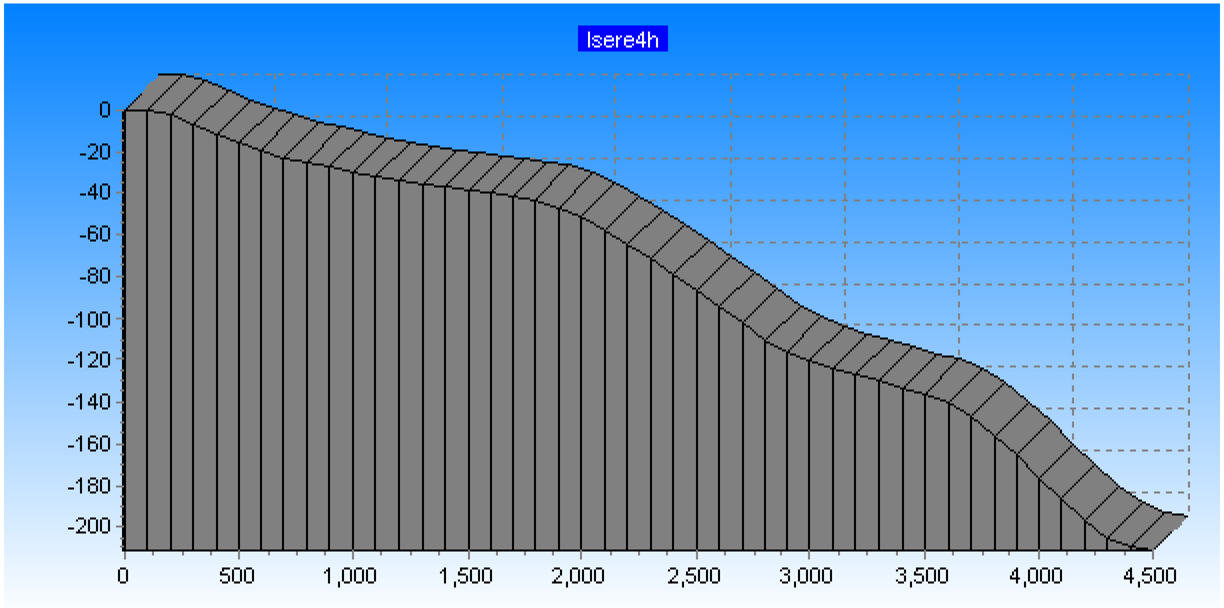
- **Generalization**: A network trained on a given data set is often able to generalize, in the sense that it can provide adequate output for new (previously unseen) input data.

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

Application example 1

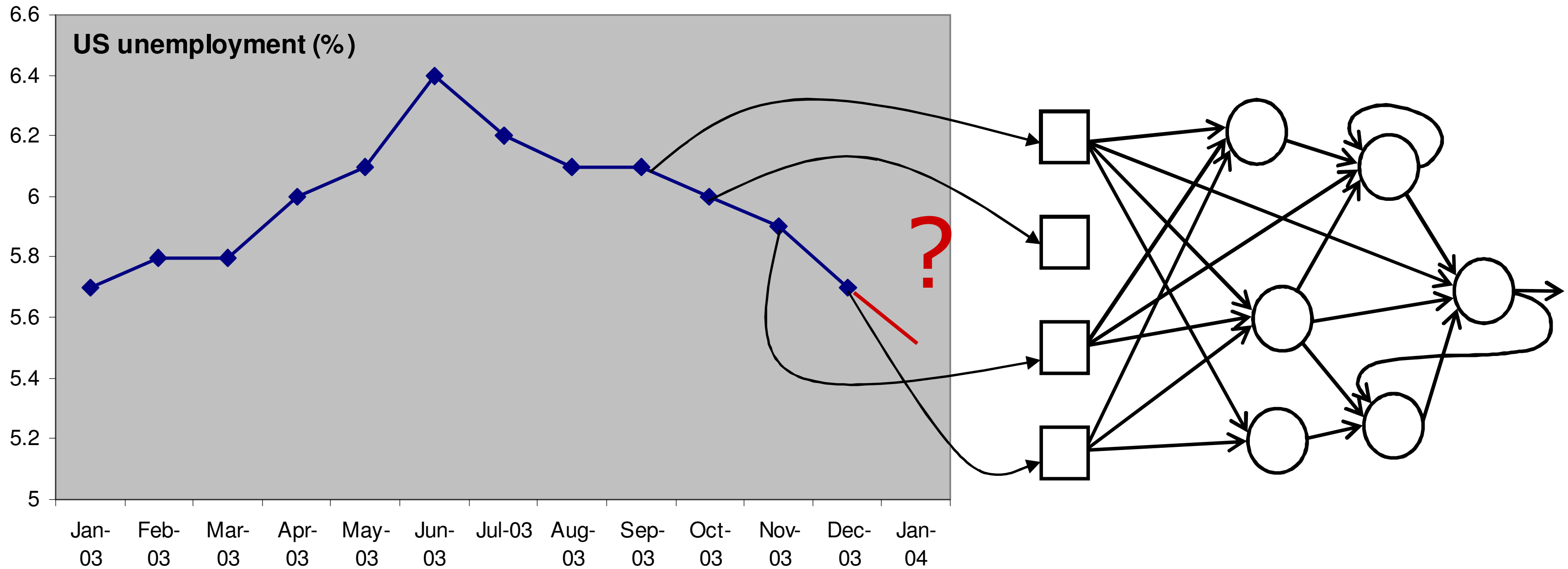
- **Optimization of braking systems in trucks** (using EAs and ANNs)





Application example 2

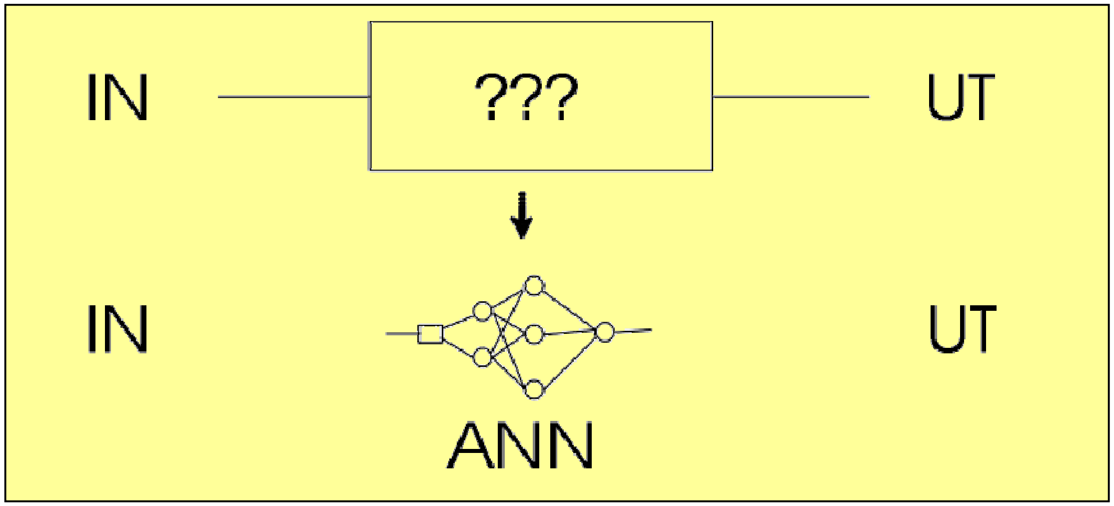
- **Prediction of time series:** The task is to make a prediction of the next value in a data set. Typical applications: finance, macroeconomics, meteorology etc.



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

Application example 3

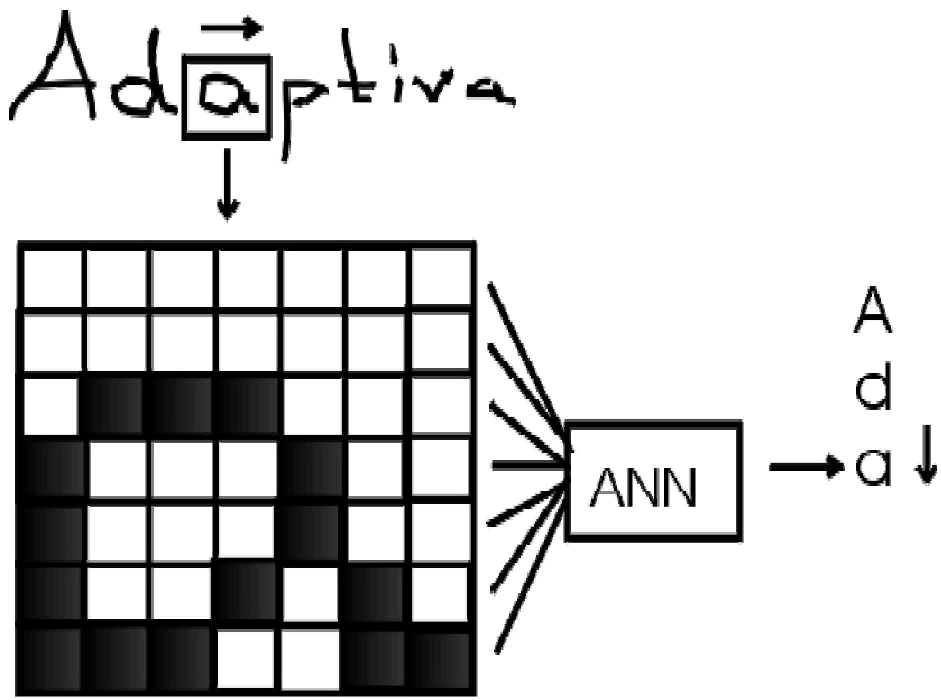
- **System identification** (model-free) of e.g. mechanical systems



- ANNs are often used in cases where it is difficult to find an analytical model.

Application example 4

- Image recognition



Brief introduction to ACO

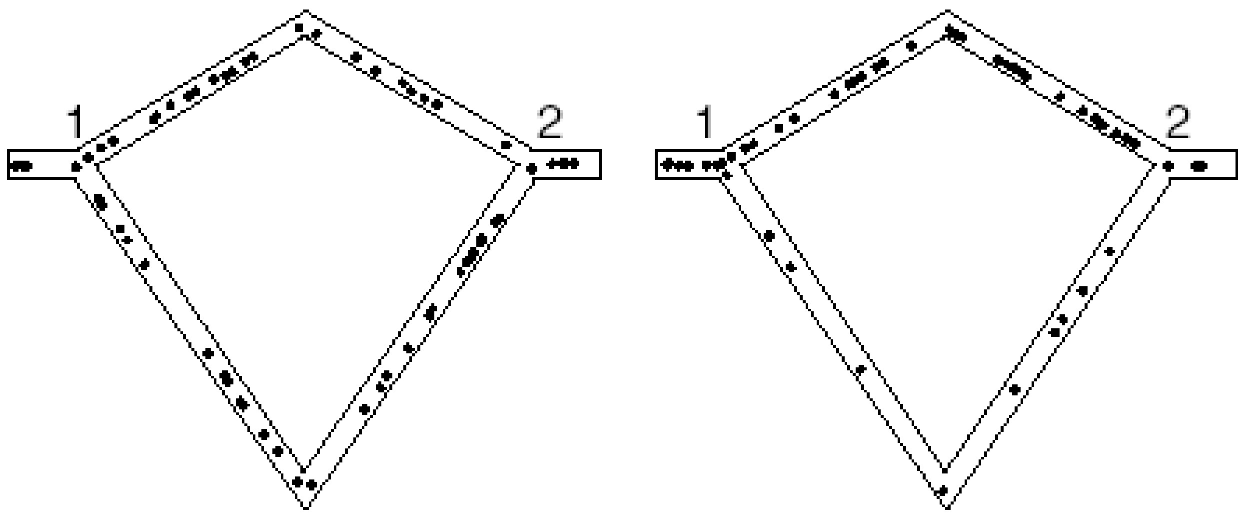
- Ant colony optimization = ACO.
- Ants are capable of remarkably efficient discovery of short paths during foraging. How do they achieve this?
- Note that ants (1) are (almost) blind, (2) have no explicit leaders
- Method: Ants deposit a trail of (volatile) *pheromones* (a chemical substance) as they move. When choosing a path, ants tend to move in the direction of highest pheromone concentration.

pp. 99-116

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

Experiment

- By laying artificial pheromones, the ants soon discover the shortest path:



pp. 99-116

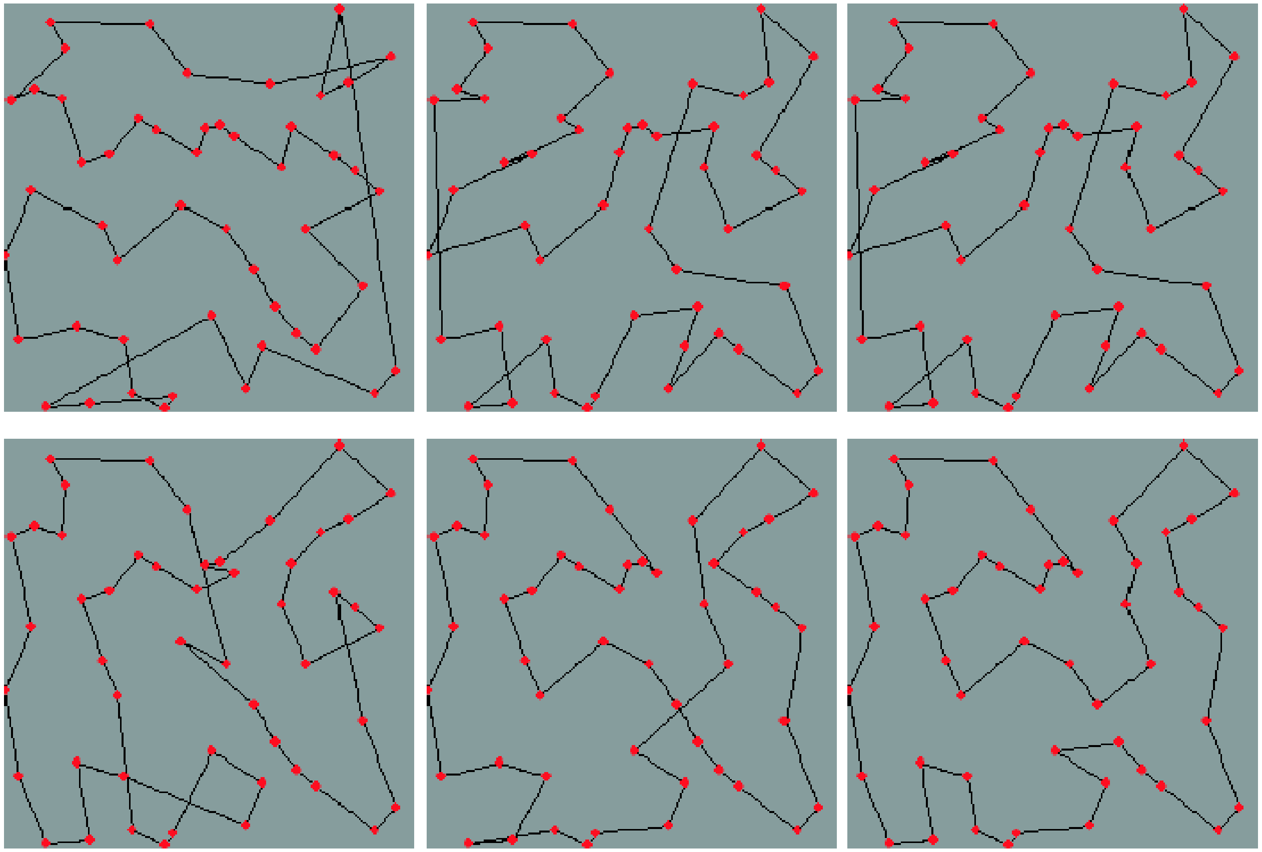
Brief introduction to ACO

- The behavior of the ants is an *emergent property* of the ant colony as a whole.
- In ACO, the same principles are used: Artificial ants deposit artificial pheromones, while attempting to find the shortest path in a graph.
- Many problems can be mapped to the problem of finding short paths. The most straightforward example is the travelling salesman problem (TSP).

pp. 99-116

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

Brief introduction to ACO

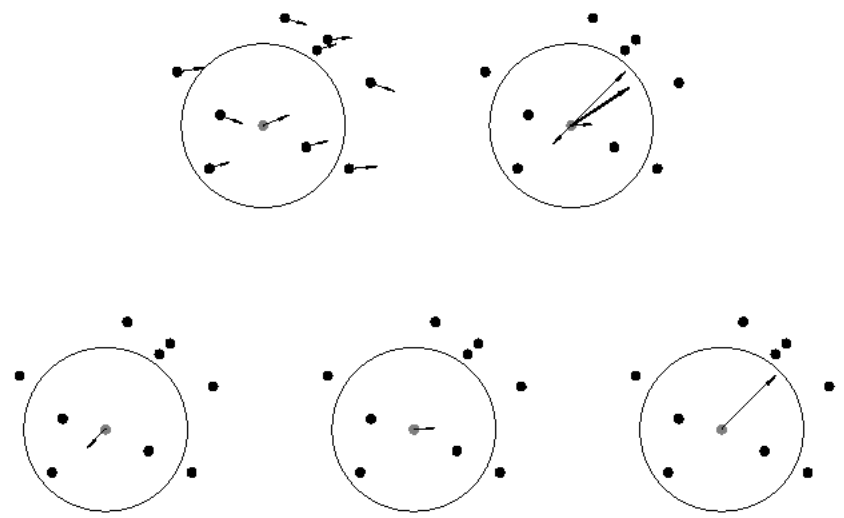


pp. 99-116

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

Brief introduction to PSO

- Particle swarm optimization = PSO
- Based on swarming behavior
- The *Boids* model: (that inspired the development of PSO)
- The movement of artificial bird-like objects (boids) is determined by three steers: *cohesion*, *alignment* and *separation*.



Brief introduction to PSO

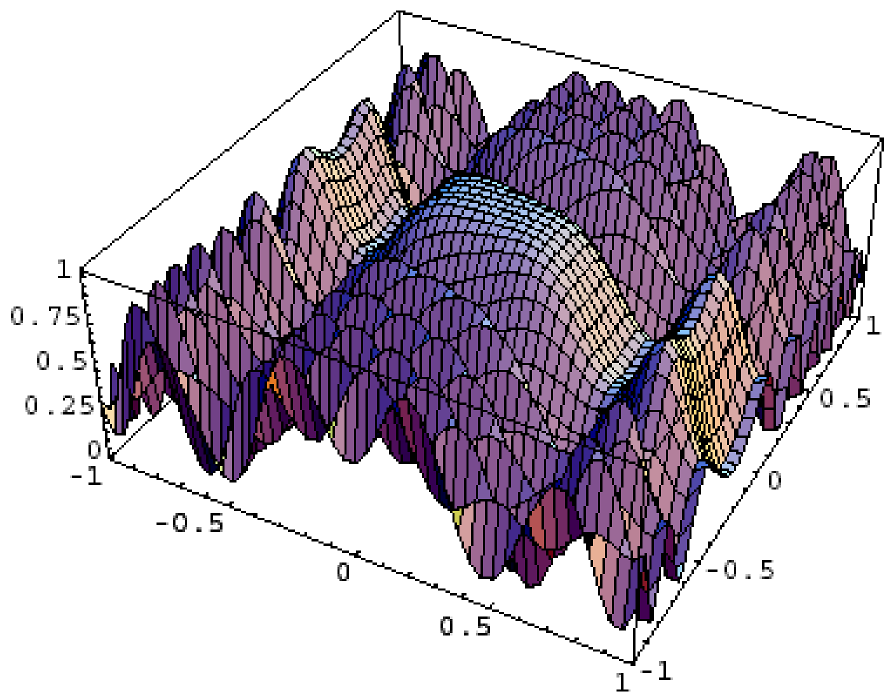
- PSO is based on the same principles: A population of particles move in the *n*-dimensional search space.
- The velocities of the *particles* in the swarm are updated according to

$$v_{ij} \leftarrow v_{ij} + c_1 q \left(\frac{x_{ij}^{\mathrm{pb}} - x_{ij}}{\Delta t} \right) + c_2 r \left(\frac{x_j^{\mathrm{sb}} - x_{ij}}{\Delta t} \right), \; j = 1, \ldots, n,$$

- ...where x_{ij}^{pb} is the best position (best value of the objective function) obtained by particle i and x_i^{sb} is the best position obtained by any particle in the swarm

Applications of stochastic opt.

- Function optimization: (particularly in problems involving many local optima)

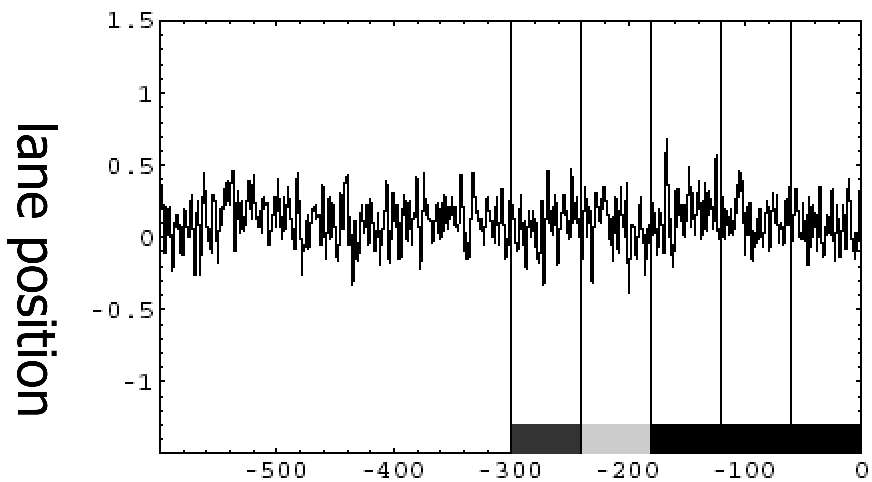


Applications of stochastic opt.

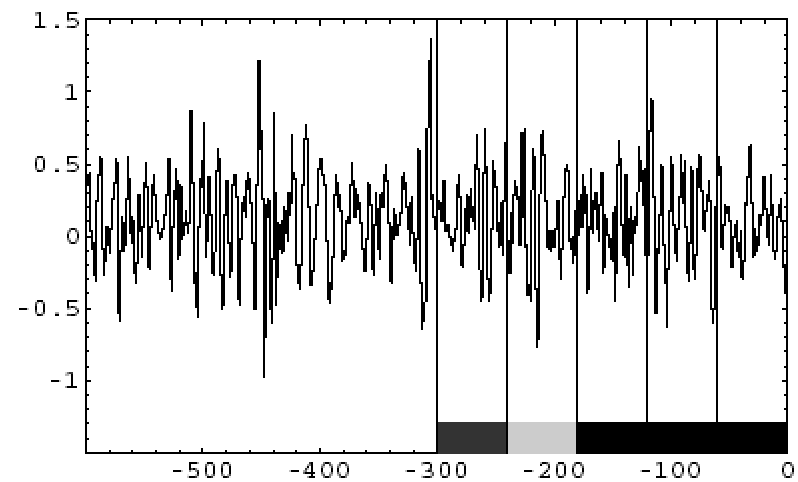
- Scheduling (e.g. airline crews)

Applications of stochastic opt.

- Classification-related problems, for example detection of sleepiness in car drivers based on time series of driving behavior:



Alert driver

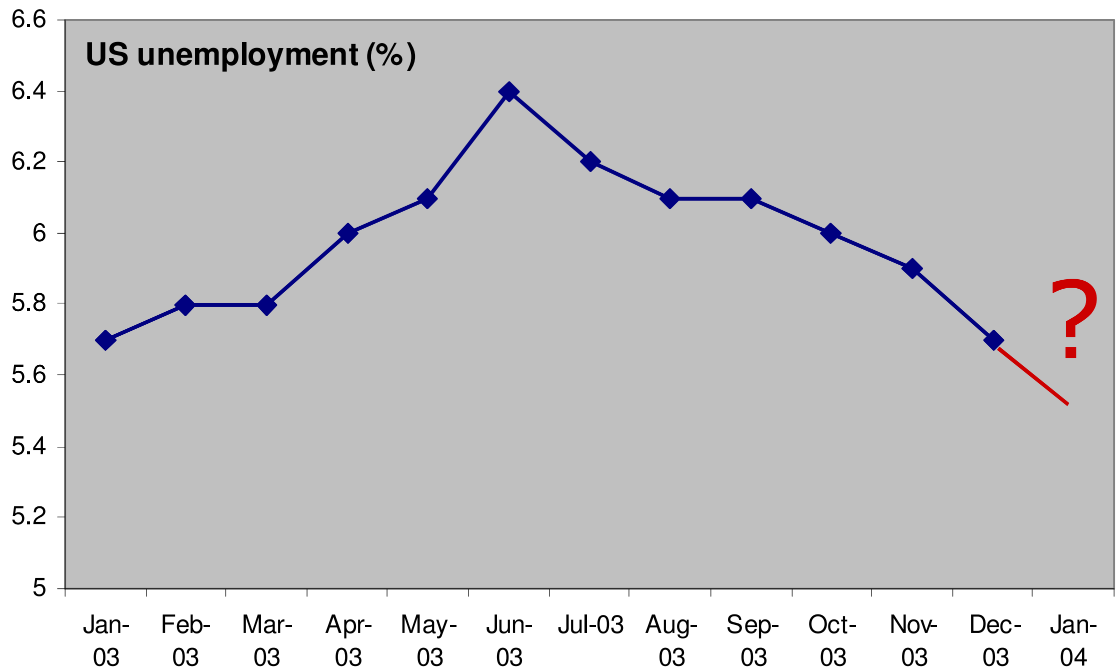


sleepy driver

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

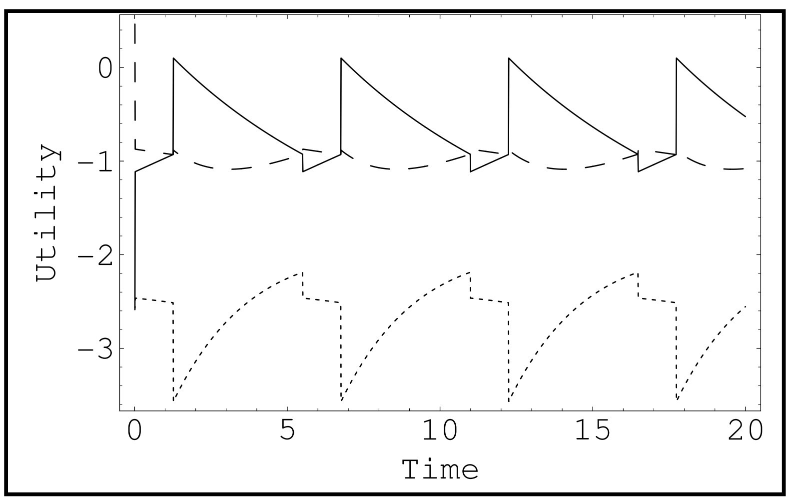
Applications of stochastic opt.

- Time series prediction: Applications in meteorology, seismology, epidemiology, finance, macroeconomics etc.



Applications of stochastic opt.

- Various problems concerning autonomous robots, e.g. behavior selection and decision-making.



Practical details

- There are two teachers: Mattias Wahde (MW) and Krister Wolff (KW)
- The course web page can be found at http://www.am.chalmers.se/~wolff/AI2/CoursePage.html
- Check the web page ***OFTEN***. There may be updates.

Practical details (schedule)

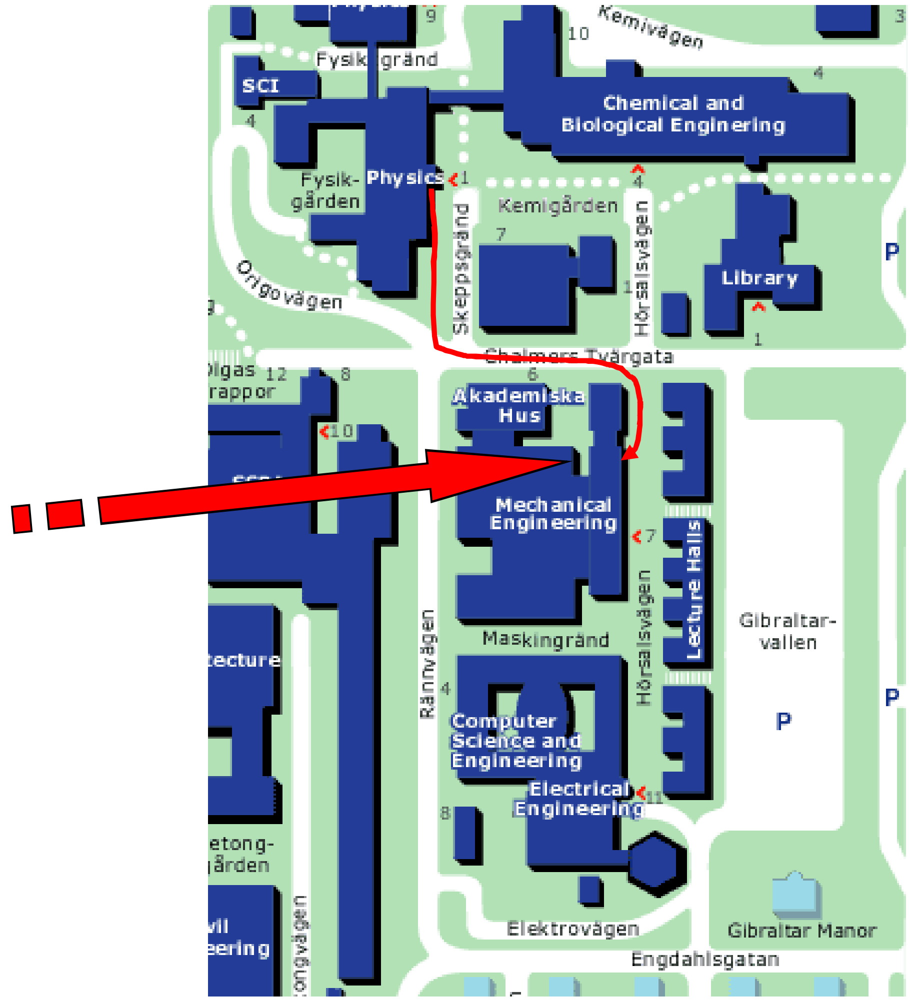
Date	Time	Room	Lecturer	Contents
20101109	13.00-15.00	Svea213	MW	Course introduction, biological basis of evolutionary algorithms (EAs)
20101111	13.00-15.00	Svea118	MW	Basics of EAs
20101116	13.00-15.00	Svea219	MW	Properties of EAs, I
20101118	13.00-15.00	Svea118	MW	Properties of EAs, II
20101123	13.00-15.00	Svea219	KW	Linear genetic programming (LGP)
20101125	13.00-15.00	Svea118	KW	Applications of EAs
20101130	13.00-15.00	Svea213	KW	Introduction to artificial neural networks (ANNs)
20101202	13.00-15.00	Svea118	KW	Ant colony optimization (ACO)
20101207	13.00-15.00	Svea213	KW	Particle swarm optimization (PSO)
20101209	13.00-15.00	Svea118	MW	Algorithm performance comparison, course summary

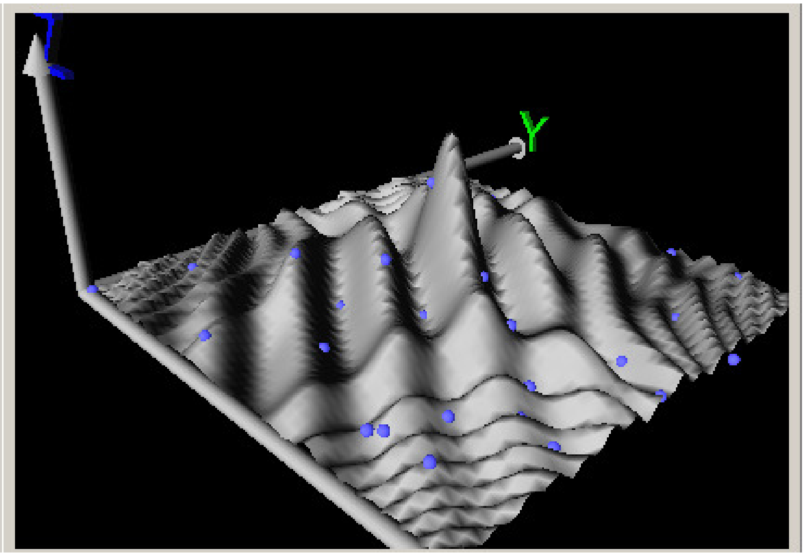
Practical details

- Examination:
  - Home problems (maximum 25p)
  - Exam (maximum 25p)
- The exam will take place on December 16, 2010, 08.30-12.30.
- Note: You need at least 10p *on the exam* to pass the course!

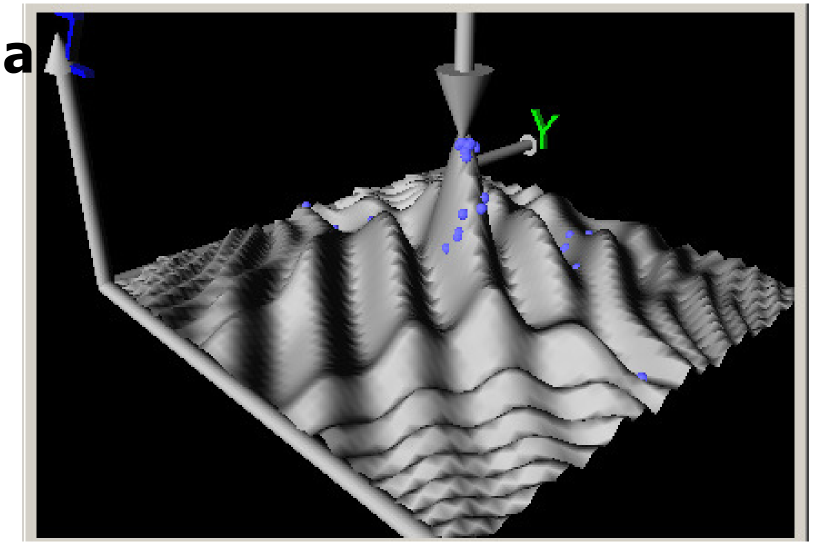
Office location

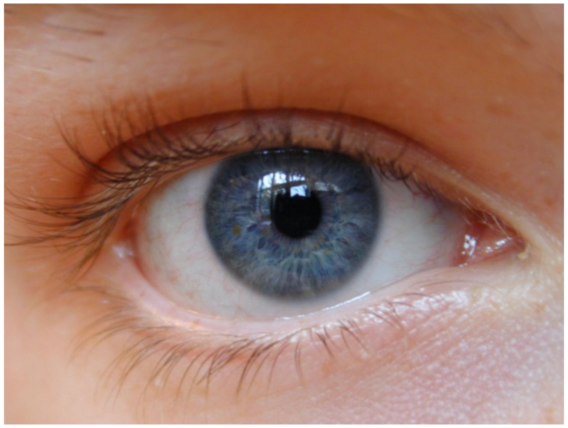
- Mechanical engineering building (nya M-huset), 1st floor.
- Enter near Café Bulten, follow the blue signs to ”Applied Mechanics” (Tillämpad mekanik)
- Dial my extension (3727) or Krister’s (3625) at the door.

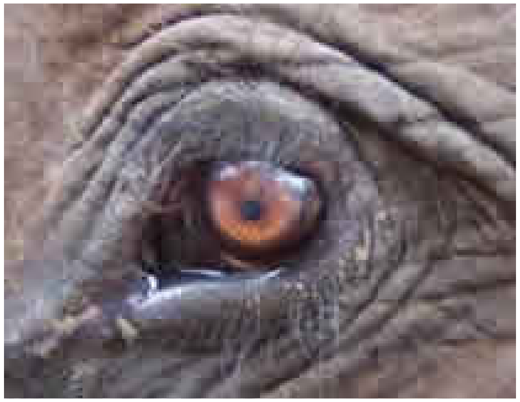


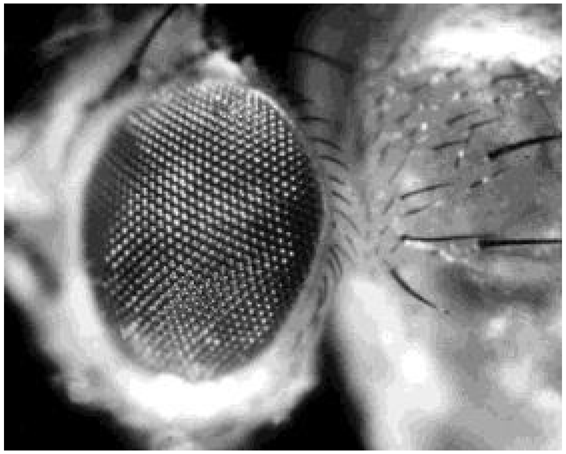


EAs are good at avoiding local optima!

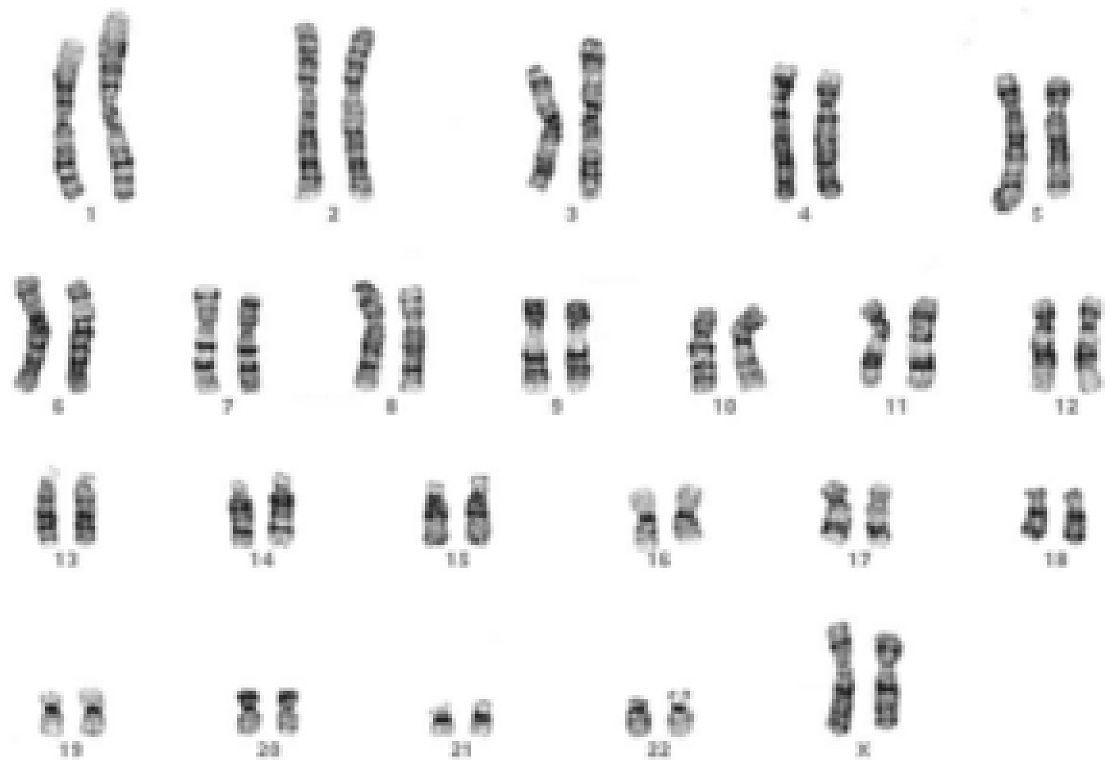






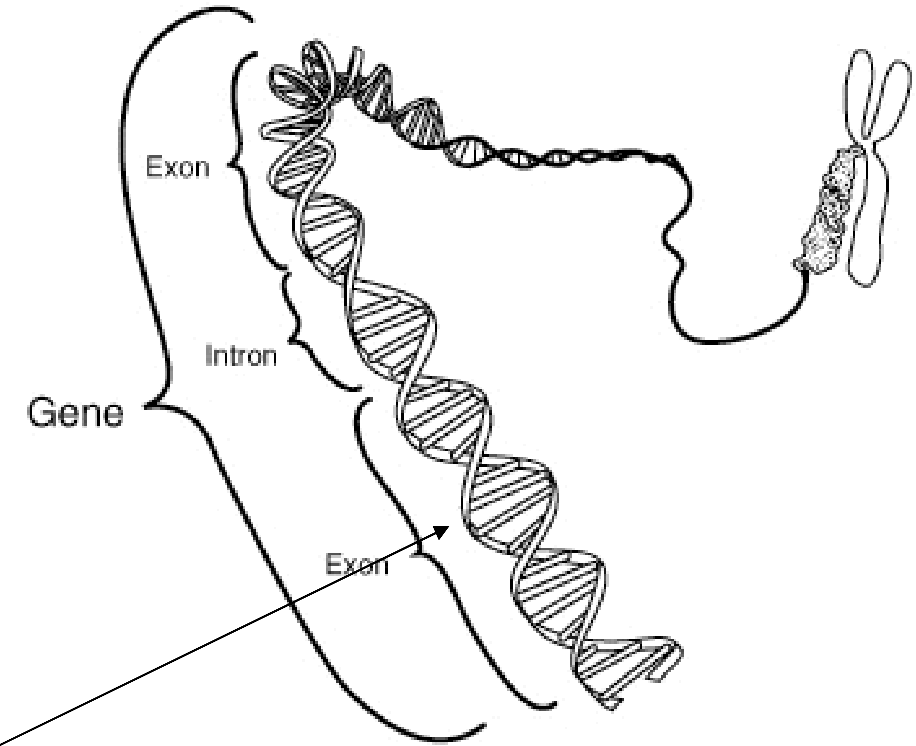


”Liger”



X

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



Sequence of base pairs:

...A-C-C-T-C-G-G-T-C-A-G...
| | | | | | | | | | |
...T-G-G-A-G-C-C-A-G-T-C...

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

Chromosome numbers in selected species

Fruit fly	8
Cat	38
Human	46
Ape	48
Horse	64
Dog	78
Carp	104

Number of genes in selected species

Bacterium	500-6,000
Yeast	6,000
Fruit fly	13,600
Human	~25,000

(However, there are many species (e.g. some fish) with more genes than humans)

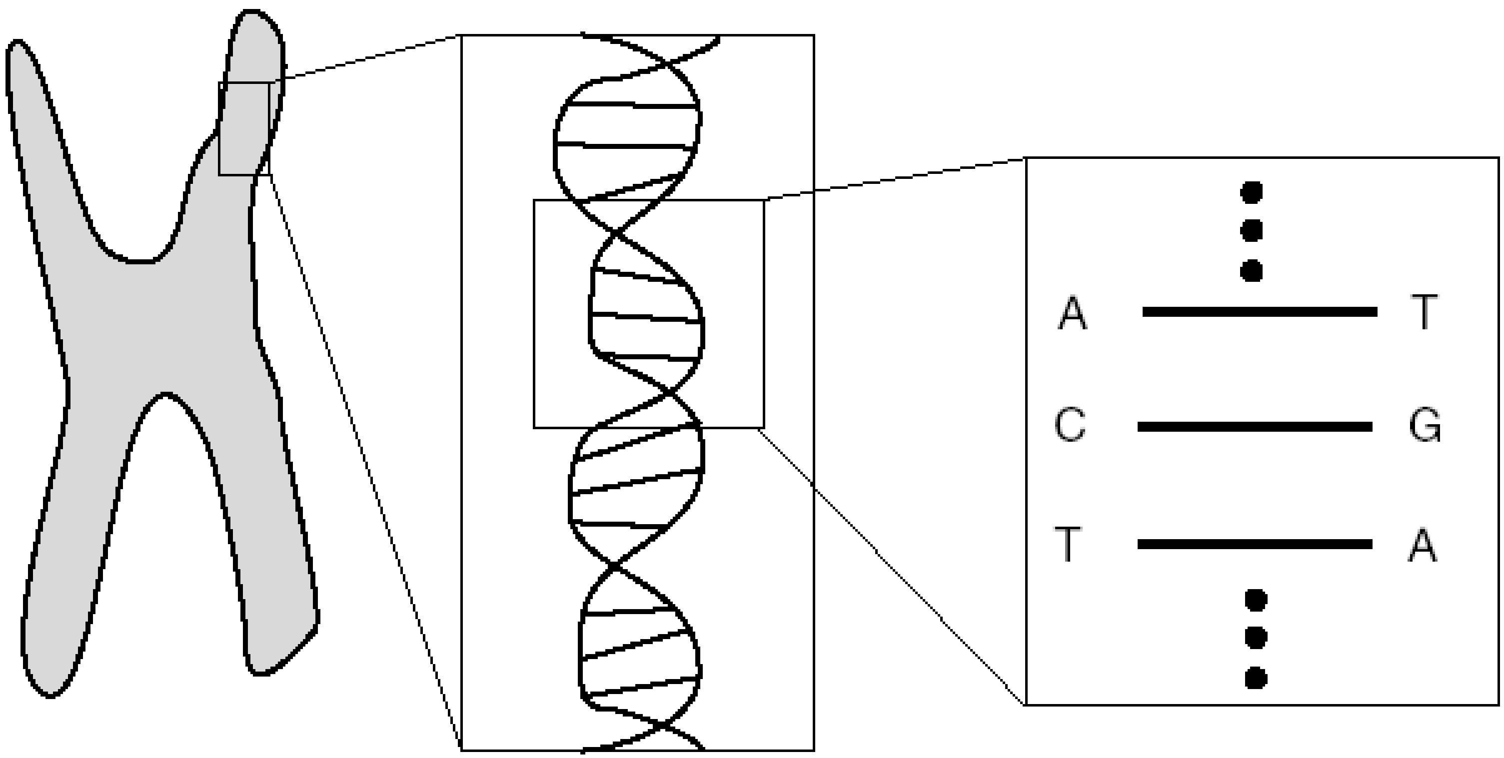
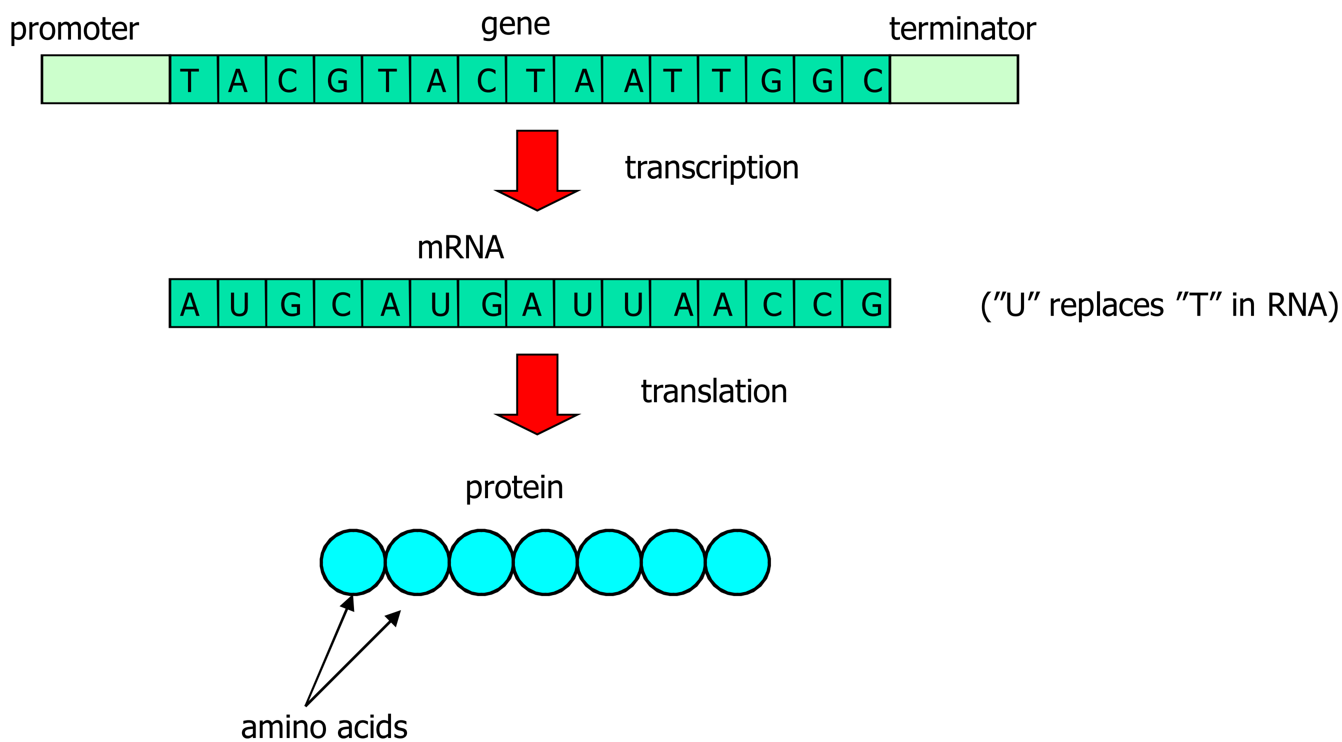


Figure 3.1: *A schematic representation of a chromosome is shown in the left side of the figure. The two blow-ups on the right show the individual base pairs. Note that A is always paired with T, and C is always paired with G.*

Genes are used for making proteins, through the two steps of *transcription* and *translation*:



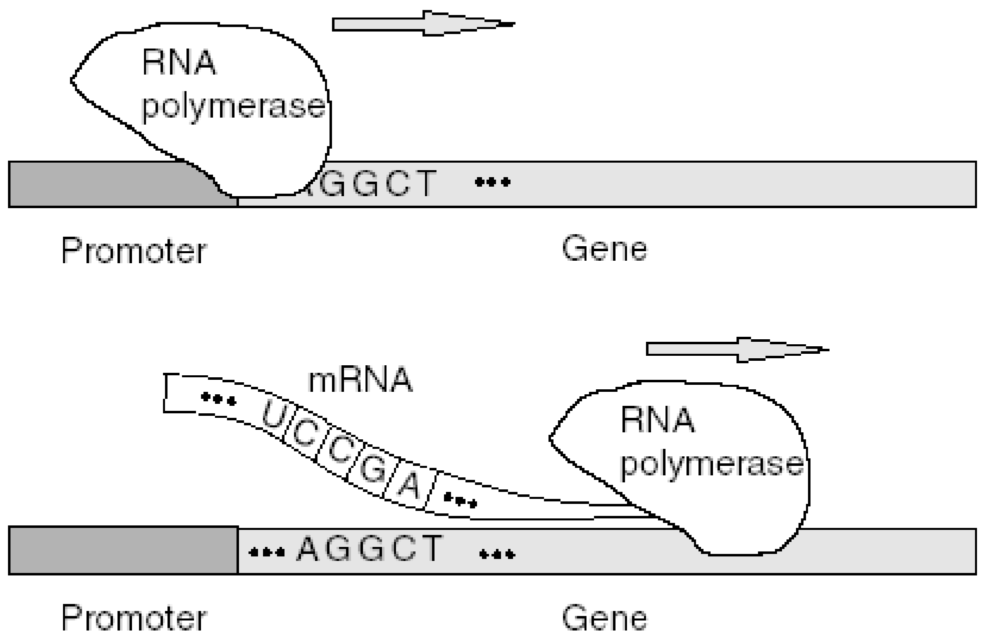
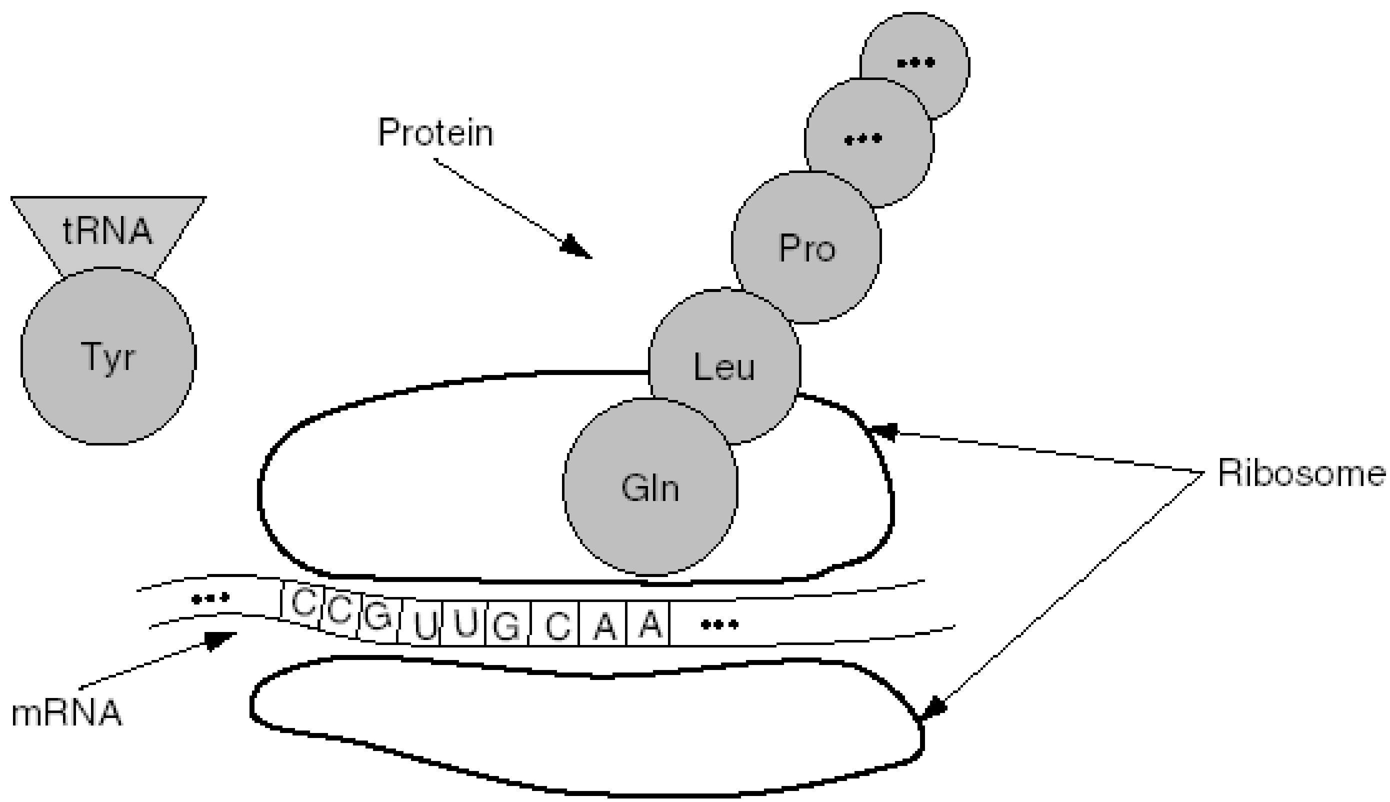
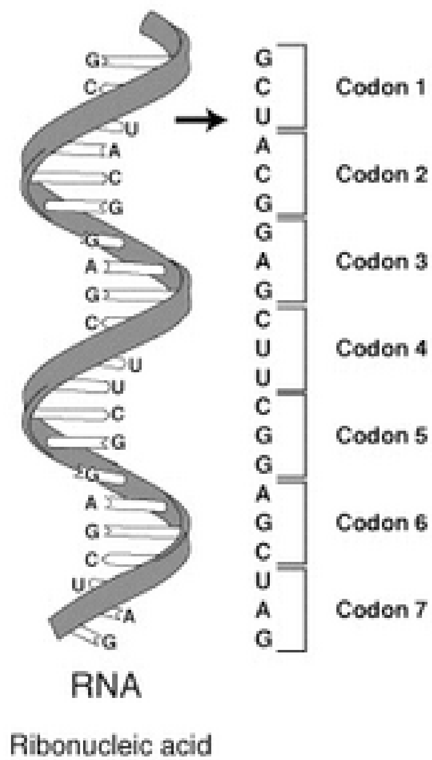


Figure 3.2: *The transcription process. The RNA polymerase binds to a promoter region and then moves along the DNA molecule, generating an mRNA molecule by joining bases available in the cell.*

Translation: (mRNA guiding the formation of proteins)



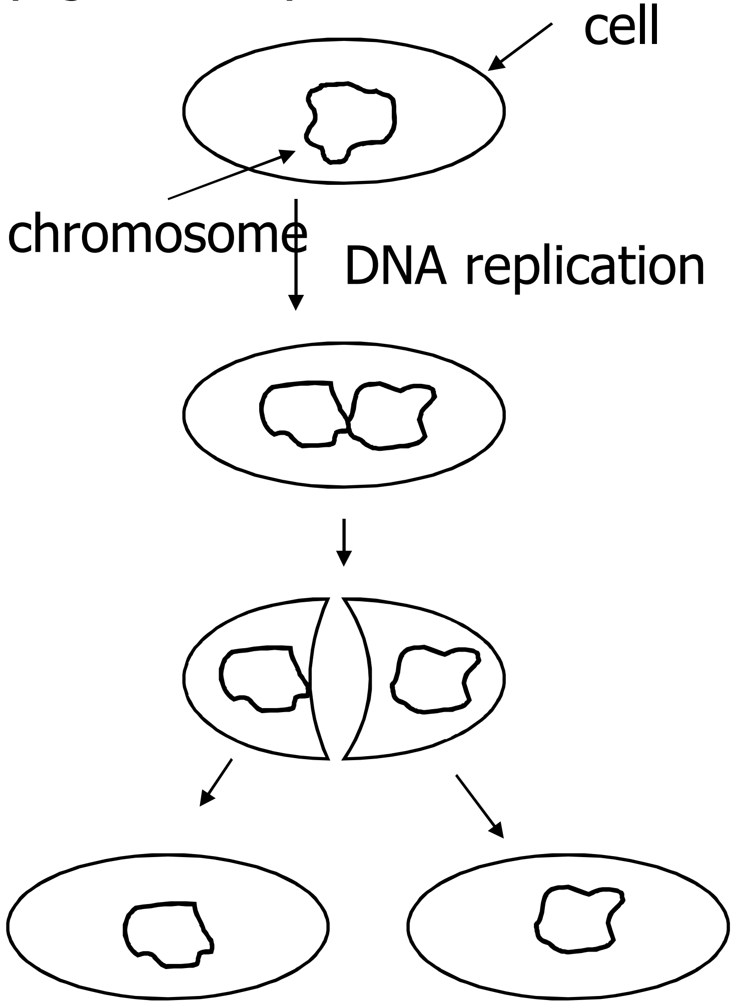


Second Position

First Position	U		C		A		G		Third Position
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C
	UUA	Leu	UCA	Ser	UAA	*Stop*	UGA	*Stop*	A
	UUG	Leu	UCG	Ser	UAG	*Stop*	UGG	Trp	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg	G
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	A
start	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg	G
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly	A
start	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly	G

(”U” replaces ”T” in RNA)

Asexual reproduction (e.g. bacteria)



1 0 1 1 1 0 1 0 0 0 1 0 1 1 0 0 0 1 0 0 0 0 0 1 1 0 0 0 1 1 0 1 0 1 1 0 0 | 1 0 1 0 0 0 0 1 0 1 1 0 0 0 1 1 0

0 1 0 1 1 0 1 1 1 0 1 0 1 1 1 0 1 1 0 0 0 1 0 1 0 1 1 0 1 0 0 1 1 0 1 0 1 | 1 1 1 0 0 0 1 0 1 0 0 1 0 1 1 0 1

1 0 1 1 1 0 1 0 0 0 1 0 1 1 0 0 0 1 0 0 0 0 0 1 1 0 0 0 1 1 0 1 0 1 1 0 0 1 1 1 0 0 0 1 0 1 0 0 1 0 1 1 0 1

0 1 0 1 1 0 1 1 1 0 1 0 1 1 1 0 1 1 0 0 0 1 0 1 0 1 1 0 1 0 0 1 1 0 1 0 1 1 0 1 0 0 0 0 1 0 1 1 0 0 0 1 1 0

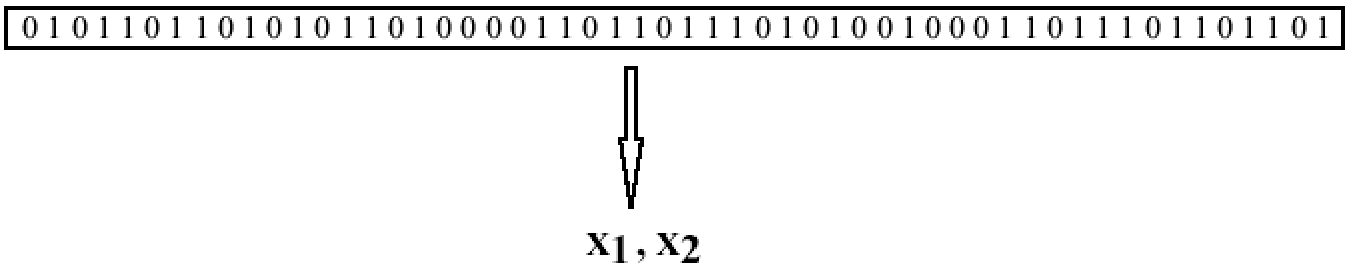
Reproduction (crossover) in EAs. (NOTE! Strongly simplified compared to the biological case!)

0 1 0 1 1 1 0 1 1 1 1 1 0 0 1 0 1 1 0 0 0 0 1 0 0 1 1 0 1 0 0 1 1 0 1 0 1 0 0 0 1 1 1 0 1 0 1 1 0 0 0 1 0 1

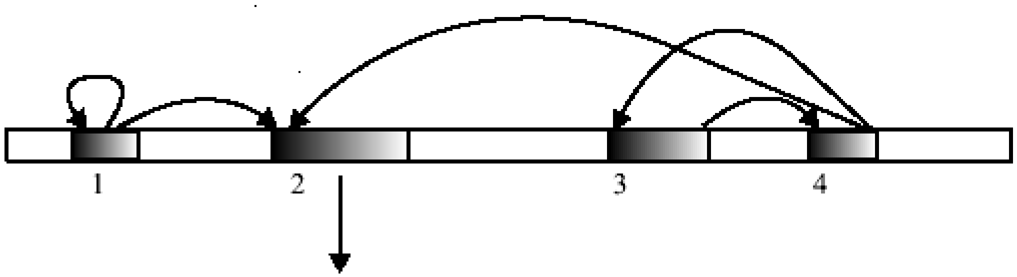
x, y

The decoding procedure in EAs is, in general, strongly simplified!

EA

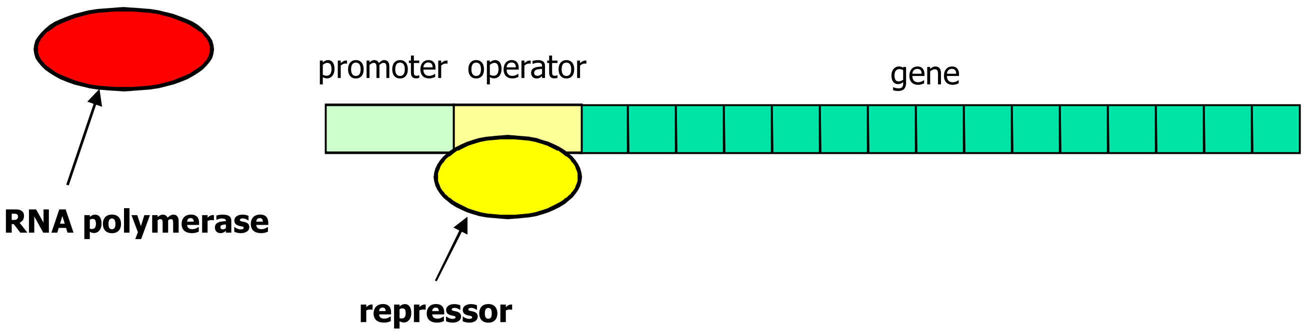

0 1 0 1 1 0 1 1 0 1 0 1 0 1 1 0 1 0 0 0 0 1 1 0 1 1 0 1 1 1 0 1 0 1 0 0 1 0 0 0 1 1 0 1 1 1 0 1 1 0 1 1 0 1
x_1, x_2

Biology

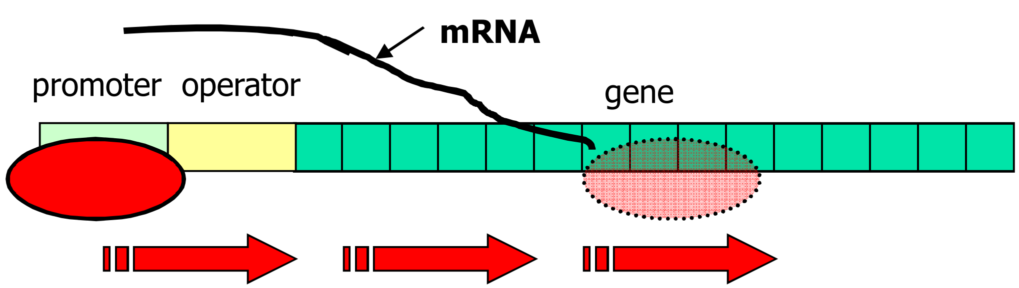

1
2
3
4

Regulatory genes: (transcription factors) genes that regulate the expression of other genes. Example of gene regulation:

Repressor protein (= the product of some other (regulatory) gene) bound to operator site: transcription is prevented



Repressor *not* bound to operator: the RNA polymerase can reach the promoter and proceed with transcription:

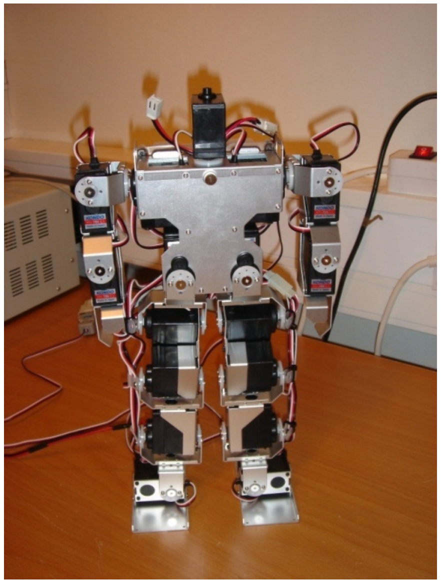


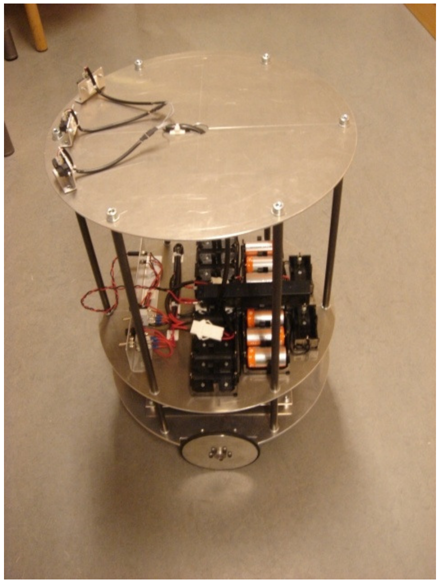
Evolution

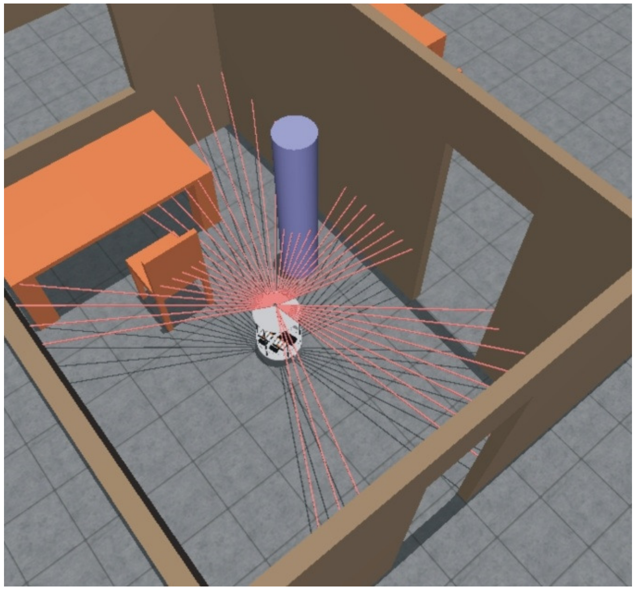
- Acts on **populations** of individuals (of a given species).
- Information is stored in the form of **chromosomes**.
- Each chromosome contains many **genes**.
- Well adapted individuals spread their genetic material (**reproduction**)
- Sexual reproduction: combination of genetic material from two individuals.
- Mutations generate new material for evolution to work with.

Adaptive systems research group

- My research group (Adaptive systems) mainly work with autonomous (freely moving) robots







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

Adaptive systems research group

- In particular, we are carrying out research regarding decision-making (behavior selection) in autonomous robots.

- We cooperate with groups in Japan (Waseda Univ., Tokyo) and Taiwan (ITRI, Hsinchu).

Adaptive systems research group

- In particular, we are carrying out research regarding decision-making (behavior selection) in autonomous robots.

- We cooperate with groups in Japan (Waseda Univ., Tokyo) and Taiwan (ITRI, Hsinchu).