

Stochastic optimization algorithms

Lecture 1, 20200901

Course introduction and motivation

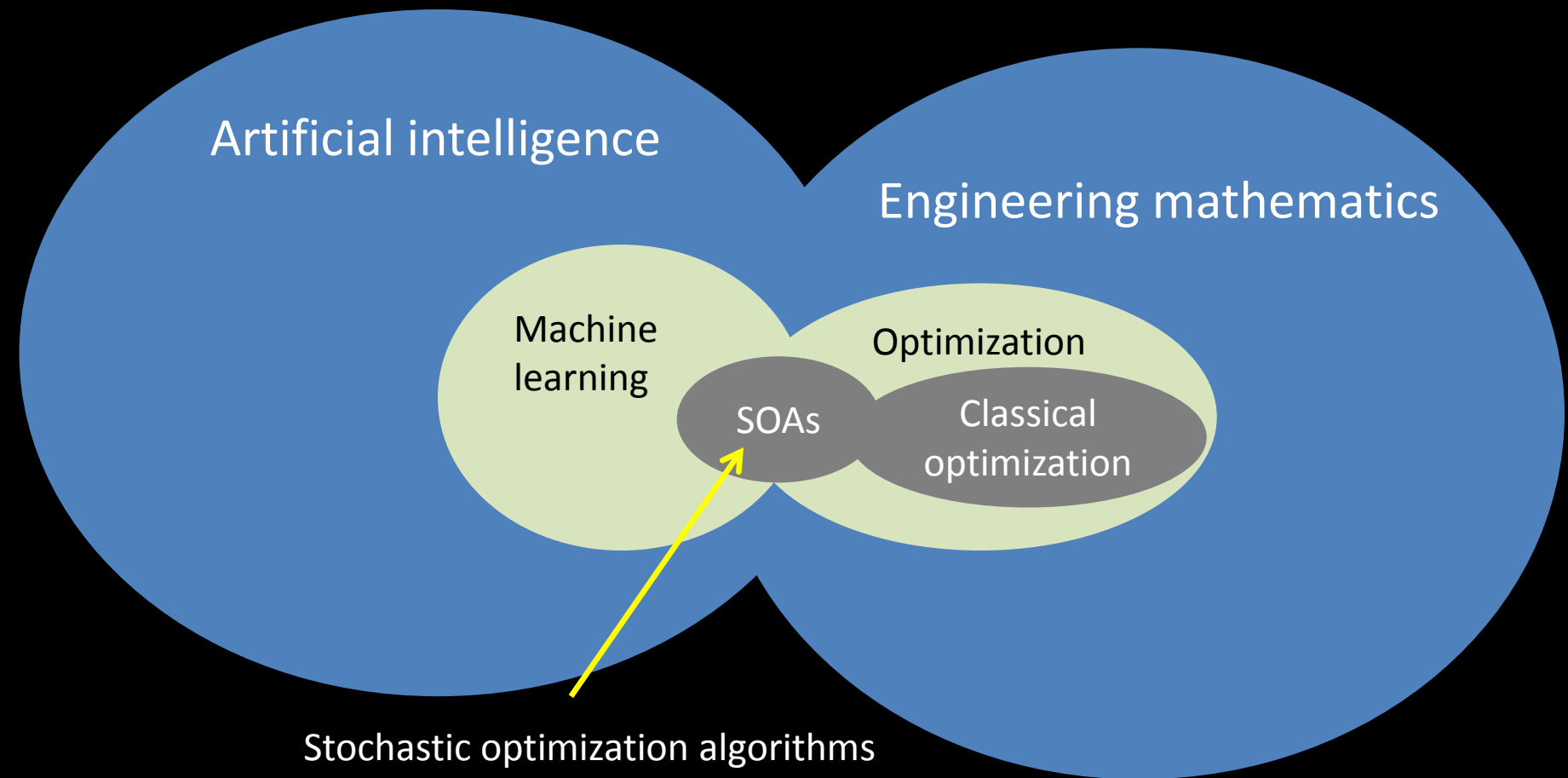
What will you learn?

- 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?

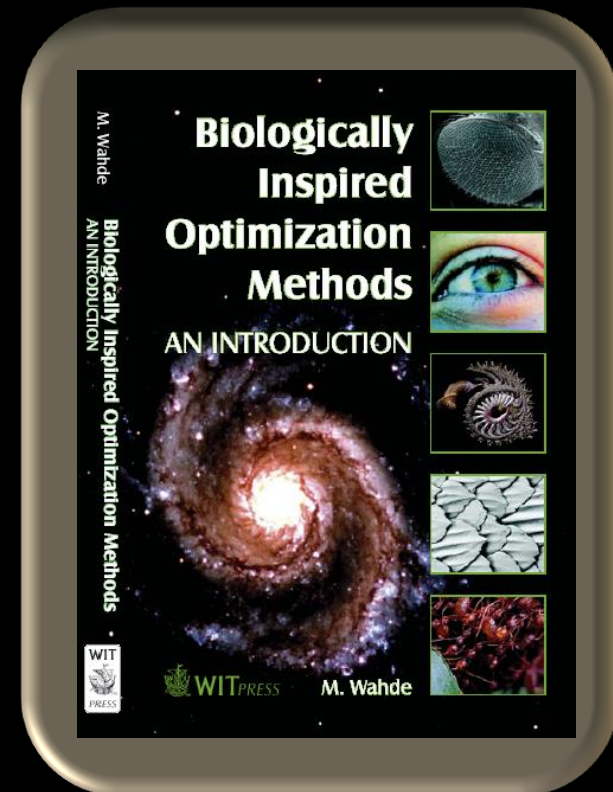
- Stochastic optimization algorithms can be used for solving many problems where classical methods are insufficient.
- Stochastic optimization methods are used more and more frequently in industry, particularly in large and computationally complex problems.
- The number of application areas is steadily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

Where does the course topic belong?



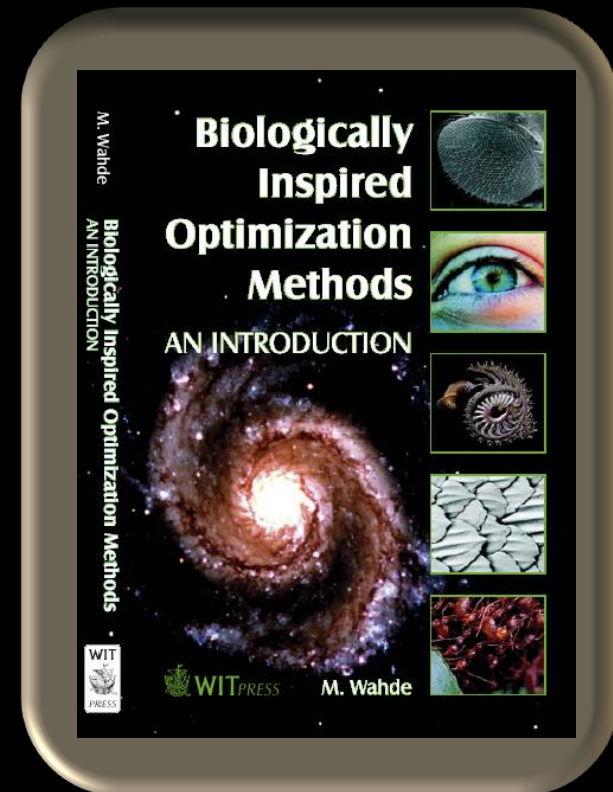
About the book...

- Available from several sources:
 - Amazon
 - The publisher (www.witpress.com)
 - Chalmers' bookstore (Cremona)
- Note that Cremona only has a limited supply...
- ...however, the book is also available as e-Book at Chalmers' library.



About the book...

- 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.. 😊)
- Note also that the book will be useful, to some extent, in other CAS courses e.g. Intelligent agents (TME 285, third quarter, highly recommended! 😊).



About the lectures...

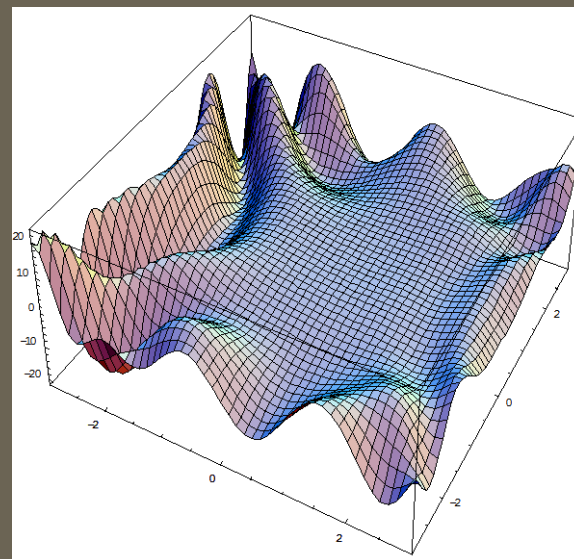
- When attending lectures, it is important that you are attentive and active, not just passively watching!
- In each lecture, I will list and review the *learning goals*.
- You can (and should) go back to the slides frequently, to make sure that you reach the learning goals.
- I will also (often) hand out a small questionnaire (quiz) which you should fill out during the lecture. You do not need to hand in the questionnaires. The answers will be posted on the web page within 24 h after the lecture.

Today's learning goals

- After this lecture you should be able to
 - Define the concept of optimization
 - List the two main classes of optimization methods
 - Give examples of biological adaptation
 - Briefly describe evolutionary algorithms
 - Briefly describe ant colony optimization
 - Briefly describe particle swarm optimization
 - Describe some applications of stochastic optimization
 - Define and describe the concepts of optima and critical points.

Optimization

- In general, optimization refers to the problem of finding the (global) minimum (or maximum) of an *objective function*.
- Sometimes (not always!) the objective function is a specific, well-defined mathematical function, $f = f(x_1, x_2, \dots)$.



Today's learning goals

- After this lecture you should be able to
 - **Define the concept of optimization**
 - List the two main classes of optimization methods
 - Give examples of biological adaptation
 - Briefly describe evolutionary algorithms
 - Briefly describe ant colony optimization
 - Briefly describe particle swarm optimization
 - Describe some applications of stochastic optimization
 - Define and describe the concepts of optima and critical points.

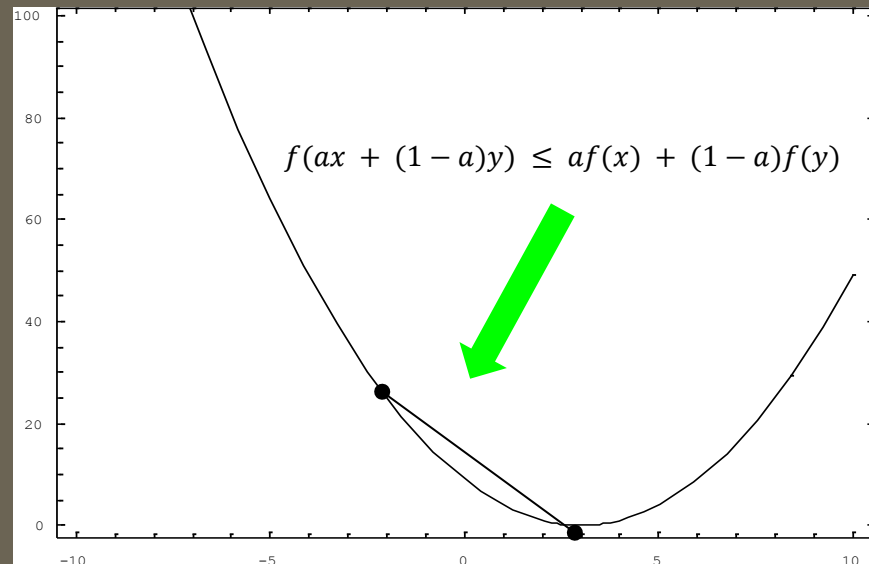
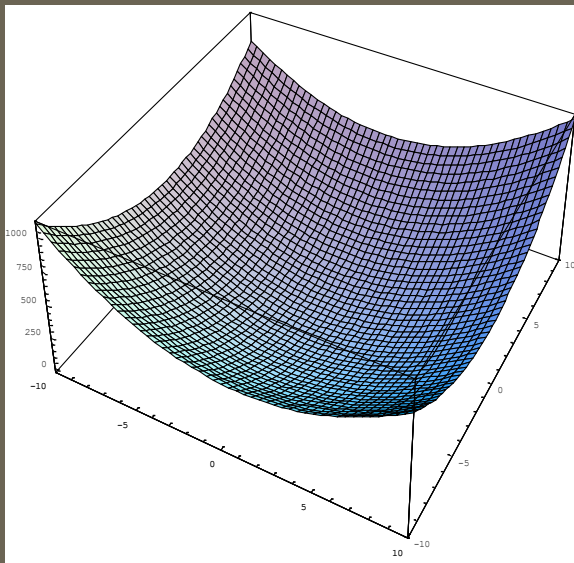


Optimization methods

- 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.

Convex functions

- Classical methods are particularly useful in *convex problems*, where any local minimum also is a global minimum.



Limitations of classical optimization

- Classical methods are less useful in cases with (for instance)
 - 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.

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*.

Today's learning goals

- After this lecture you should be able to
 - Define the concept of optimization
 - **List the two main classes of optimization methods**
 - Give examples of biological adaptation
 - Briefly describe evolutionary algorithms
 - Briefly describe ant colony optimization
 - Briefly describe particle swarm optimization
 - Describe some applications of stochastic optimization
 - Define and describe the concepts of optima and critical points.



Biologically inspired computation

- In general, the solution to many engineering problems can be found in biological systems.
- Examples of correspondences:
 - Image processing \Leftrightarrow vision (e.g. eyes of higher animals)*
 - Graph search \Leftrightarrow foraging (ants)
 - Decision-making \Leftrightarrow choice of activities (animals in general)
 - Radar \Leftrightarrow echo-location (bats)
 - Optimization \Leftrightarrow adaptation (darwinian evolution)
 - (applied) artificial intelligence \Leftrightarrow Animal intelligence

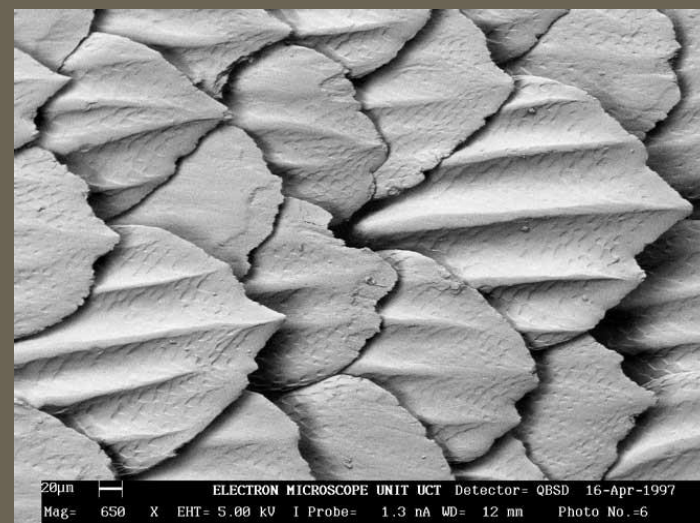
* (" \Leftrightarrow " should here be read as *corresponds to*)

Biologically inspired computation

- 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).

Example 1: Shark skin

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

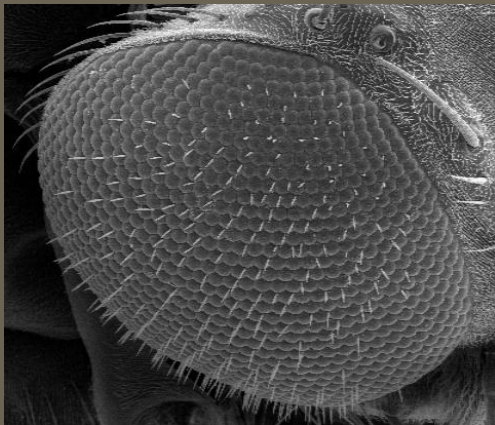


Example 1: Shark skin

- The rib-like structures affect the interaction between the surface of the shark and the surrounding water, essentially reducing drag.
- A 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.

Example 2: Evolution of the eye

- The evolution of eyes is another interesting 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.



Example 3: Swarming

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



Example 3: Swarming

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



Today's learning goals

- After this lecture you should be able to
 - Define the concept of optimization
 - List the two main classes of optimization methods
 - **Give examples of biological adaptation**
 - Briefly describe evolutionary algorithms
 - Briefly describe ant colony optimization
 - Briefly describe particle swarm optimization
 - Describe some applications of stochastic optimization
 - Define and describe the concepts of optima and critical points.



Biologically inspired optimization

- Many different optimization methods have been developed based on biological phenomena, and new methods appear continuously.
- In this course, we will study three such methods:
 - Evolutionary algorithms (inspired by evolution)
 - Ant colony optimization (inspired by cooperative behavior)
 - Particle swarm optimization (inspired by swarming)

Brief introduction to EAs

- Evolutionary algorithms (EAs), are a class of stochastic optimization methods (loosely) based on darwinian evolution.
- In these algorithms, one maintains a set of candidate solutions (rather than just one) to the problem at hand.
- The candidate solutions (individuals) are evaluated, and are then subjected to a set of evolutionary operators, to generate new individuals, which are then evaluated etc.

Darwinian evolution

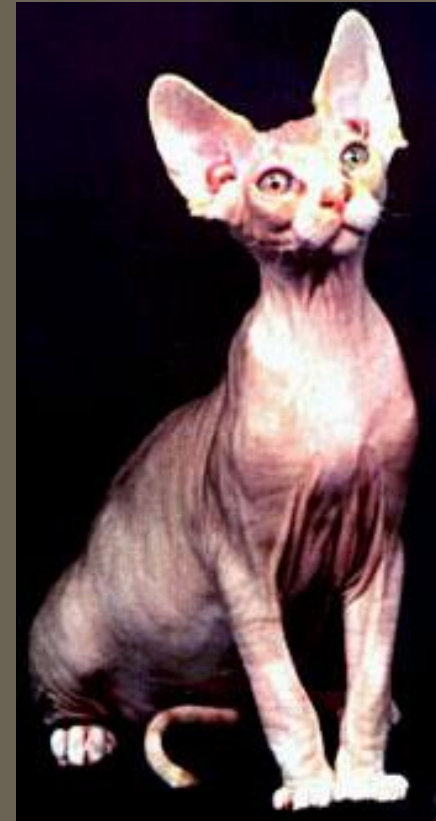
- *Darwinian evolution*, is a process involving *gradual, hereditary* changes of biological organisms, over long periods of time.
- Individuals that are well adapted to their environment have a high probability of generating offspring (*biological selection*).



pp. 35-98

Darwinian evolution

- *Mutations* are random changes to the *genotype*.
- Mutations sometimes lead 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

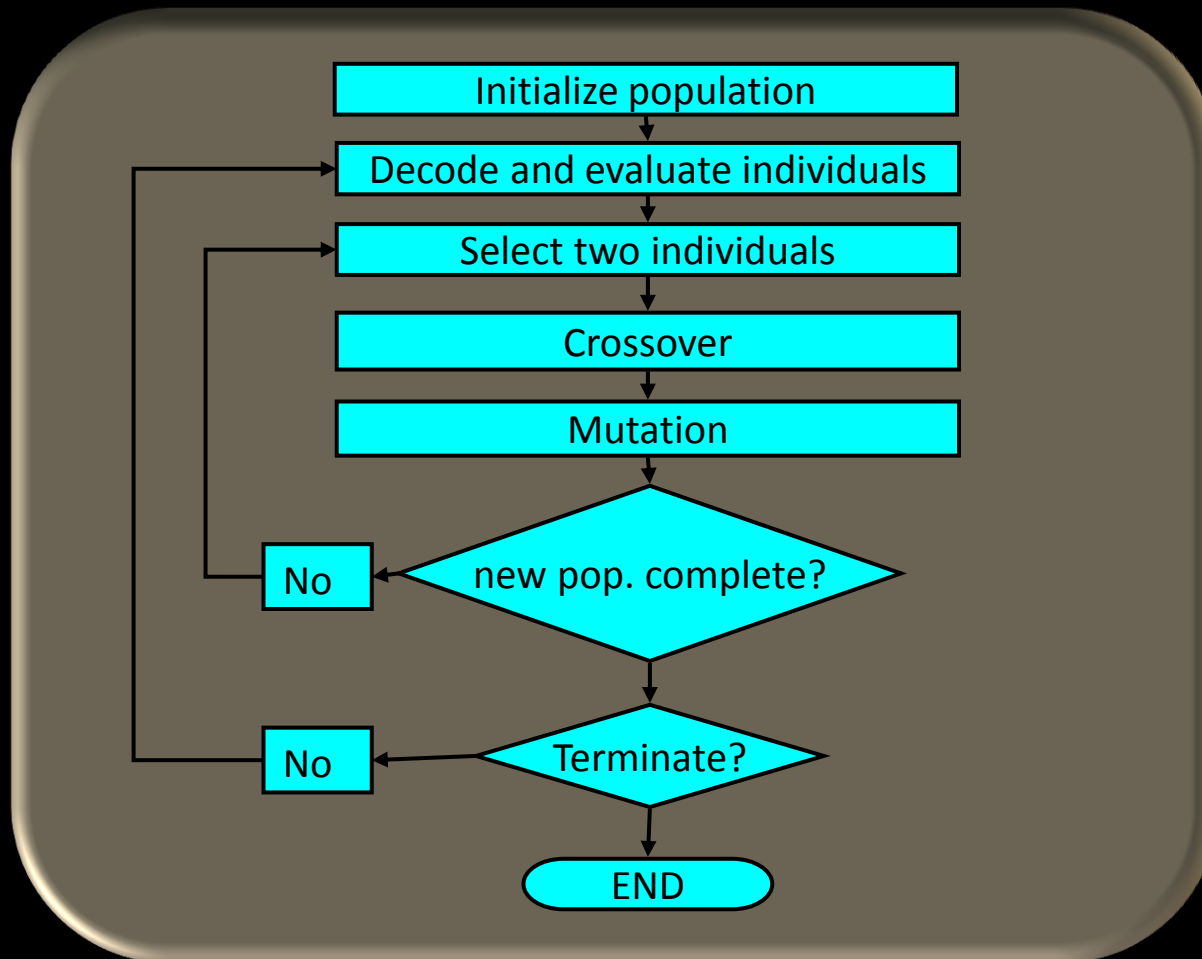
Brief introduction to EAs

- 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

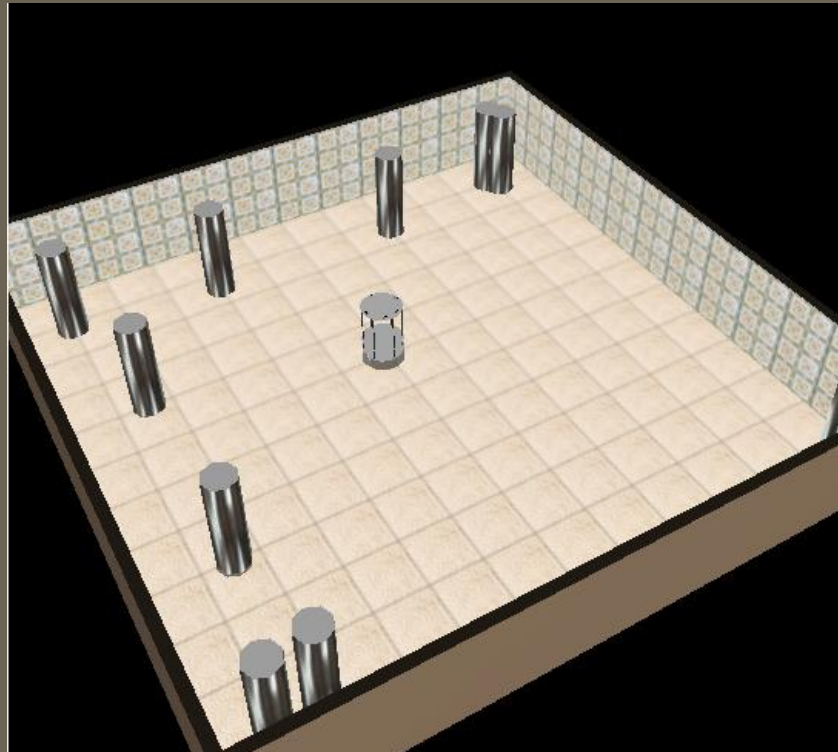


Summary



pp. 35-98

Simple EA example: Robot behavior



Simple EA example: Robot behavior

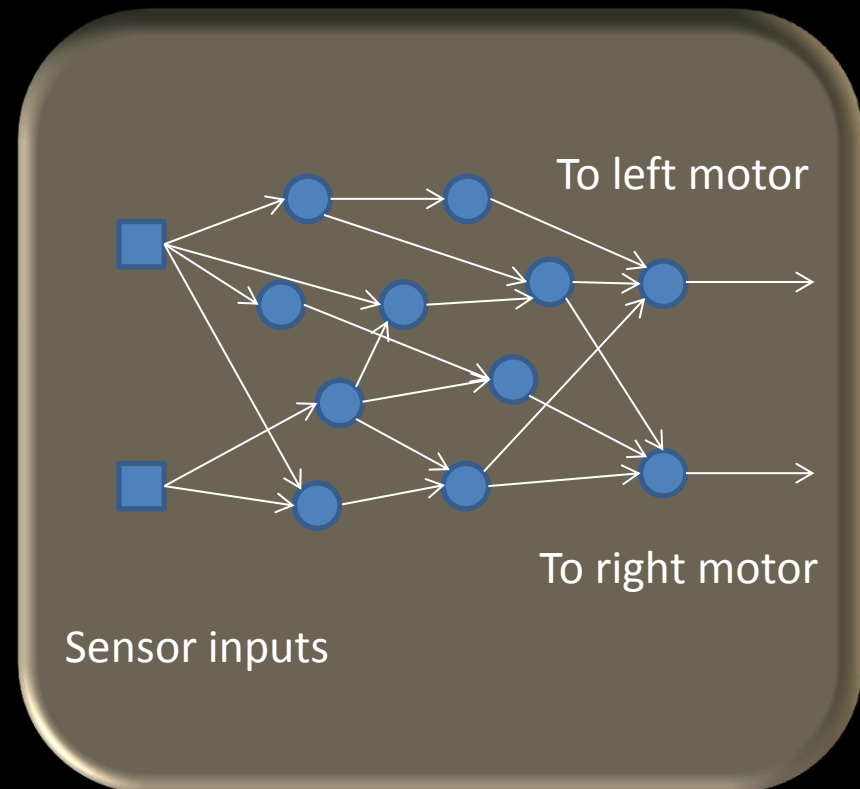
- Consider the following (toy) robotics problem:
 - A robot equipped with distance (IR or sonar) sensors is required to move as far as possible without colliding with any of the moving obstacles (cylinders).
- The navigation method of the simulated robot is *evolved* (rather than programmed), using an evolutionary algorithm.

Simple EA example: Robot behavior

- Evaluation measure (objective function): Distance moved until collision or maximum time reached.
- Thus, the robot is not given *any* information or feedback regarding its individual actions. It gets only an overall measure of its performance.
- The maximum allowed time is gradually increased.

Simple EA example: Robot behavior

- The robotic brain is structured as a continuous-time artificial neural network.
- The actions of the robot are determined by the structure and parameters of the network.



Simple EA example: Robot behavior

- Videos shown:
 - Generation 1 (random networks)
 - Generation 5
 - Generation 37
 - Generation 59
 - Generation 104
 - Generation 153

Simple EA example: Robot behavior

- Note, again, that the robot learns even rather complex actions without direct specification of what it should do.
- However, note that in a realistic application, one would need to use a more complex setup:
 - Multiple evaluations for each individual (to avoid adaptation to special conditions).
 - Perhaps penalties for near misses, when the robot passes an obstacle almost without margin.
 - Better objective function, to avoid excessive turning.
 - Probabilistic (rather than deterministic) obstacle motion.

Simple EA example: Robot behavior

- However, probabilistic obstacle motions imply that several evaluations would be required for each robot (time-consuming)
- Other aspects:
 - Gradual increase in the duration of each robot evaluation.
 - Gradual variation (often increase) in the size of the neural network defining the robotic brains.
- Thus, even though the algorithm (GA) itself is not difficult to understand, using it successfully requires some experience, which you will get by taking this course! 😊

Today's learning goals

- After this lecture you should be able to
 - Define the concept of optimization
 - List the two main classes of optimization methods
 - Give examples of biological adaptation
 - **Briefly describe evolutionary algorithms**
 - Briefly describe ant colony optimization
 - Briefly describe particle swarm optimization
 - Describe some applications of stochastic optimization
 - Define and describe the concepts of optima and critical points.



Brief introduction to ACO

- Ant colony optimization (ACO) is a class of optimization methods inspired by the cooperative behavior of ants.
- ACO problems are generally formulated as search problems on graphs (i.e. finding the shortest path).

Ants

- Ants are capable of very efficient cooperative behavior.
- An example of cooperative behavior is the discovery of short paths (between the nest and a food source) during foraging (food search).
- How do ants achieve this?
- Note that most ants are
 - (Almost) blind
 - Have no explicit leaders



pp. 99-116

Ants

- 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.
- The behavior of the ants is an *emergent property* of the ant colony as a whole.

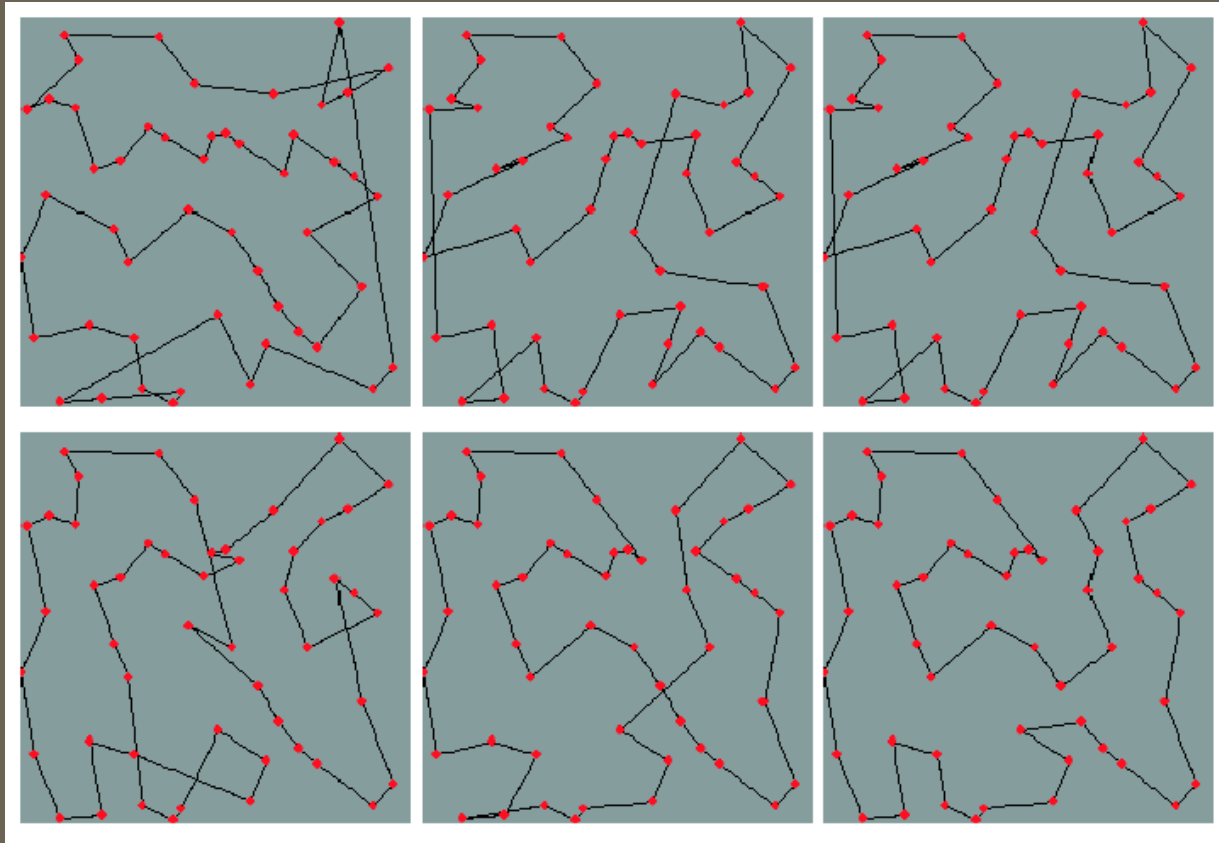
Brief introduction to ACO

- In ACO, the same principles are used: Artificial ants deposit artificial pheromones, while attempting to find the shortest path in a graph.
- As in EAs, one maintains a population of artificial ants.
- Ants that find short paths deposit more pheromone than ants that find long paths.
- Ants move probabilistically, in proportion to the amount of pheromone along the various accessible paths.

pp. 99-116



Brief introduction to ACO



pp. 99-116

Today's learning goals

- After this lecture you should be able to
 - Define the concept of optimization
 - List the two main classes of optimization methods
 - Give examples of biological adaptation
 - Briefly describe evolutionary algorithms
 - **Briefly describe ant colony optimization**
 - Briefly describe particle swarm optimization
 - Describe some applications of stochastic optimization
 - Define and describe the concepts of optima and critical points.



Brief introduction to PSO

- Particle swarm optimization (PSO) is based on swarming behavior in nature, a phenomenon that is common in many species of (for example) bird and fish.



pp. 117-138

Brief introduction to PSO

- Just as in EAs and in ACO, in PSO one maintains a swarm (a population, of so-called *particles*) that move in the n -dimensional search space.
- Thus, each particle defines a vector \mathbf{x} such that the objective function $f(\mathbf{x})$ can be evaluated.
- The velocities of the *particles* are updated such they preferentially move towards those positions where particles have previously obtained a good result.
- Once velocities are available, new positions can be found etc.

pp. 117-138



Today's learning goals

- After this lecture you should be able to
 - Define the concept of optimization
 - List the two main classes of optimization methods
 - Give examples of biological adaptation
 - Briefly describe evolutionary algorithms
 - Briefly describe ant colony optimization
 - **Briefly describe particle swarm optimization**
 - Describe some applications of stochastic optimization
 - Define and describe the concepts of optima and critical points.

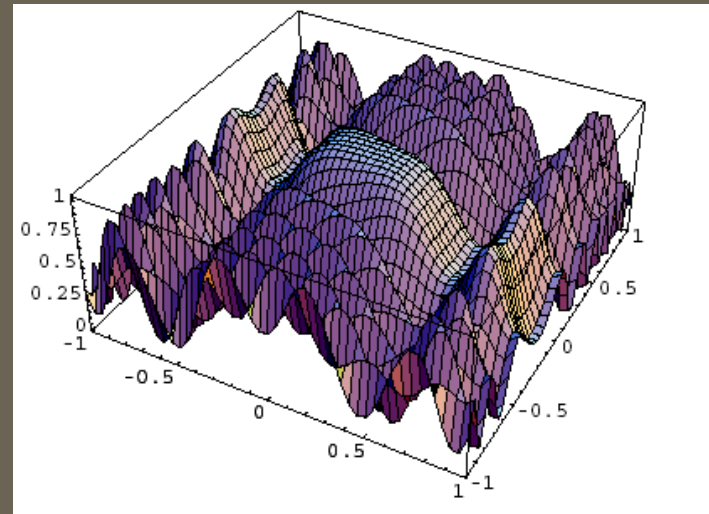


Applications of stochastic optimization

- On the following slides I will list just a few examples of applications of stochastic optimization algorithms (SOAs).
- The selection involves problems that I have worked with over the years – thus, it is not an unbiased sample of all possible applications. There are *many* more! 😊
- We will discuss many different applications during the course.

Function optimization

- Stochastic optimization is particularly useful in problems involving many local optima, where classical methods can get stuck.
- In many cases where SOAs are applied, the objective function consists of many variables (thousands or more).



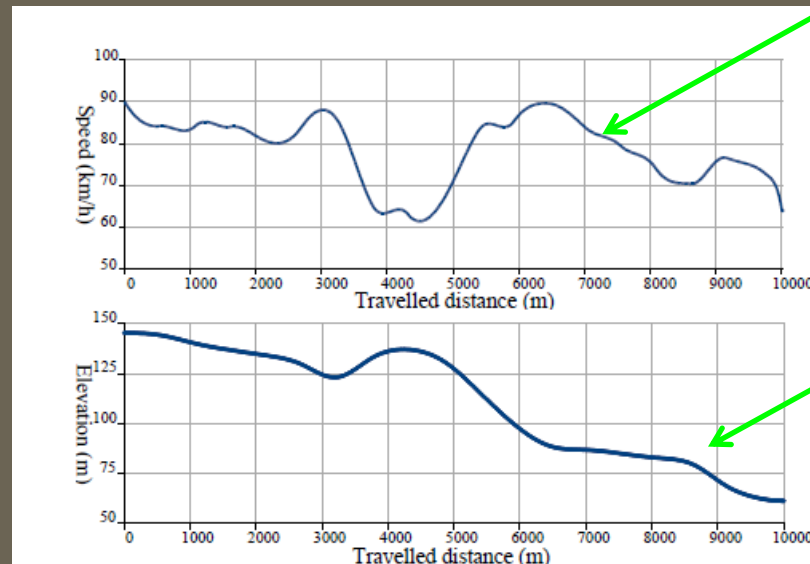
Scheduling

- SOAs can be used in many scheduling problems, for example airline crew scheduling.



Engineering optimization

- An example: Truck fuel consumption reduction by using optimized speed profiles.

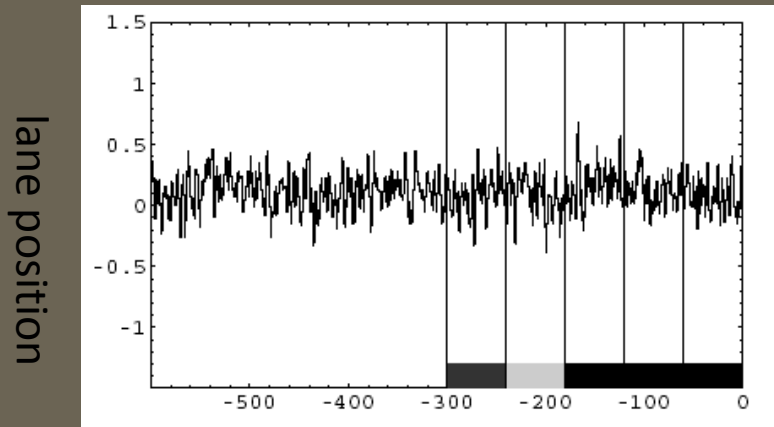


Speed profile,
optimized using
a genetic algorithm

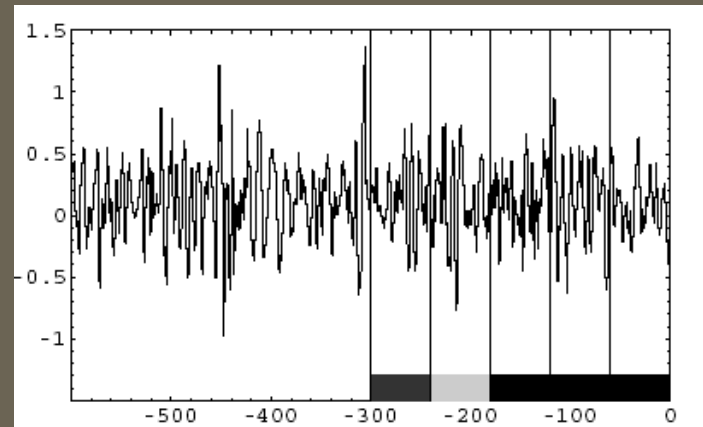
Road profile

Classification

- For example, detection of sleepiness in car drivers based on time series of driving behavior.
- Two classes: Sleepy and alert.



Alert driver



sleepy driver

Time series prediction

- Example: Financial time series prediction



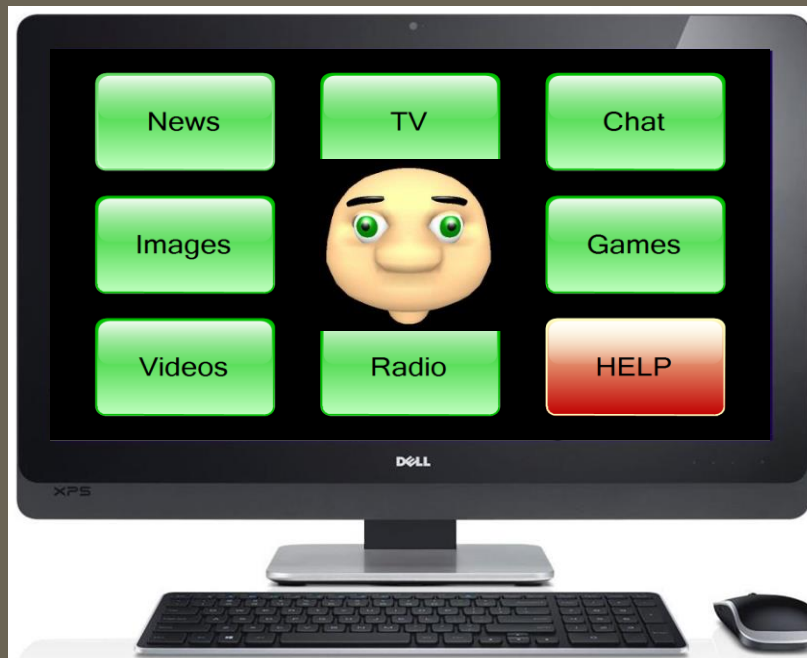
Today's learning goals

- After this lecture you should be able to
 - Define the concept of optimization ✓
 - List the two main classes of optimization methods ✓
 - Give examples of biological adaptation ✓
 - Briefly describe evolutionary algorithms ✓
 - Briefly describe ant colony optimization ✓
 - Briefly describe particle swarm optimization ✓
 - **Describe some applications of stochastic optimization** ✓
 - Define and describe the concepts of optima and critical points.

Applied artificial intelligence group

- In my research group (Applied artificial intelligence, AAI), we currently work with three main topics
 - *Intelligent software agents* (conversational agents)
 - *Autonomous robots* (currently working on an interactive assistance robot for Chalmers' library)
 - *Autonomous vehicles* (driveable surface detection, platooning, fuel consumption minimization, trajectory planning etc.)
- Almost all our work involve stochastic optimization methods, in some way.

Intelligent agents



Autonomous robots



Autonomous vehicles



Practical details: Lectures

- This year, due to the pandemic, there will not be standard classroom lectures.
- Instead, each lecture will consist of a video presentation directly followed by an interactive Zoom session.
- You should strive to view the video (roughly 2 x 45 minutes) *before* the corresponding Zoom session.
- The course book (see slides 4-5) defines the course in principle, but it is essential for you also to assimilate the information given during lectures. Thus, make sure to participate actively in the Zoom sessions.

Practical details: Schedule

- Lectures (see also the course web page).
 - Tuesdays: Video published at 07.00, Zoom session 11.00-11.45
 - Wednesdays: Video published at 17.00 the day *before* the lecture, Zoom session 09.00-09.45
 - Fridays: Video published at 07.00, Zoom session 11.00-11.45
- An extra session (Matlab programming *for stochastic optimization*) is planned for Tuesday Sept. 8 (evening).


Practical details: Teacher availability

- During the course, you may ask questions at any time, not only during the Zoom sessions. You can pass by our offices (see map below) or send an e-mail.
- Please don't hesitate to ask questions, any time. You may also ask the course assistants.
- When e-mailing, preferably use your Chalmers or GU mail address (to avoid that the e-mail ends up in the spam box).

Practical details: Web page

- The course web page can be found at <https://chalmers.instructure.com/courses/10274>
- Check the web page *often*. There will be frequent updates (programs, home problems, FAQ etc.)

Practical details: Web page

 [FFR105](#) > [Syllabus](#)

[lp1 HT20](#)
[Home](#)
[Grades](#)
[People](#)
[Modules](#)
[Syllabus](#)

FFR105 Stochastic optimization algorithms

[Jump to today](#)

Course memo

FFR105, Stochastic optimization algorithms, SP1, Autumn 2020, 7.5p

The course is offered by the Department of Mechanics and Maritime Sciences

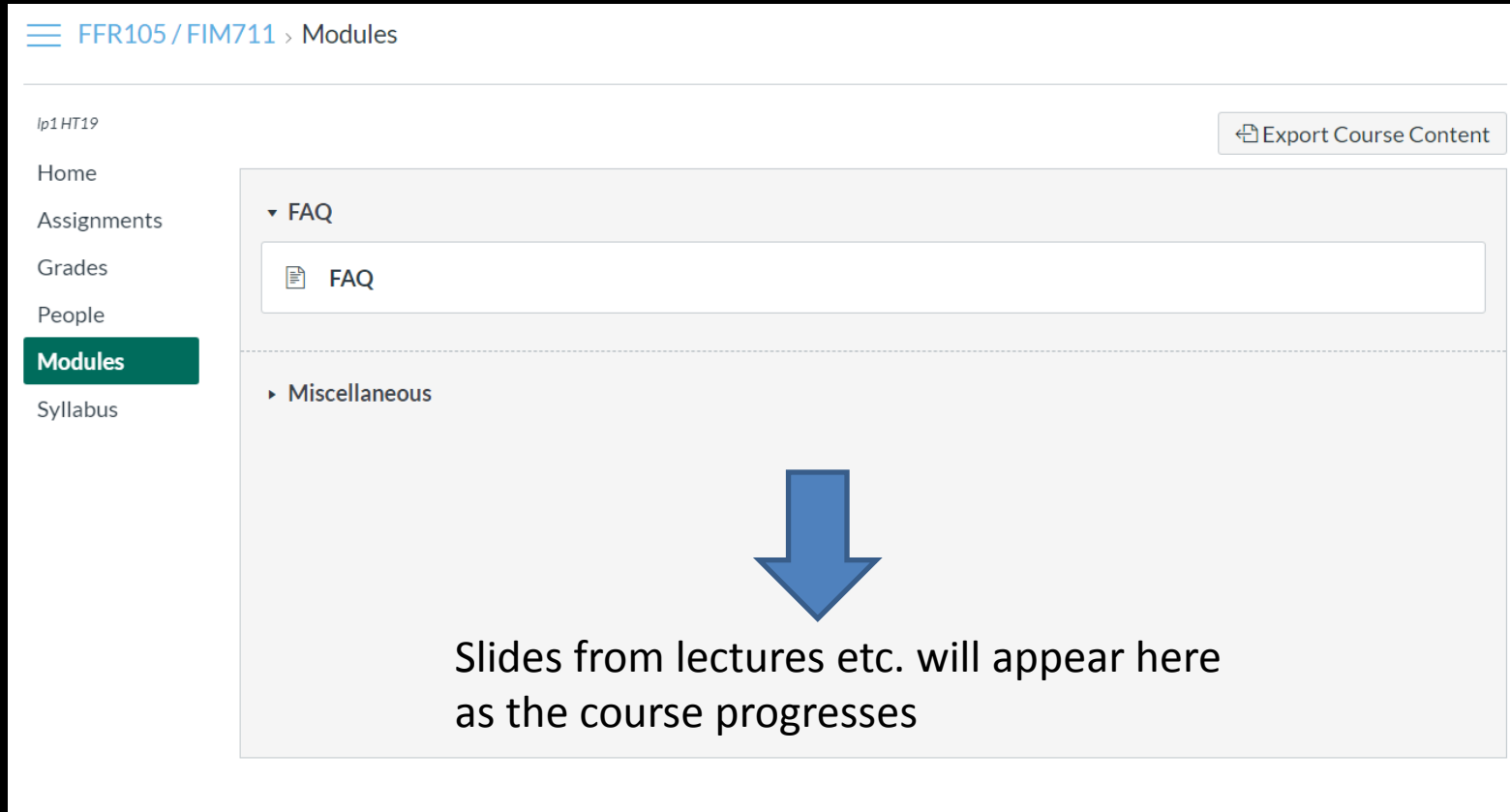
General information

This year, with the ongoing pandemic, the course will be given in a different way than usual. There will be no classroom lectures. Instead, each lecture will consist of a video with a recorded presentation of the lecture (uploaded at the scheduled time of the lecture; see below), directly followed (just after the end of the scheduled lecture time) by an interactive session on Zoom, where I will discuss the contents of the lecture with the students, and also answer any questions that the students might have. In addition, students are welcome to ask questions or give feedback at any time (see also Contact details below).

Contact details

During the course, we strive to be available as much as possible. You are welcome to ask questions at any time, either in the Zoom session associated with each lecture or at other times, and you may also ask questions via e-mail or telephone. You are always welcome at our offices. However, due to the pandemic, we can only receive one (1) student at a time, and (of course) only if you are symptom free. You do not need to make an appointment, but since we are not always in our offices (in the current situation, we mostly work from home) it's a good idea to first check that we are there (e.g. via e-mail or telephone).

Practical details: Web page



FFR105 / FIM711 > Modules

ip1 HT19

Home

Assignments

Grades

People

Modules

Syllabus

Export Course Content

▼ FAQ

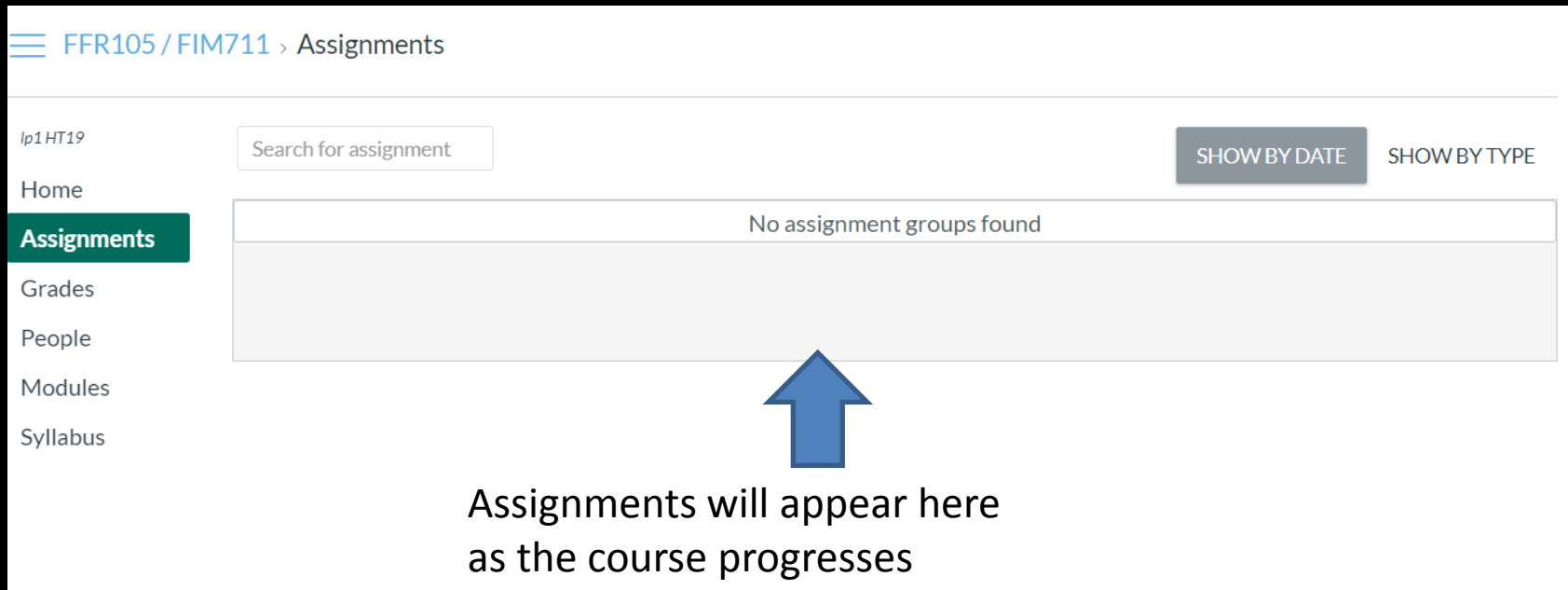
FAQ

► Miscellaneous

↓

Slides from lectures etc. will appear here
as the course progresses

Practical details: Web page



FFR105 / FIM711 > Assignments

lp1HT19

Search for assignment

SHOW BY DATE SHOW BY TYPE

Home

Assignments

Grades

People

Modules

Syllabus

No assignment groups found

Assignments will appear here
as the course progresses

Practical details: Examination

- The examination will be determined by your result on
 - Two sets of home problems. Maximum score: $10+15 = 25$ p.
 - A written exam at the end of the course: Max. score: 25p.
- In order to pass the course, you must, as a minimum ...
 - ...correctly solve an introductory programming problem.
 - ...get at least 10p on the exam.
 - ...hand in satisfactory solutions to the mandatory parts of the home problems.

Practical details: Examination

- Concerning home problems, you may discuss with other students, but you **must hand in your own solution.**
Do not copy the work of any other student.
- In this course, I only care about your **individual abilities**, not your ability to work in a team etc.

Practical details: Examination

- For those who aspire to higher grades, the grade limits are, for Chalmers:
 - Grade 3: Up to 32.5p
 - Grade 4: 33 to 41.5p
 - Grade 5: 42p and up
- For GU students, the grade limits are:
 - Grade G: Up to 38.5p
 - Grade VG: 39p and above
- ECTS grades: A \Leftrightarrow 5, B \Leftrightarrow 4, C \Leftrightarrow 3, D \Leftrightarrow weak 3 (below 25p)

Practical details: Office location

- Mechanical engineering building (nya M-huset). Enter near Café Bulten, follow the stairs (one floor up).
- If needed, dial my extension (3727).
- Go right as you enter the corridor.

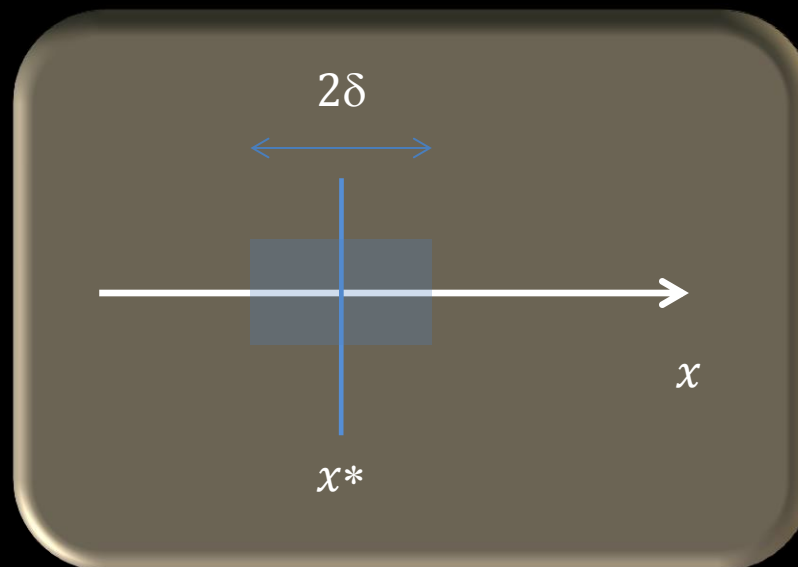


Classical optimization methods

- This first week, we will consider classical optimization.
- Starting next week, and for the remainder of the course, we will consider stochastic optimization.
- For some of you, the topics of the first week will be a repetition – no harm in that! 😊
- **Very important:** Make sure that you (also) *follow the lectures next week*. Crucial for understanding SOAs!

The concept of a *minimum*

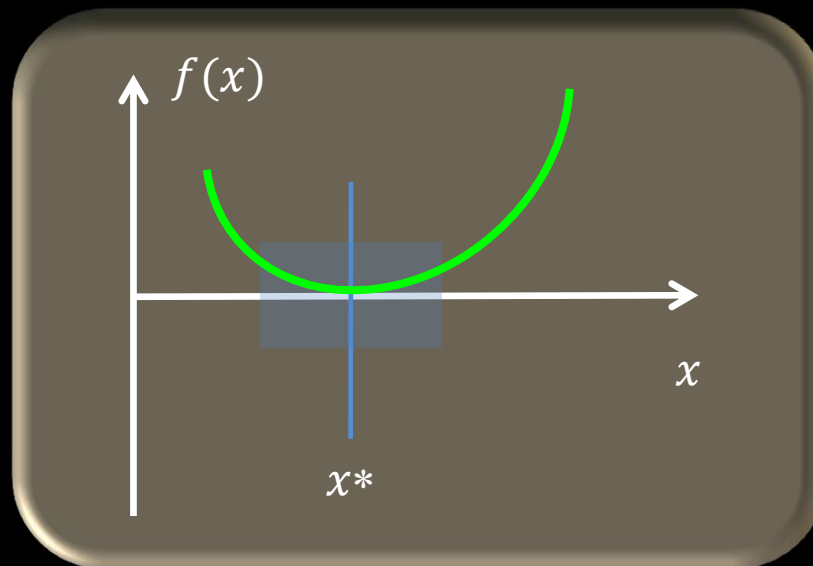
- Let $I(x^*, \delta) = \{x: |x - x^*| < \delta\}$, where $\delta > 0$, denote a neighbourhood of x^*



The concept of a *minimum*

- Local minimum at a point x^* if the function takes larger values around x^* , i.e. if

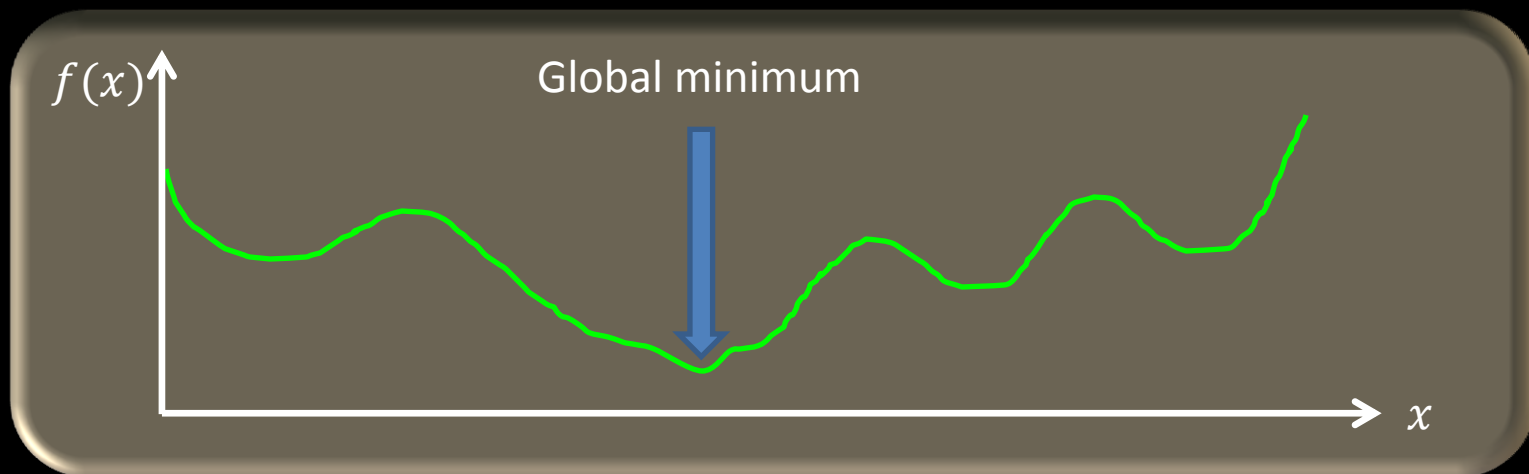
$$\exists \delta > 0: f(x) > f(x^*) \quad \forall x \in I(x^*, \delta), \quad x \neq x^*.$$



pp. 9-10

The concept of a *minimum*

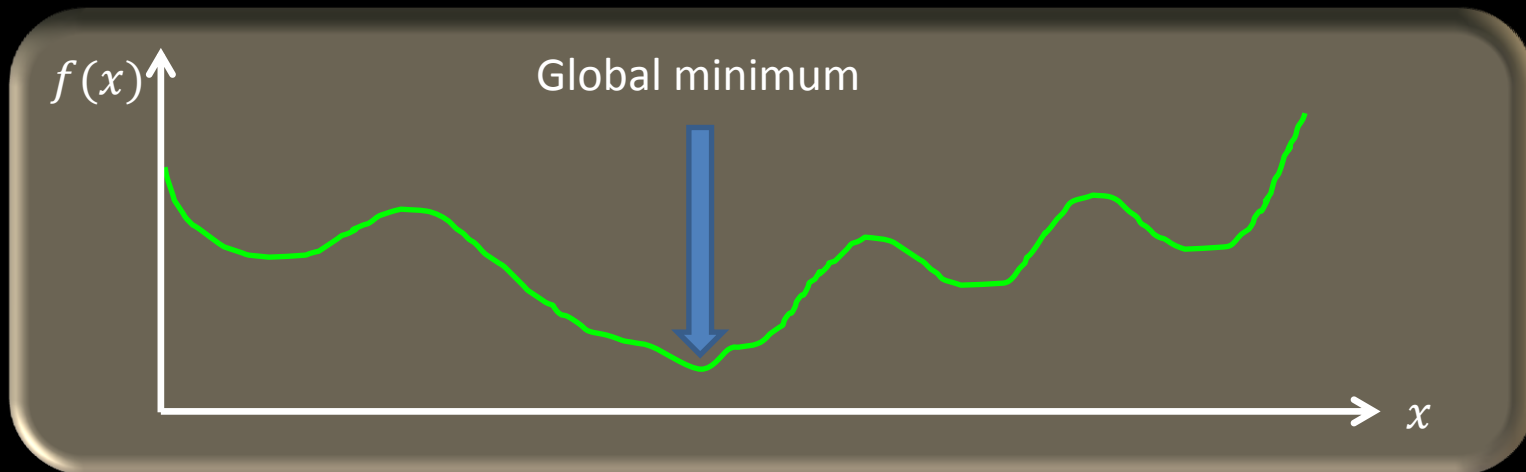
- Maxima can be defined analogously.
- Both concepts can be generalized to several dimensions.
- A function f has a global minimum at x^* if
$$f(x) \geq f(x^*) \quad \forall x \in D$$



pp. 9-10

Global *minimum*

- A function f has a global minimum at x^* if $f(x) \geq f(x^*) \forall x \in D$



- A function can have any number of global optima (minima or maxima). Example: $f(x) = \sin(x)$.

Objective function

- The quantity whose global minima (or maxima) one is searching for (in optimization).
- Maximization of $f \Leftrightarrow$ minimization of $-f$.
- No restriction to consider only minimization.

Constraints

- In many problems, the decision variables \mathbf{x} are not allowed to take any values in R^n .
- *Constraints*:
 - *Inequality*: $g_i(\mathbf{x}) \leq 0, i = 1, \dots, m$
 - *Equality*: $h_i(\mathbf{x}) = 0, i = 1, \dots, k$
- *Feasible point*: A point \mathbf{x} that satisfies all $m + k$ constraints.

Constrained optimization problem

- Minimize $f(\mathbf{x})$, $\mathbf{x} \in S$ where S is the *set of feasible points*.

- Note:

$$g_i(\mathbf{x}) \geq 0 \iff -g_i(\mathbf{x}) \leq 0$$

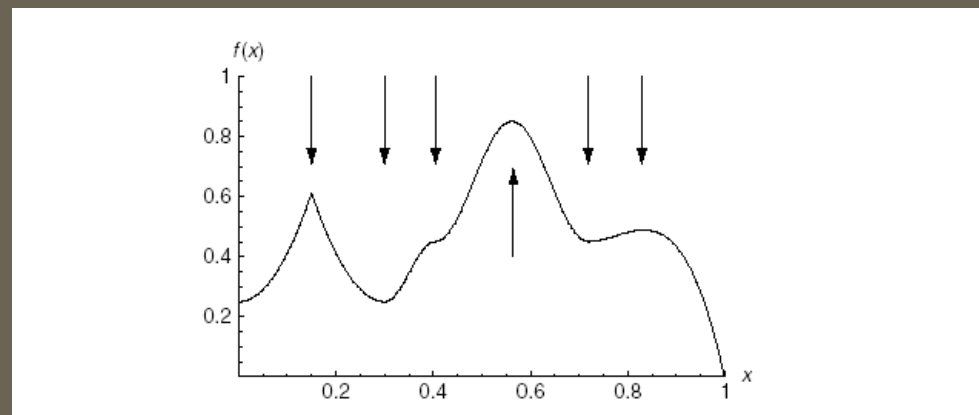
...and

$$h_i(\mathbf{x}) = 0 \iff \{h_i(\mathbf{x}) \leq 0 \text{ and } -h_i(\mathbf{x}) \leq 0\}$$

- (Taxonomy, pp.11-12, read by yourselves).

Properties of local optima

- *Critical point*: Occurs where either
 1. $f'(x^*) = 0$ (in which case x^* is a *stationary point*) or
 2. $f'(x^*)$ does not exist
- The local optima of f occur at critical points.



pp. 12-14

Nature of stationary points

- If there exists a $\delta > 0$ such that
 1. $f(x)$ is continuous in $I(x^*, \delta)$,
 2. The derivative $f'(x)$ exists in $I(x^*, \delta)$,
 3. $f'(x) < 0 \forall x \in]x^* - \delta, x^*[$,
 4. $f'(x) > 0 \forall x \in]x^*, x^* + \delta[$,... then $f(x)$ has a local minimum at x^* .
- Maximum: defined analogously.
- Some stationary points are **saddle points** (neither minima nor maxima). Example: $f(x) = x^3$ at $x = 0$.

Today's learning goals

- After this lecture you should be able to
 - Define the concept of optimization ✓
 - List the two main classes of optimization methods ✓
 - Give examples of biological adaptation ✓
 - Briefly describe evolutionary algorithms ✓
 - Briefly describe ant colony optimization ✓
 - Briefly describe particle swarm optimization ✓
 - Describe some applications of stochastic optimization ✓
 - **Define and describe the concepts of optima and critical points.** ✓