Part MMDCCXV

Future of Informatics - Chapter 6

CHAPTER 6:

NEW, INFORMATICS DRIVEN,

GENERAL METHODOLOGY

for SCIENCE, TECHNOLOGY, MEDICINE, ...

CONTENTS

PROLOGUE

PROLOGUE

CHANGING THE MAIN PARADIGM of SCIENCES

We are currently witnessing an important change in Science.

From

Galilean, Mathematics-driven, science,

that dominated science since its modern beginnings, and during which the paradigms, concepts, methods, tools and results of Mathematics have dominated,

to

Post-computer, Informatics-driven, science

at which paradigms, concepts, methods, tools and results of Informatics dominate.

LANGUAGES in WHICH NATURE IS WRITTEN

- Galileo: Nature is written in the language of mathematics.
- Wolfram: Nature is written in the language of computation.
- New view: Nature is written in the language of information processing.

CURRENT MAIN BASIC METHODOLOGIES

There are currently two basic methodologies sciences, technologies,.... have:

- Experimental methodology (with observations and experiments as main tools)
- Theoretical methodology (with reasoning, deduction and proofs as main tools)

ORIGINS

Modern science was borned when two basic scientific methodology, experimental and theoretical, started to be clearly formulated and developed.

This happened in 17th century

ORIGINS of MODERN SCIENCE

Modern science and its two main methodologies are usually seen as having its origin in 17-teen century in

- critical rationalism of Descartes (1596-1650);
- practical empiricism of Galileo Galilei(1564-1642); and
- breathtaking contributions of Isac Newton (1642-1727).

Galileo's main position was a concentration on:

- making carefully chosen observations and experiments;
- measuring things quantitatively and using quantitative methods to process results of observations and experiments,
- thinking mathematically and
- submitting theoretical outcomes/hypothesis to the nature verification/experiments.

DEEP THOUGHTS of FOUNDERS

- Measure what is measurable, and make measurable what is not.

  Galileo Galilei
- All truths are easy to understand once they are discovered. The point is to discover them.

  Galileo Galilei
- The Bible shows the way to go to heaven, not the way heaven go.

  Galileo Galilei
- We cannot teach people anything, we can only help them discover it within themselves.

  Galileo Galilei
- I think, therefore I am. R. Descartes

GALILEO GALILEI

INFORMATICS as a BASIS of a NEW METHODOLOGY

Informatics, as a symbiosis of a scientific and a technology discipline, develops also basic ingredients of a fundamentally new, in addition to deduction/theory and observations/experiments, the third basic methodology for all sciences, technologies, medicine, and society in general.

This new, informatics-based, methodology provides a new way of thinking and new languages for science and other areas, extending the Galilean experiments/mathematics-based approach to science and technology to new heights.

NEW, INFORMATICS-DRIVEN, METHODOLOGY

Informatics-driven methodology subsumes and extends the role and improves tools mathematics used to play and offer in advising, guiding and serving other scientific and technology disciplines and society in general.

BASIC FEATURES of INFORMATICS-DRIVEN METHODOLOGY - BRIEFLY - I.

- Modeling - Design of information processing models of nature, men made, and virtual phenomena and processes.
- Simulation - Utilization of information processing models to study natural, men-made and virtual phenomena and processes.
- Visualisation and animation of data, objects and processes.
- Virtualization - of systems and processes including designs of the virtual reality systems and worlds.
- Mining of information and knowledge from huge data streams and sets.

BASIC FEATURES of INFORMATICS-DRIVEN METHODOLOGY - BRIEFLY - II.

- Design of systems for mechanized problem solving, reasoning, learning and proofs designing as well as proofs checking.
- Searching (sophisticated searching as an alternative to deep knowledge based reasoning)
- Design of artificial systems, especially artificial intelligence and life systems, as a way to understand natural and nature-made systems.
- Development of methods to design, analyse and verify complex (information processing) systems.

BASIC FEATURES of INFORMATICS-DRIVEN METHODOLOGY - BRIEFLY - III.

- Design of algorithms, study of their performances and study of the inherent complexities of computational, communication and description systems, as a way to get a deep understanding of various phenomena and their interrelations.
- Design, analysis and comparison of descriptional languages and systems and of the relations between objects and their specifications.
- Study of the real world through the study of all (also virtual) information processing worlds.
- Development of tools to mechanized research results creation and presentation.
- Development of platforms for worldwide cooperation of research and development communities.

FEATURES of INFORMATICS-DRIVEN METHODOLOGY - MORE DETAILS - I.

Modelling: - concepts, theory, tools and methodologies for specification, design, verification and analysis of information processing models, especially of very and ultra complex human-made and nature-created systems and processes.

Simulation: - theories and methods of design, validation and verification of simulations, as well as of the display of their outcomes using imaging, visualisation and animation tools.

Visualisation: - and animation theories, methodologies and tools.

Virtualization: - formalization, digitalization, (multilayer) generalization, (structural) composition/decomposition, reduction as well as approximation and other abstraction-driven methodologies and tools the importance of which has been verified in the development of mathematics;

Design of AI systems: - The design of (productive, clever, cooperative, emotionally intelligent and trustworthy) robots and other intelligence intensive systems;

Design of artificial human limbs, organs and bodies; that match or outperform those of human bodies.

FEATURES of INFORMATICS-DRIVEN METHODOLOGY - MORE DETAILS - II.

Data, information and knowledge mining: An application of data mining, pattern matching, information retrieval, knowledge discovery and hypothesis formation as well as of learning systems to huge sets or streams of data provided by sensors or obtained due to the digitalization processes.

Automatization of reasoning: The design of systems for (automated) problem solving, reasoning, theories creating, proof checking and theorem proving and of other systems to deal with the tasks that used to belong to the human domain only;

Design techniques: The development of methods to specify, design, analyse, verify, modify and maintain very complex, parallel and distributed (information processing) systems.

Algorithm design and analysis techniques: The development of methods to design algorithms and protocols, to analyse their performance and to explore the inherent complexity of computational, communication and description problems, as well as to study complexity classes and their mutual relations - as a way to get a deeper understanding of various phenomena concerning efficiency and of their interrelations.

POWER of NEW METHODOLOGY - I.

- It brings new dimension to both old methodologies;
- It brings into new heights an enormous power of modeling and simulations;
- It utilises an enormous power of visualisation;
- It utilizes an enormous potential that the study of virtual worlds brings for understanding of the real world;
- It utilises an enormous power of (sophisticated) search techniques.

POWER of NEW METHODOLOGY II

- It utilises the observation that designs of information processing models (robots) bring deep understanding of the real life phenomena;
- New, informatics-driven, methodology, allows us also to change the centuries old vision of the science.

  Namely, the vision that science is to achieve its goals mainly by discovering, isolating and studying very primitive phenomena and processes.

POWER of NEW METHODOLOGY - III.

- New informatics-based methodology develops not only tools to exploit better the existing knowledge, but also contributes to the development of new tools to discover new knowledge.
- New methodology is used, as a very powerful and inspiring one, also in all areas of scholarship, learning and art.
- Power of this new methodology is so large also for such areas of society as politics and justice that one can expect quite soon materialisation of the famous Leibniz's vision of an ending of future legal disputes : "Gentlemen, let us calculate".

POWER of NEW METHODOLOGY - IV.

- New methodology starts to make hard some (most) of the soft-sciences.
- New methodology provides intellectual frameworks and tools to consider investigation also of complex and large systems, in their full complexity.

REDUCTIONISM versus HOLISM - I.

Reductionists view the entire universe, from its beginning in time and from particles (or Planck scale) to galaxies, as being governed by simple rules that are comprehensible by human mind.

Design of microscopes and telescopes, that allowed to see and explore microcosmos and macrocosmos, and their enormously increased performances, that were behind of many discoveries, much supported the reductionists view of science.

REDUCTIONISM versus HOLISM - II.

Reductionism lead to the view that physics is the only fundamental natural science, chemistry is applied physics, life is created by complex chemical reactions and various particular steps in the biological evolution, as speciesation and emergence of humanity, are mainly due to various random events and historical circumstances.

Non-reductionists position is to a large extend based on the possibility of computers to model interactions of huge number of simple elements and to build computer models so complex that they are not understandable by human mind and that this brings very new tools for science to ask and to answer new questions.

CASE STUDY I -MODELLING and SIMULATION

CASE STUDY I

-MODELLING and SIMULATION

DEVELOPMENT of GOALS of SCIENCE - DRIVING METHODOLOGIES - I.

- Classical Greek Period - methodology was driven by philosophy. The starting and basic assumption was that phenomena of nature are not under the control of Gods, and that human mind is powerful enough to discover basic qualitative axioms (truths) and deduction rules. The goal was to understand CAUSES OF PHENOMENA of nature and to find out WHY THINGS HAPPEN.

DEVELOPMENT of GOALS of SCIENCE - DRIVING METHODOLOGIES - II.

- Galileo time - methodology was driven by mathematics and by a belief that nature should be the source and verifier of our basic and derived knowledge.The starting and basic assumption was that God created mathematically expressible world and let it to run according to mathematically expressible rules. The goal of science was to discover that rules (and by that to demonstrate geniality of the God).
  Understanding developed that it is not absolutely necessary to know the causes of phenomena, but that is to a large extent sufficient to know explicitly important RELATIONS among some quantitative characteristics OF PHENOMENA. They can be sufficient to make useful and important predictions concerning the behaviour of the phenomena of physical nature and to understand HOW THINGS HAPPEN.

VIEWS of GALILEO

- Galileo: "It does not seem expedient to me to investigate what may be the cause of acceleration; the chief concern must be to explore the law according to which acceleration takes place.
- But hitherto I have not been able to discover the cause of those properties of gravity from phenomena and I frame no hypothesis ..... To us it is enough that gravity does really exists, and act according to the laws which we have explained, and abundantly serves to account for all the motions of the celestial bodies, and of our see."

DEVELOPMENT of GOALS of SCIENCE - DRIVING METHODOLOGIES - III.

- Post-computer era - methodology is driven by informatics . The starting and main assumption is that we often do not need to know neither causes of phenomena nor explicitly relations between their important quantitative characteristics, it is sufficient to design an (evolving - capable to learn) oraculum, an information processing MODEL OF PHENOMENA, that can be used to answer, or even visually demonstrate, sufficiently well, important questions about the phenomena.

WISDOMS

- Theories (Models) are lies that tells you true.
- The art is a lie that helps us to see truth.

  Pablo Picasso
- An important question is: how complex can be simple systems and how simple can be complex systems?

CASE STUDY - I. MODELING and SIMULATION

CASE STUDY - I.

MODELING and SIMULATION

MODELS

Models are always idealisation, approximation, guesses because our knowledge of nature is always incomplete and approximate at the best.

For models of phenomena of nature we have always to ask how much are they based on observation and measurement of accessible phenomena; how much are they based on informed judgement and how much on convenience?

Models are most useful when they are used to challenge existing formulations, rather than to validate or verify them. Models always capture only some essences of the phenomena they model - one should not assume that they could capture the whole phenomena.‘

MODELLING and SIMULATION - I.

- Modeling and simulations are key methods in such disciplines as earth-sciences, environmental sciences, (system) biology, astronomy, economics, social sciences where one needs an understanding of complex phenomena produced by interactions of various processes and/or neighbouring elements.
- Using simulation one can follow processes that are otherwise invisible because of being too small, or too slow, or too large, or too fast, or too dangerous.

MODELLING and SIMULATION - II.

- Simulations and modeling provide a completely new view of reality that is the closest to reality that we can have.
- More complex modelled systems are and more modifiable they are, for example by changing (or even reversing) the flow of time, or by zooming, or changing the flow and amounts of various resources, better understanding of the reality can be obtain through them.
- In general modeling and simulations allow in very effective way to use virtual worlds to get deep insights into the real world.

MODELLING and SIMULATION - III.

- Some complex phenomena of nature, society or of people cannot be dealt with using experimental methods due to the practical or even ethical considerations. Modeling and simulations are then the main available methodology to study them.
- Modeling and simulation of the human brain is very challenging tasks. It is expected to bring huge benefits for medicine (to treat brain diseases and disorders via drugs and implants, to deal with memory, hearing and vision problems,...) and also for informatics (to inspire the design of computers with novel types of parallelism, connectionism and distributivness as well as of the AI systems of novel types).

A BIT of HISTORY - I.

- The first large scale and influential computer simulation was done in 1953 by E. Fermi, J. Pasta and S. Ulam in Los Angelos laboratory - to understand behaviour of large non-linear systems and their supposed tendency to decay to states of ever greater disorder.
- One of the basic questions concerning simulation/modelling is for which natural phenomena and processes we can do simulation/modelling by systems that are significantly simpler.
- A variety of interesting results have been obtained concerning classical simulation of certain classes of quantum phenomena, especially circuits.

A BIT of HISTORY - II.

- The Blue Brain project of École Polytechnic in Laussane, which started in 2005, aims at the simulation of the entire human brain down to the molecular level.
- One of the most powerful current computer system, Roadrunner, with a theoretical (demonstrated) performance 1.7 (1.1) petaflops and occupied space 260 m^2, has been developed by the IBM for Los Alamos National Laboratory to model the decay of US nuclear arsenal.

SIMULATION and MODELLING of CELLS

- Cells can be seen as well-organized autonomous systems of many discrete interacting components and biology accumulated huge amounts of knowledge about their basic components.
- This is especially true for the main components/macromolecules of cells: nucleic acids polymers - DNA and RNA - proteins and membranes.
- However, all that provides very little insights into how cells work as a whole and how they process information.
- What is therefore needed is to understand cells and their main components, and also other biological elements, as systems of interacting components that abstract from their chemistry.
- To do that through continuous mathematical models would require to simulate huge amounts of differential equations and that is unfeasible.
- A way out seems to be to create automata models of cell behaviour in the form of several discrete, concurrent, asynchronous and heterogeneous models of interacting elements.
- The interaction models of the above type are some of the most difficult for Informatics to handle and analyse, and so deep results of the concurrency theory and of other areas of scientific Informatics need to be used.

BRAIN MODELING

- In US, during period 1991-2004, they run big project "Human brain project" oriented on simulation of the brain.
- In Europe, in 2013, European commission launched 1,2 milliard EUR project entitled "Human Brain Project" oriented on brain simulation, see Chapter 10, with following goals: to gain fundamental insides into what it means to be human; to develop new treatments for brain diseases and to develop revolutionary new ICT - see Chapter 9.

EMERGENCE of NEW SIMULATION METHODOLOGY - I.

An understanding that modeling can be seen as a key element of new methodology goes actually back also to von Neumann who said:

- The sciences do not try to explain, they hardly try to interpret, they mainly make models. By a model is meant a mathematical construct which, with the addition of certain verbal interpretations, describes observed phenomena. The justification of such a mathematical construct is solely and precisely that is expected to work.

EMERGENCE of NEW SIMULATION METHODOLOGY - II.

The above vision can nowadays be better formulated by the following "improvement" and "modernisation" of von Neumann thought.

- The sciences do not try to explain, they hardly try to interpret, they mainly make models. By a model is meant an informatics system which, with the addition of certain verbal interpretations, describes observed phenomena. The justification of such models/systems is solely and precisely that is expected to work.

MODELS and REALITY

Simulations and modeling provide a completely new view of reality that is often closest to reality that we can have.

More complex such models are and more modifiable such simulations are, for example by changing (or even reversing) the flow of time, or by zooming, or changing the flow and amounts of various resources, better understanding of the reality can be obtain through them.

In general modeling and simulations allow in a very effective way to use virtual worlds to get deep insights into the real world.

USEFULNESS of HUGE MODELS

Through models we can create systems that can have enormous predictive and explanatory value in spite of the fact they are behind human potential to understand them fully and globally.

We can however, using informatics tools that enhance our intellectual capabilities, to understand them locally and to make them to give us answers to important inquires concerning phenomena they model.

SUCCESS STORIES of MODELLING

- Computer modeling and simulation of big molecules and their dynamics is one of the big success stories of computer modeling and simulation with big impacts on design of useful molecules and drugs. Several factors are behind. Powerful (parallel) computers, good sets of test data and deep knowledge and expertise available. All that open new windows to see the world of molecules and gave rise also to new fields such as protein dynamics.
- One of the big goal of modeling is to understand how biological systems work. For example cells. Biological systems can be seen as well-organized autonomous systems of many discrete interacting components.

FUNDAMENTAL QUESTIONS CONCERNING MODELLING

- Which processes can be simulated sufficiently well by systems that are (much) simpler than systems to be simulated?

EXAMPLES of MODELLING and SIMULATIONS

- Simulations of battles, impacts of natural or men-made disasters, earth and atmosphere processes and so long require the largest computer power available;
- Simulations for entertainment industries, especially motion movies, computer and video games, belong to the most interesting, stimulating and challenging ones.
- Improvement in the area of computer technology, but also in the design of algorithms and software systems open usually new domains where modeling and simulations are feasible.
- Technically, simulations can be deterministic or randomized, steady-state or dynamic, discrete or continuous.

CASE STUDY II - VISUALISATION

CASE STUDY II - VISUALISATION

WISDOMS

- A picture is worth of thousand words.

  Anonymous
- Of all our inventions for mass communications, pictures still speak the most universally understood language.

  Walt Disney
- Pictures and shapes are but secondary objects and please or displease only in the memory.

  Francis Bacon

ESSENCE and GOALS of VISUALISATION

Visualisation is science, technology and art to use graphic to represent and display data, information and knowledge via images - graphic is here a communication medium.

Rendering is the process of generating an image (of two, three or more dimensional objects or processes) from their high-level descriptions in a strictly defined formal language using computer programs.

A high level description/representation of an image uses various (of a lower level) primitive elements. In a schematic drawing, they may be line segments, curves; in a 3D rendering they may be triangles and polygons in space and in a graphical user interface they may be windows and buttons. Descriptions should contain geometry viewpoint, textures, lighting and shading information

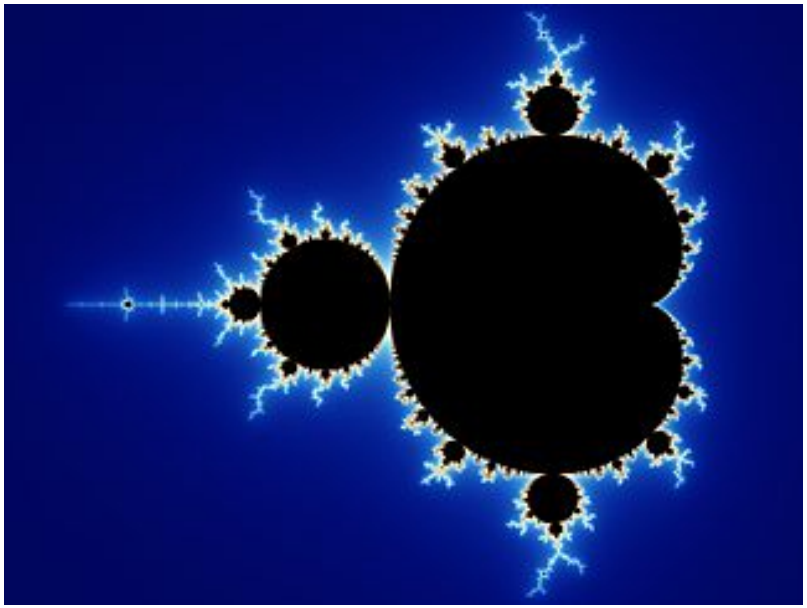
THREE GOALS of VISUALISATION

On one side visual representations are to take the advantage of the human eye's broad bandwidth into the mind to allow users to see, explore and understand large amounts of information and complex processes at once.

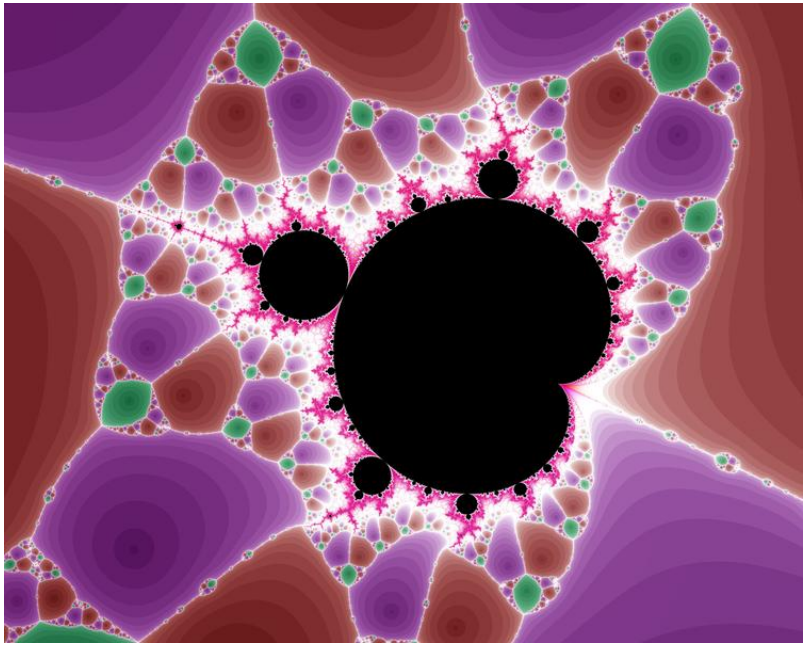
On the second side, visualisation helps us to see those parts of our cosmos and microworld that are invisible by our unaided eyes.

On the third side, visualisation can help us to get new understanding of phenomena not only of the real world, but also of the world of formal science, mathematics and information processing worlds and to develop in these areas new knowledge, hypotheses and theories discovery tools.

MANDELBROT SET - I.



MANDELBROT SET - II.



MANDELBROT SET GENERATION

Mandelbrot set is a set of values of c in a complex plane for which the orbit of 0 under the iteration of complex quadratic polynomial

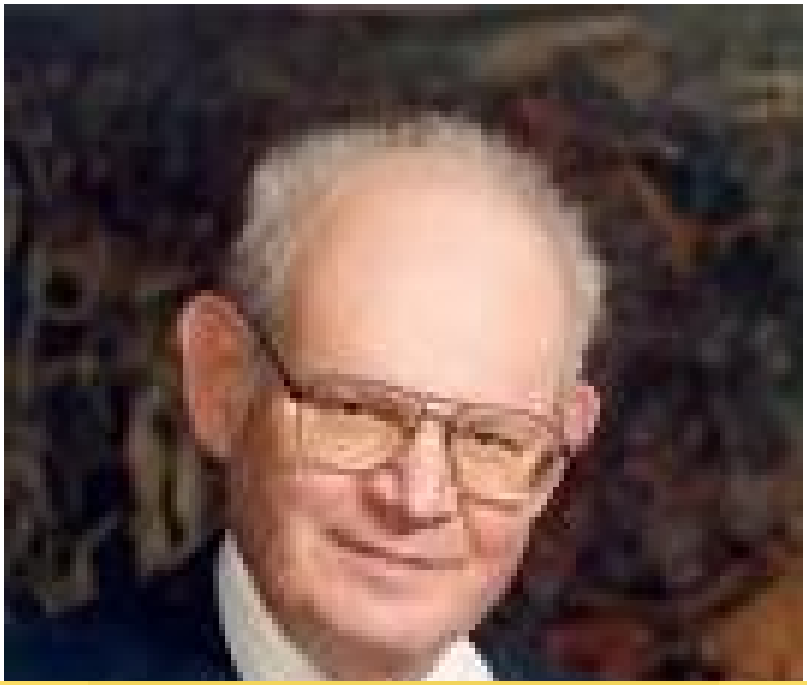
$$z_{n+1} < - - z_n^2 + c$$

remains bounded.

That is a complex number c is a part of the Mandelbrot set if, when starting with $z_0 = 0$ and applying the iteration repeatedly, the absolute value of z_n remains bounded however larger n is.

Treating real and imaginary parts of complex numbers as coordinates pixels of points are colored depending on how fast sequence created by iteration converges.

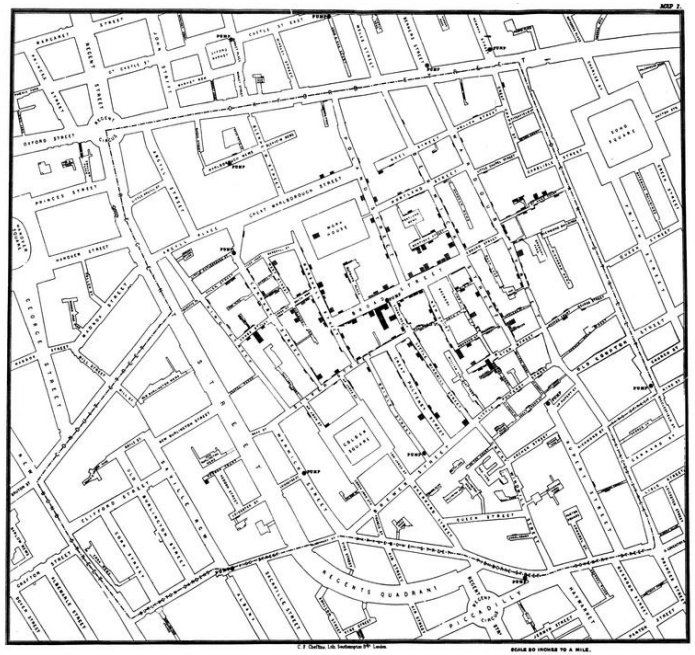
MANDELBROT



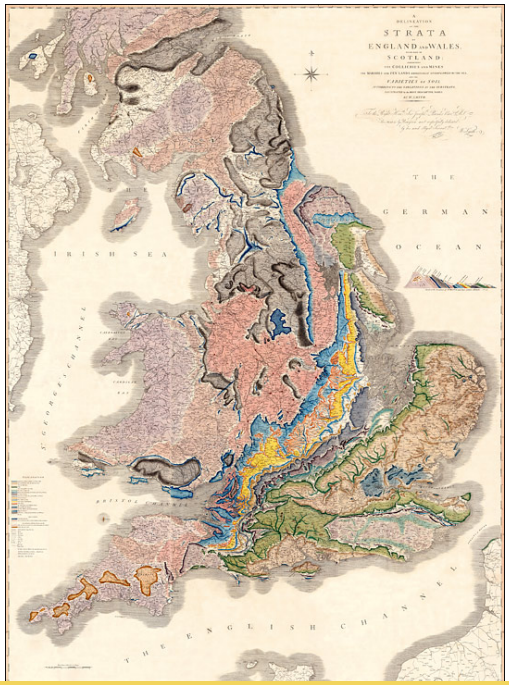
HISTORY of VISUALISATION

- Cave paintings can be see as first attempts for visualisation.
- Geographical maps of the world were very important products of visualisation (since the third century BC).
- William Playfair (1786) is usually mentioned as one defining various types of diagrams to visualise data (line graphs, bar charts, pie charts,...).
- In 1815 a geological map of England was published by William Smith - a map that was claimed to change the world (of visualisation).
- Visualisation of fractal and elements of nonlinear phenomena, especially attractors brought again a new dimension to visualisation.

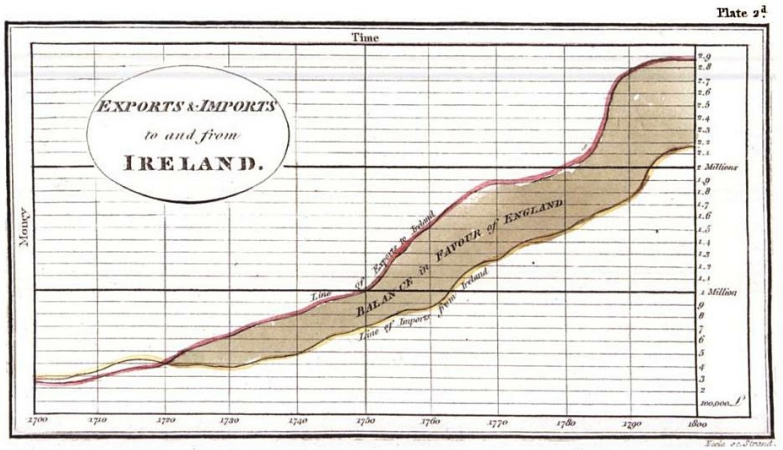
JOHN SNOW's MAP of 1984 CHOLERA OUTBREAK in SOHO



SMITH's GEOLOGICAL MAP of GREAT BRITAIN



PLAYFAIR's VISUALISATION



ANOTHER VIEWS of VISUALISATION

Visualisation can also be seen as a way of utilising methods derived from the study of human perception, graphical design, art and usability analysis.

Understanding of the human perception and of the potential of imaging technology are the key extreme factors in the development of visualisation principles, forms, methods and tools and also in using visualisation.

Visualisation is of importance also because we are visual beings who use sight as one of our key senses for data and information understanding and knowledge discovery.

A MORE TECHNICAL VIEW of VISUALISATION

In a more technical way, one can say that the goal of visualisation systems is to transfer logical description of complex information (logical elements and their relations as well as operations for their manipulations), by encoding their elements and their relations through perceptual elements and their relations.

The encoding has to be done in such a way that good/better understanding of complex information can be visually obtained - one can therefore transform logical descriptions of complex information to perceptual elements, procedures and their accompanying graphics.

VISUALISATION as a TOOL - I.

- Visualisation is nowadays one of the key tools to bring new quality into the learning and knowledge discovery process in practically all areas of human and societal activities.
- An original idea was to make static visualisations. New methods and technology allow to create dynamic and interactive visualisations and also visualisation displays in such a form that they can be well processed and understood not only by humans, but also by machines - computers, robots and so on.

VISUALISATION as a TOOL - II.

- While early visualisations were static objects, printed on paper or other fixed medium, modern visualisations are very dynamic processes, with the user controlling virtually all stages of the procedure, from data selection and mapping control to color manipulation and view refinement.
- While most of the visualisations are two dimensional, they often try to depict more-dimensional objects and not only the four-dimensional ones where time is the fourth dimension, but even more dimensional. 3D printers bring again a new dimension into the attempts to depict three dimensional objects and their relations.

VISUALISATION versus COMPUTER GRAPHIC

Computer graphics and visualisation have different goals though visualisation uses graphics much.

Visualisation is the application of graphics to display data by mapping data to graphical primitives (points, lines, areas and volumes) and rendering the display.

Computer graphic is predominately focused on creation of the interactive synthetic images and animations of three-dimensional objects most often where visual realism is one of the primary goals.

Visualisation does not emphasize visual realism as much as effective communication of information.

Moreover, many type of visualisation do not deal with physical objects and those that do are often communicating attributes of the objects that are not normally visible - such as material stress.

TYPES of VISUALISATION

There is a variety of important types of visualisation.

- Scientific (data) visualization
- Many-dimensional objects visualization
- Flow visualizations
- Geo-visualization
- Information visualization
- Software visualization
- Music visualizations

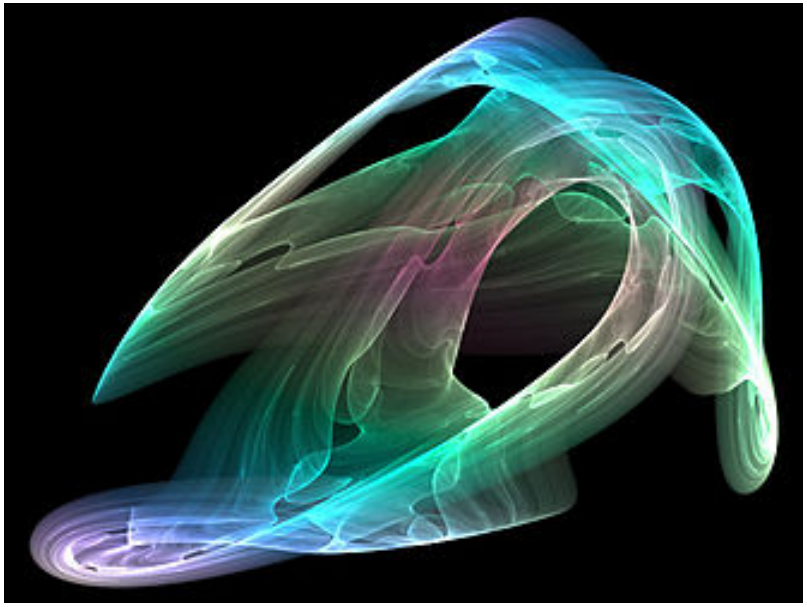
WHY IS VISUALISATION of SCIENTIFIC DATA of IMPORTANCE

Main reasons why such scientific simulations are of large importance are the following ones:

- Visualisations usually compress a lot of data into a single image where spatial distribution of data and coloring could much speed up comprehension of data and their analysis;
- Visualisations create images that allow to reveal easily correlations between different quantities, both in space and time;
- Visualisations allow to see data selectively and interactively - all in real time;
- Visualisations can furnish new space-like structures and by an observation of their creation and changes in time one can get important insights into their dynamics;

ATTRACTOR

An attractor is a set toward which a variable, moving according to a dictates of a dynamical system , evolves over time.



SCIENTIFIC DATA VISUALISATIONS APPLICATIONS

Visualisation is of large importance and can even be the key tool in huge amount of applications. It is even possible for an application to be totally driven by its visualisation. The main application areas of visualisation can be seen as follows:

- In the natural sciences (physics (especially astronomy, cosmology,...): biology (cell and molecular biology);...)
- In the earth and environmental sciences.
- In formal sciences;
- In applied sciences;
- In medicine - medical reconstructions for diagnoses and at surgeries
- In health care, transport, communication, economics,..

CURRENT VISUALISATION TOOLS

- Currently there are in widespread use several software tools that allow to exhibit complex three-dimensional data sets, graphs of very sophisticated functions and to use graphics in powerful new ways to display data.
- Examples of such tools are Geometer's sketchpad, Mathematica, Maple, Matlab.
- Such tools allow to see things more clearly, in more details and in a new way.

SLIP CASE

- Automobile company Volve used to work with clay models to design car bodies and used to be a long process.
- In the early 1980s, company created their team of experts to "computerize" the process, but they failed.
- More successful was the team headed by brilliant mathematician Bjorn Dahlberg, holder of prestigious Salem Prize in harmonic analysis.
- Dahlberg used sophisticated techniques from differential equations, differential geometry, linear programming, convex surface theory, harmonic analysis and other parts of mathematics to design software SLIP that revolutionized car surface design process.

DE-VISUALISATION MOVEMENTS - I.

- In the Golden era of Greek mathematics visualisation of geometrical objects played an important role.
- An introduction of analytical geometry around 1640, that initiated exploration of many dimensional objects, changed view on importance of visualisation.
- For long time visualisation did not have large importance because visualisation tools were primitive.
- There was also a sort of antivisualisation movement at the beginning of 20th century, represented especially well by an attempt of so called Bourbaki movement in France (after 1930), to make consequently picture-less presentation of all fundamental mathematics.
- It was only (namely) due to advances in computer and printer technologies that visualisation started to play a very important role again.

CASE STUDY III - ALGORITMISATION AND COMPLEXITY CONSIDERATIONS - I.

1. Finding a new data structure or an algorithm can revolutionize the way scientists think about a problem and can even create a research area. Some examples:

- Strassen's exponentially faster matrix multiplication algorithm (1969).
- Discovery of NP-complete problems by Cook (1971).
- Freivalds' matrix multiplication testing randomized algorithm (1979);
- Shor's quantum factorization algorithm (1994).

CASE STUDY III - ALGORITMISATION AND COMPLEXITY CONSIDERATIONS - II.

2. Study of inherent complexity (computational, communication and descriptional) of phenomena and especially feasibility issues is an important methodology to influence and guide research directions.

Complexity considerations played the key role especially in the following two areas:

- Quantum information processing - it was due to complexity outcomes that it got clear that this area may have big potential. The complexity theory keeps having deep impacts on the whole development in quantum information processing and communication. One can even say that one of the goals of the quantum complexity theory is to challenge our basic intuition how the physical world behaves.
- Security - in a broad sense - computational complexity consideration plays key role in modern secret key and also in public-key cryptography and in all other areas of security. , secrecy, anonymity, privacy and trust.

WHY von NEUMANN DID NOT DISCOVER QUANTUM COMPUTING

von Neumann is one of the fathers of both quantum physics and modern computers. It is therefor natural to ask why he did not come with the idea to develop quantum computers.

Very likely it is because:

- No complexity theory was known (and needed)
- The concept of randomized algorithms was not known.
- No public key cryptography was known (and needed).
- Progress in physics and technology was far behind what would be needed to make even rudimentary implementations.

ALGORITHMS GENERATING INDUSTRIES

- Google's 35 billion-euro business model largely rests on algorithms such as Page ranking - a fix-point algorithm that made WEB searches practical.
- The music industry has been revolutionized by standards such as MP3 and its defining algorithms and data structures.

FAMOUS ALGORITHMS - I.

- Binary search
- Dynamic programming
- Euclid algorithm
- Fast Fourier transform
- Gauss elimination algorithm
- Finite element methods
- Monte Carlo algorithms
- Newton method
- QR-algorithm to compute eigenvalues
- Random walk algorithms
- Runge-Kuta method
- Simplex method for linear programming

FAMOUS ALGORITHMS - II.

- Dijkstra's short path algorithm
- Krylov subspace iteration method
- Metropolis algorithm
- Page ranking algorithm
- Quicksort (T. Hoare, 1962)
- RSA encryption and signature
- Shor factorization algorithm
- Viterbi algorithm

POWER of COMPLEXITY THEORY as a METHODOLOGY - I.

- Extraordinary power of methods and outcomes of the computational complexity theory, as of a tool for advances of science, has been especialy well demonstrated in the development of quantum information processing and communication.
- It was mainly due to the outcomes of the quantum complexity theory that an understanding has emerged that quantum computers could pay off. Physics had at that time no tools to see that quantum computers could pay of.
- The complexity theory keeps having deep impacts on the whole development of the area of quantum information processing and communication including quantum cryptography.

POWER of COMPLEXITY THEORY as a METHODOLOGY - II.

- One can even say that one of the goals of the quantum complexity theory is to challenge our basic intuition how the physical world behaves and to understand two of the great mysteries of the 20th century: what is the nature of quantum mechanics and what are the limits of computation.
- The complexity theory approaches can also lead us to ask better questions about the quantum nature – nontrivial, but answerable questions – questions that put old quantum mysteries in a new light. Quantum computational complexity also allows us to demonstrate that some physical phenomena are not possible - by showing that possibility of these phenomena would allow us to solve efficiently complete problems for some computational complexity classes that are believed to be larger than the classes **P** or **BPP** or **BQP**.

CLASSIFICATION of IMPOSSIBLE - COMPLEXITY THEORY APPROACH - I.

Some of the ways to show that some quantum event or phenomenon E is impossible:

- To show that E would imply superluminal communication.
- To show that E would violate NO-cloning theorem.
- To show that E would imply that problems from some class (likely) larger than **P** would be solvable in polynomial time
  (and that also allows to classify impossible tasks).

CLASSIFICATION of IMPOSSIBLE - EXAMPLES

- Abrams and Llyod (1998) showed that under the assumption that quantum mechanics is non-linear, more exactly that Weinberg's model of quantum mechanics is valid, one can solve in polynomial time **NP**-complete problems.
- Aaronson (2004) has shown that if arbitrary one-qubit nonlinear gates can be implemented without an error, then **PSPACE**-complete problems can be solved in polynomial time.
- Aaronson has also shown that if so-called *post-selection* is allowed, then **PP**-complete problems can be solved in polynomial time.

THE ROLE of QUANTUM COMPLEXITY THEORY

- One of the goals of quantum complexity theory is to challenge our basic intuition how physical world behaves.
- Quantum complexity theory is of great interest because one of its goals is to understand two of great mysteries of 20th century: what is nature of quantum mechanics and what are the limits of computation.
- It would be astonishing if a merge of such important areas would not shed light on both of them and would not bring new great discoveries.
- Taking complexity theory perspective can lead us to ask better questions about quantum nature – nontrivial, but answerable questions, which put old quantum mysteries in a new light even if they fall short of answering them.
- One way quantum complexity theory obtains insights into the quantum world is through the study of main quantum complexity classes **BQP, QMA** and their complete problems.

CASE STUDY IV - PROOFS and AUTOMATED REASONING

PROOFS

WHAT IS A PROOF?

The concept of proof is one of the most basic ones in modern science.

Any discovery and development of a broader view of this concept, as well as of tools to make or verify proofs, can therefore have far-reaching impacts on science.

Some citations:

- A nice proof makes us richer (Jurij Manin).
- A proof is whatever convinces me (S. Even)
- The glory attached to the creativity involved in finding proofs, makes us to forget that it is much less glorified procedure of verification which gives proofs their value. (O. Goldreich)
- Proofs depend much on the audience. We prove things in a social context and address them to a certain audience. W. P. Thurston
- Proof is lingua franca of mathematics.
- Proof is a rhetoric device for convincing someone else that a mathematical statement is true.

WISDOMS

- There are proofs that date back to Greeks that are still valid today.

  Andrew Willes

- In a sense, mathematicians have been most advanced by those who distinguish themselves by intuition rather than by rigorous proofs.

  Felix Klein

PROOFS and THEIR DEVELOPMENTS - I.

- Introduction of the concept of the proof, and deductive mathematics, is often considered as the main contribution and feature of Greek mathematics from its Classical period, especially due to Euclid's **Elements**.
- Euclid was the first who systematically used precise definitions, axioms and strict rules of logic and who systematically proved every statement.
- This idea was already almost completely abandon during Alexandrian period of Greek mathematics.
- Actually for almost 2000 years achievements of the classical Greek period, concerning methodology of mathematics, were virtually ignored, as not much useful.

PROOFS and THEIR DEVELOPMENTS - II.

- Cicero said: *The Greeks held the geometer in the highest honor; accordingly, nothing made more brilliant progress among them than mathematics. but we have established as the limit of this art its usefulness in measuring and counting.*
- The Greek virtue of insisting on exact concepts and proofs can now be sen as a defect so far as creative mathematics was concerned.
- At the end of 17th century mathematicians had virtually dropped the idea of clearly defined concepts and deductive proofs. Mathematics was so inspired by developments of science and vice verse that a fusion of mathematics and vast areas of science was detectable

PROOFS and THEIR DEVELOPMENTS - III.

- In the 18th century mathematics and physics virtually merged, and the physical meaning of the mathematics guided the mathematical steps and often supplied partial arguments to fill in non-mathematical steps.
- The concept of a proof in mathematics, as of a device to communicate truth of some assertion, can be seen as an equivalent of the concept of reproducible experiment in physics and chemistry.

PROOFS and THEIR DEVELOPMENTS - IV.

- The above period of mathematics, in which mathematicians did not care about foundations of their work is sometimes called heroic age of mathematics. Some views of that time:
  - Such subtleties as the Greeks worried about we no longer need.
  - *Why to go to the trouble of proving by complicated reasoning things which one never doubts in the first place, or of demonstrating what is more evident by means of what is less evident.*
- On a more philosophical basis, mathematicians, believed that they are unearthing the mathematical design of universe and therefore they were confident in the truth of their outcomes.

PROOFS and THEIR DEVELOPMENTS - V.

- Descartes, one of founders of modern science, believed that axioms of mathematics are true and mathematical reasoning is sound on the conviction that God would not deceive us, so that to deny the truth and clarity of mathematics would be to deny God.
- Actually from 200 B.C. till 1870 almost all mathematics rested on an empirical and pragmatic basis.
- At the end of 17th century mathematicians had virtually dropped the idea of clearly defined concepts and deductive proofs.

PROOFS and THEIR DEVELOPMENTS - VI.

- At the end of 17th century mathematics was so inspired by the developments of science and vice verse that a fusion of mathematics and vast areas of science was detectable and mathematics started to rely more and more on scientific results to justify its own procedures.
- In the 18th century mathematics and physics virtually merged, and the physical meaning of the mathematics guided the mathematical steps and often supplied partial arguments to fill in non-mathematical steps.

PROOFS and THEIR DEVELOPMENTS - VII.

- By 1900 the goal of establishing mathematics rigorously was formulated by Hilbert. The rigorization was (believed to be) achieved by axiomatising various branches of mathematics.
- At the end of the 19th century a variety of strange functions have appeared as well as paradoxes in set theory and it got clear that a level of rigor in definitions, theorems and proofs has to be increased
- At the beginning of the second half of 20th century it has been discovered that Hilbert's notion of proofs is too strong - some simple stated theorems could have proofs only longer than number of particles in universe.

PROOFS and THEIR DEVELOPMENTS - VIII.

The existence of culture of theorems and proofs can be seen as rigorous and well-tested standard for formulation and recording ideas and discoveries that have as a consequence the fact that mathematical ideas and discoveries stood up under the test of time.

MATHEMATICS VERSUS PHYSICS from PROOFS POINT of VIEW

In one of the recent papers: J.Borwein et all (1995): Experimental mathematics: A discussion, the preoccupation of mathematics with proofs is questioned:

Mathematics is seen as consisting of two parts

- theoretical mathematics - speculative as theoretic physics
- rigorous mathematics - giving evidence (using proofs) as experimental physics does.

VIEWS of ZEILBERGER - I.

Computers have already started doing to Mathematics what the telescopes/microscopes did to astronomy/biology.

In the future not all mathematicians will care about absolute certainty, since there will be so many exciting new facts to be discovered: mathematical pulsars and quasars that will make Mandelbrot set to look like a Galilean moon.

We will have (both human and machine) professional theoretical mathematicians who will develop computational paradigms to make sense out of the empirical data and who will keep Fields medal along with (both human and machine) experimental mathematicians.

Will there still be a place for mathematical mathematicians?

Dorov Zeilberger

HILBERT's CONCEPT of PROOFS

The currently dominating view of the proof is due to Hilbert:

A proof in a formal system F is a sequence of statements such that each of them is either

- an axiom of F;
- or it can be derived from previous statements of the sequence using one of the finitely many deduction rules of F

DIFFICULTIES WITH the HILBERT's CONCEPT of PROOF

- **Gödel Incompleteness result.** For any sufficiently powerful formal systems there are correct statements the truth of which cannot be proven in that system.
- **Chaitin Incompleteness result**: In any proof systems one can show randomness only a finitely many random strings.
- **Computational complexity result** There are simple Boolean formulas each proof of which has to be longer than the number of particles in universe.
- **Proof puzzle** What is more difficult

  to find a proof or **to verify a proof**

  This is a problem equivalent to the **P=NP** problem.

VERY LONG PROOFS - I.

1799 Ruffini published a more than 500 pages proof of "Abel-Ruffini theorem" - in 1824 N. Henrik published another 6 pages proof.

1974 Thomson published more than 700 pages classification of "N-groups"

1974 Deligne published more than 700 pages proof of "Ramanuyan and Weil conjecture"

1974 Appel and Haken published 741 pages proof of "4-colour theorem" with strong computer support.

1980 Almgren published 1728 pages proof of ...

1983 Gorenstein, Lyon and Aschbacher published 890 pages proof of "Tichotomy theorem"

1983 Hejbal published a 1322 pages proof of "General Selberg trace formula".

VERY LONG PROOFS - II.

2000 Almgren published 955 pages proof of "Regularity theorem".

2004 Aschbacher and Smith published 1221 pages classification of "Quasi-thing groups"

2004 Classification of finite simple groups has 10,000 to 20,000 pages

- As a very long proof has been considered
  - in 1900 a proof having 100 pages;
  - in 1950 a proof having 200 pages;
  - in 2000 a proof having 500 pages.
- In many cases traditional proofs may be set aside in favour of experimentation by testing of thousands or millions of examples by computers.
- There are attempts to make a proof in a distributed way using thousands of provers and it is expected that that can take even 20 years.

VIEWS of ZEILBERGER - II.

I can envision an abstract of a paper (around 2100) that reads, "We show in a certain precise sense that ..... is true with probability larger than 0.99999 and that its complete truth could be determined with a budget of $ 10 billions."

As absolute truth becomes more and more expensive, we could sooner or later come to accept the fact that only few non-trivial results could be known with old-fashioned certainty.

PROBABILISTIC PROOFS

Three new types of proofs have played a crucial role in theoretical informatics in the last 15 years:

- Interactive and non-interactive (one- and many-prover) proof (systems)
- Zero-knowledge proof systems.
- Probabilistic (holographic) proof (systems).

PROBABILISTIC PROOF SYSTEMS

The concept of probabilistic checkeable proofs (PCP), or *transparent* or *holographic* proofs, is another great/shocking idea concerning proofs.

Informally, PCP proofs are proofs that are written down in such a way that one needs to look only to (very) few randomly chosen bits of it in order to find out whether the proof is correct with (very) high probability.

The hard task is, however, to encode a given proof in such a way that randomized checking is possible.

Famous **PCP-Theorem** says that every **NP**-complete problem/language admits a probabilistically checkeable polynomially long proof.

PCP THEOREM

Intuitively, the PCP-theorem says that for some fixed (and universal) constant k, for every n, any mathematical proof of length n can be rewritten as a (different) proof of length $poly(n)$ that is formally verifiable on 99% by a randomized algorithm that makes only k queries to the proof.

One can also prove that each proof can be rewritten in such a way that it is enough to check 11 randomly chosen bit in order to verify the proof with probability at least $\frac{1}{2}$.

AUTOMATED and SEMI-AUTOMATED PROOF SYSTEMS and PROOFS CHECKERS - I.

- A view of proofs as "whatever that convinces me" has started to be recently again pursueded.
- The development of proof assistance systems and proof checkers has started already in 1966.
- Proof checkers are getting faster and faster more powerful to such an extend that they have a chance significantly increase proving power of human.
- Proof checkers started to proof theorems beyond the proving power of mathematicians at that time.

AUTOMATED and SEMI-AUTOMATED PROOF SYSTEMS and PROOFS CHECKERS - II.

- First famous result that created a lot controversy was computer assisted proof (741 pages) of the four colour theorem in 1976.
- Another famous result was the proof of Robbin's conjecture in 1996 that had resisted attempts for proofs for 60 years.
- The outcomes of proof checkers started to indicate that the concept of creative thinking may need to be challenged - proof checkers can take very different path to come to the same conclusions.
- As a consequence it has been discovered that there is a very thin line between mechanical and creative actions.
- Some proofs are so complex that they can be designed and checked only with the help of computers and proofs checkers.
- It starts to be feasible the situation that most of the proofs could be done by proof checkers to such an extend that the main role of mathematicians will be shifted to hypothesis and theorems formulation.
- As the result, mathematics research, as it is known today, as proof-centered, may disappear to a large extend. Mathematicians may focus mainly on hypothesis and theorem formation and proofs will be left to theorem provers.

ROBBINS CONJECTURE

In the 1930s, Herbert Robins conjectured that 10 usual simple axioms of Boolean algebra are equivalent to the following three axioms

$$x \cup (y \cup z) = (x \cup y) \cup z$$
$$x \cup y = y \cup x$$
$$\overline{\overline{x \cup y} \cup \overline{x \cup \bar{y}}} = x$$

(Last axiom is called "Robbins equation").
The conjecture was shown to be true first by William McCune from the Argonne National Laboratory in 1996 using the prover EQP (Equational Theorem Prover).

A human readable proof was later produced by Allen L. Mann in 2003.

PROOF CHECKERS

- Some proof checkers or proof assistant, as Mizar, try to imitate usual informal proofs and automate their most technical parts.
- Some other, as Metamath, work in a very different way, and can be used with every sort of formal systems - Metamath makes no assumption about the used logic - it is a substitution algorithm.
- One of the goals at the design of proof checkers is to proof as many, and as difficul as possible, theorems - there are yearly competitions in that.
- Another goal is to design as small as possible, but still powerful, proof checkers.

AUTOMATED THEOREM PROVERS

- There is currently a variety of automated theorem provers. Perhaps the most known are provers "E" for full first order logic, Otter (that developed into Prover 9), Vampire, Waldmeister.
- Most famous success (already in 1996) was the proof (after 8 days of work) of the already mentioned Robbins conjecture.
- Since that time a variety of theorem has been proven in (semi)groups, projective geometry and various logics.
- Intel and other companies use automated theorem provers regularly to verify correctness of their implementation of various operations (as division).
- Another applications of theorem provers are in software

PROOF ASSISTANT

- Proof assistants allow to take a usual proof and to turn it into the formal proof.
- Coq - a proof assistant which can automatically extract an executable program from specifications.
- HOL Light - a proof assistant for classical first order logic.
- There is nowadays a whole industry of the people who use computers to search axiomatic systems for new true statements and their proofs.

FOUR-COLOR PROBLEM STORY

- It is the problem to decide whether each only-countries map on the sphere can be colored with 4 colors only.
- The problem was formulated in 1852 and published in 1878.
- In 1879 A. Kempe published a proof that 4 colors are enough - in 1890 proof was found to be wrong.
- Problem got famous, but only partial solution were found - for example that 4 colors are enough for maps with at most 25 countries.
- In 1974 K. Appel and W. Haken from University of Illinois claimed that they solved positively 4-colors problem using 1200 hours of university supercomputer.
- There is still controversy whether such a proof can be accepted because no human can check it and several errors have been found (and always corrected) in the program used.
- In 1994 another computer assisted proof of four-color theorem has appeared and so far no error was found in the program used.
- In 2004 the 1994 proof was checked using a "mathematical assistant".

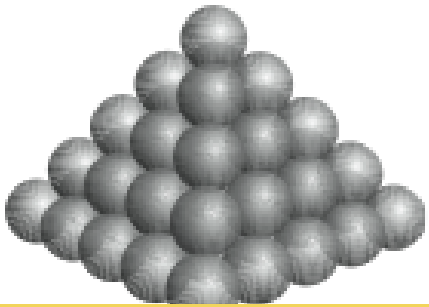
FlysPecK PROJECT

- FlysPecK project has as the goal to make formal proof (computer made with all details formally verified) for Thomas Halles' solution of Kepler's conjecture.
- The Kepler conjecture from 1611 asserts that the density of packing of congruent spheres in 3D is never greater than $\pi/\sqrt{18}$.
- Conjecture remained unsolved for almost 400 years until it was "solved" in 1998 by Halles using a lot of computer computations.
- To verify Halles' proof a group of 12 prominent mathematicians was established in 1999. After 4 years of work they declared that they are 99% sure that the proof is correct.
- In 2005 Halles started an initiative to make a complete formal version of his proof.
- It is expected that to turn the usual Halles' proof into a formal proof will require about 20 years of work of scientists from all over the world and will cost at least 5 millions of dollars.

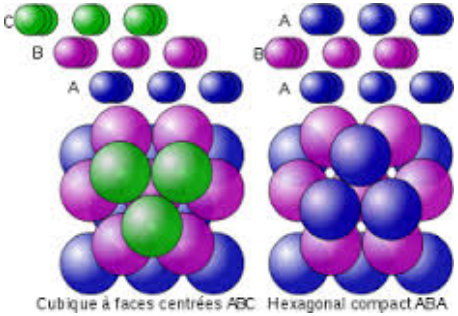
KEPLER's CONJECTURE

In 1961 Kepler formed his conjecture that no arrangement of equally sized spheres filling space has a greater average density (74%), than that of face-centered cubic packing or face-centered cubic packing. (Random packing has density 65%.)

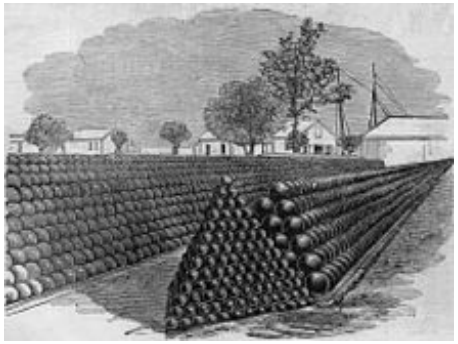
In 1831 Gauss showed that Kepler's conjecture holds if only regular packing arrangements are considered.



OPTIMAL PACKING of BALLS



Cubique à faces centrées ABC Hexagonal compact ABA



CASE STUDY V - ARTIFICIAL LIFE

- Basic question: Is life characterized by what it does or by what is made of?
- The goal of the artificial life research is to help to broaden the empirical database of the biology as a science.

MAIN APPLICATIONS of INFORMATICS BASED METHODOLOGY

- 20th-century science has been very successful, especially because its leading sciences, physics and astronomy, could work and make use of the laws well expressible through mathematical language using formulas and equations.
- After a shift of emphasis to the other areas of science with very complex systems, for example in case of life-, earth- and atmosphere-sciences, it started to be clear that it is not feasible to search for their understanding through the usual mathematical concepts as formulas and equations.
- In particular the situation has changes very significantly with a shift of emphasis in science to sciences that have to explore very complex phenomena not exhibiting regularity, symmetry and periodicity. For example, phenomena being described using nonlinear dynamics. Phenomena where quite clearly the approaches offered by mathematics are no longer adequate.
- The possibility to obtain huge amounts of data and to perform huge amounts of computations and to model very complex phenomena and to very complex visualisations have also radically changed science and technology, and ....

GRAND CHALLENGES of INFORMATICS DRIVEN METHODOLOGY - I.

- The development of tools to do modeling, simulation and visualisation of complex phenomena and processes.
- The development of methods and tools to specify, design, analyse and verify huge software systems.
- The development of tools to mechanized reasoning, learning and computation.
- The development of tools to discover the underlying simplicity in complex systems.

GRAND CHALLENGES of INFORMATICS DRIVEN METHODOLOGY - II.

- The development of methods and tools to manage huge data streams and sets.
- Truth-worthiness of the available digital networks information. This is already a huge problem and its size and importance are rapidly increasing, due to the decentralisation of the process of internet information creation.
- Openness problem for produced information and knowledge. A big challenge is to deal with scientific, engineering and methodological problems related to the task of digitalization and storing of increasingly growing amounts of available information and knowledge in a proper form that would allow its searching, mining, retrieving and verification by anyone and anytime in a reasonable time. In other words, a huge challenge is to provide world-wide open access to all available information.
- To make informatics thinking, methodologies and tools a basis of current education process on all levels and in all areas.
- To personalise learning processes.

APPENDIX

APPENDIX

WHAT GÖDEL's INCOMPLETENESS THEOREM ACTUALLY TELLS US?

- Gödels incompleteness theorem, that is usually seen as putting some limits on mathematics actually says "In order to get more you have to assume more".
- What does that actually (practically) implies in the era of super supercomputers?
- Adding one (some) additional axioms and to make deductions from them is technically, likely, not too much. However, due to the limiting capability of human brain it can be a lot?

INFORMATICS and PHILOSOPHY - I.

- Another big challenge for scientific Informatics is to revitalise philosophy of science and to find out what are goals, methods and tools of science in the new, informatics driven, era of science.
- Revolutions in physics and biology at the end of 19th century were accompanied by intensive philosophical discussions concerning goals, methods and tools of science.
- At the end of 20th century philosophy of science was almost dead.It was not fashionable to talk about science. The mode shifted to doing science.

INFORMATICS and PHILOSOPHY - II.

- Now with new scientific methodology most of the old questions come up in a new setting.
- For example, the goal of the science was to get an understanding of nature through understandable models and reductionism was the main method.
- It is getting clear that some key phenomena of nature and society are so complex that any reasonable and useful model of them has to be too complex to be fully understandable by people.
- It seems that our goals have to change - we can and have to be, in some cases happy with models that are too complex to be fully understandable by people if we would be able to work with them, with the help of

INFORMATICS and PHILOSOPHY - III.

- Informatics driven methodology is starting to have big impacts also on the philosophy of science. On the way we see goals of science, the character of scientific research, scientific methods, ways knowledge is produced, character of knowledge and of the scientific truth, level of precision we drive for and are happy with and so on.
- For example, one of the basic assumption of science was a belief that natural science should concentrate on the study of simple systems that can be sufficiently well specified by very few qualitatively different parameters and so they can be handled by human mind.
- Physics has been then the most successful because

NEW TYPES of SCIENTISTS

New methodology creates also a new class of scientists - in addition to theoreticians and experimentalists we have scientists that understand and master informatics-driven methodology and have different approach to and use different tools for knowledge acquisition, presentation (especially visualisation).