Universidad
Carlos III de Madrid
www.uc3m.es

AUTOMATA THEORY AND FORMAL LANGUAGES

UNIT 1: INTRODUCTION TO AUTOMATA THEORY AND FORMAL LANGUAGES

David Griol Barres - Computer Science Department – UC3M - dgriol@inf.uc3m.es

OUTLINE

- Why Automata Theory? History and Origin
  - *Disciplines of Computing*
  - *What "Computer Science" is?*
  - *Can any problem be solved using a computer?*
  - *Which are the capabilities and limitations of a computer?*
- Relationship with other knowledge areas
- Machines, languages and algorithms

OUTLINE

- **Why Automata Theory? History and Origin**
  - ***Disciplines of Computing***
  - ***What “Computer Science” is?***
  - ***Can any problem be solved using a computer?***
  - ***Which are the capabilities and limitations of a computer?***
- Relationship with other knowledge areas
- Machines, languages and algorithms

Why Automata Theory? History and Origin

- Disciplines of Computing according the *IEEE Educational Activities Board*:
  - Computer Engineering, Computer Science, Information Systems, Information Technologies, Software Engineering.

- **Computer Science**
  - Despite the enormous breadth of computer science, there are nonetheless concepts and skills that are common to computing as a whole.
  - "*All computer science students must learn to integrate theory and practice, to recognize the importance of abstraction, and to appreciate the value of good engineering design*".

Computing Curricula 2005. The Overview Report.
http://www.acm.org/education/curric_vols/CC2005-March06Final.pdf

Definitions of Computer Science

1. Science whose main objective is the resolution of problems by means of a computer. (i.e. digital system).

- *Norman E. Gibbs , Allen B. Tucker* (1986): *“the main objective of* ***Computer Science*** *has to be understood as the* ***systematic study of algorithmic processes and data structures that describe and transform information*** (***formal properties***)*, and not the development of programs”*

Gibbs, N. E. and Tucker, A. B. 1986. A model curriculum for a liberal arts degree in computer science. Commun. ACM 29, 3 (Mar. 1986), 202-210.

Definitions of Computer Science

2. Similar meaning that “Informática” in Spanish:

- “*Study of computers and all the phenomena related to them*”. (***Herbert Simon and Allen Newell***).
- “*Knowledge related to computers and computation”* (CSAB, *Computing Science Acreditation Board)*.

Why Automata Theory? History and Origin

Computer Theory

“Science whose main objective is to provide an answer to a set of questions that can be summarized in”:

It is previous to the invention of computer (even transistor).

Mathematical properties and applications of software and hardware.

- How to develop a program to solve a specific problem?
- Does the program actually solve the problem?
- How long and how many resources are needed for its resolution? (*temporal and spatial complexities*)
- Which computing model to use (Imperative, OOP, Logic Prog., etc.)
- Which problems **can and cannot be solved** using a computer?

Why Automata Theory? History and Origin

Direct application of concepts of Computer Theory

- Videogames: design of the behavior of the different characters.
- Compilers and Natural language Processing:
  - Lexical analysis.
  - Pattern recognition.
  - Parsing methodologies (e.g., grammars, XML).
  - Design and modification of new programming languages.
- Implementation of robust protocols (e.g., security systems).
- Cryptography (protocols).
- Construction of easiest and more efficient computer systems.
- Design (Sequential Machines → Code).
- SW to design and evaluate digital circuits.
- Question answering (web).
- SW to verify that a system has a finite number of states.

Why Automata Theory? History and Origin

- ... "I think that to some extent, we are victims of our own success. There are so many ways of earning money, so many things to try, as many exciting addresses, which sometimes make us forget to think about the fundamentals of our discipline... We have to maintain a continuous between giving a total dominance to our tendency to try and use anything, and at the same time designing our experiments using the scientific method, making sure that these foundations arewell balance... **We must remember that it is a scientific discipline and not only a branch of high technology"**...

Richard M. Karp
(Turing Award 1985)

Why Automata Theory? History and Origin

To learn more…

Bibliography

Educational Activities Board of IEEE

http://www2.computer.org/portal/web/education

- Computing Curricula 2005: The Overview Report

  "Describes general principles and commonalities among all of the specific disciplines described in the other four volumes".

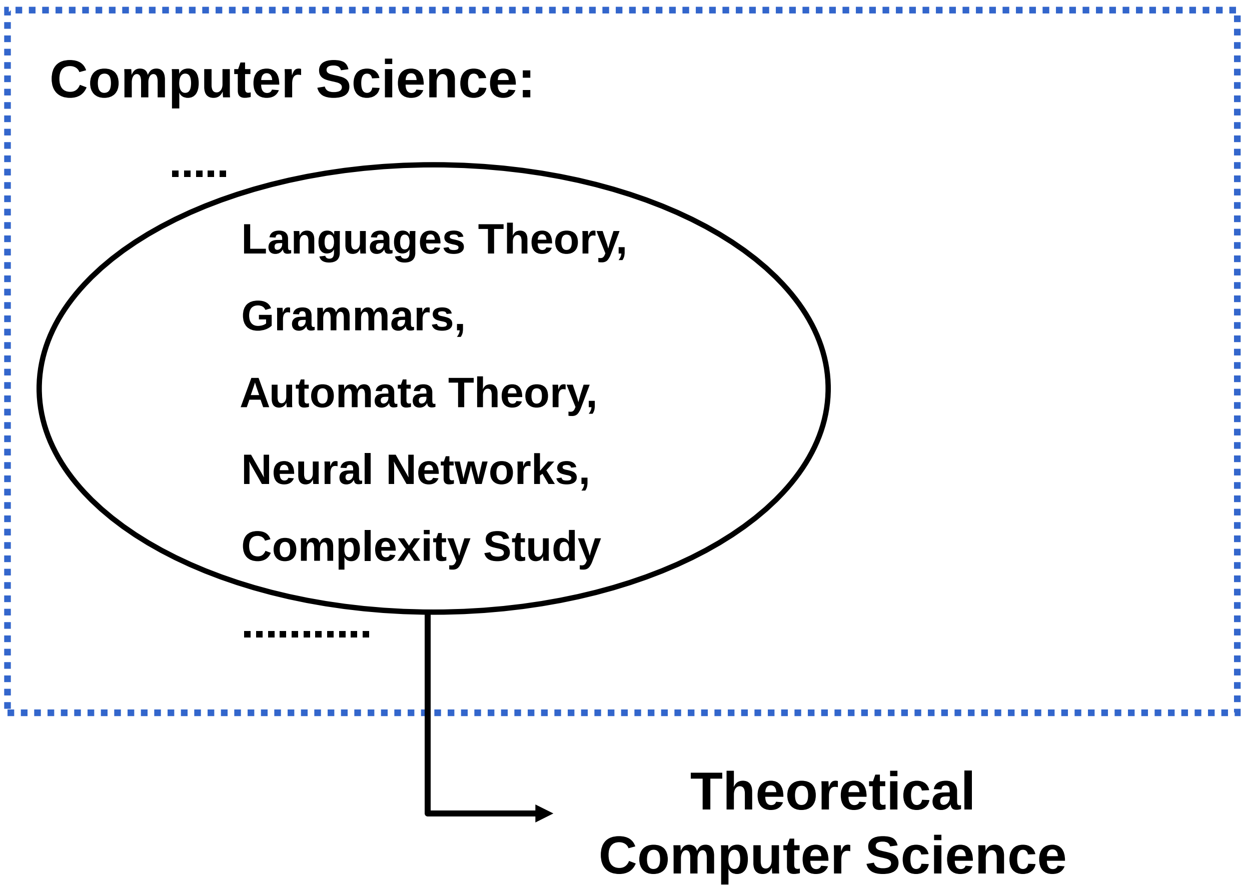
- Computing Curricula 2001 Computer Science

  "Outlines recommendations for undergraduate programs in Computer Science."

OUTLINE

- Why Automata Theory? History and Origin
  - *Disciplines of Computing*
  - *What “Computer Science” is?*
  - *Can any problem be solved using a computer?*
  - *Which are the capabilities and limitations of a computer?*
- **Relationship with other knowledge areas**
- Machines, languages and algorithms

Relationship with other knowledge areas



Computer Science:

.....

Languages Theory,

Grammars,

Automata Theory,

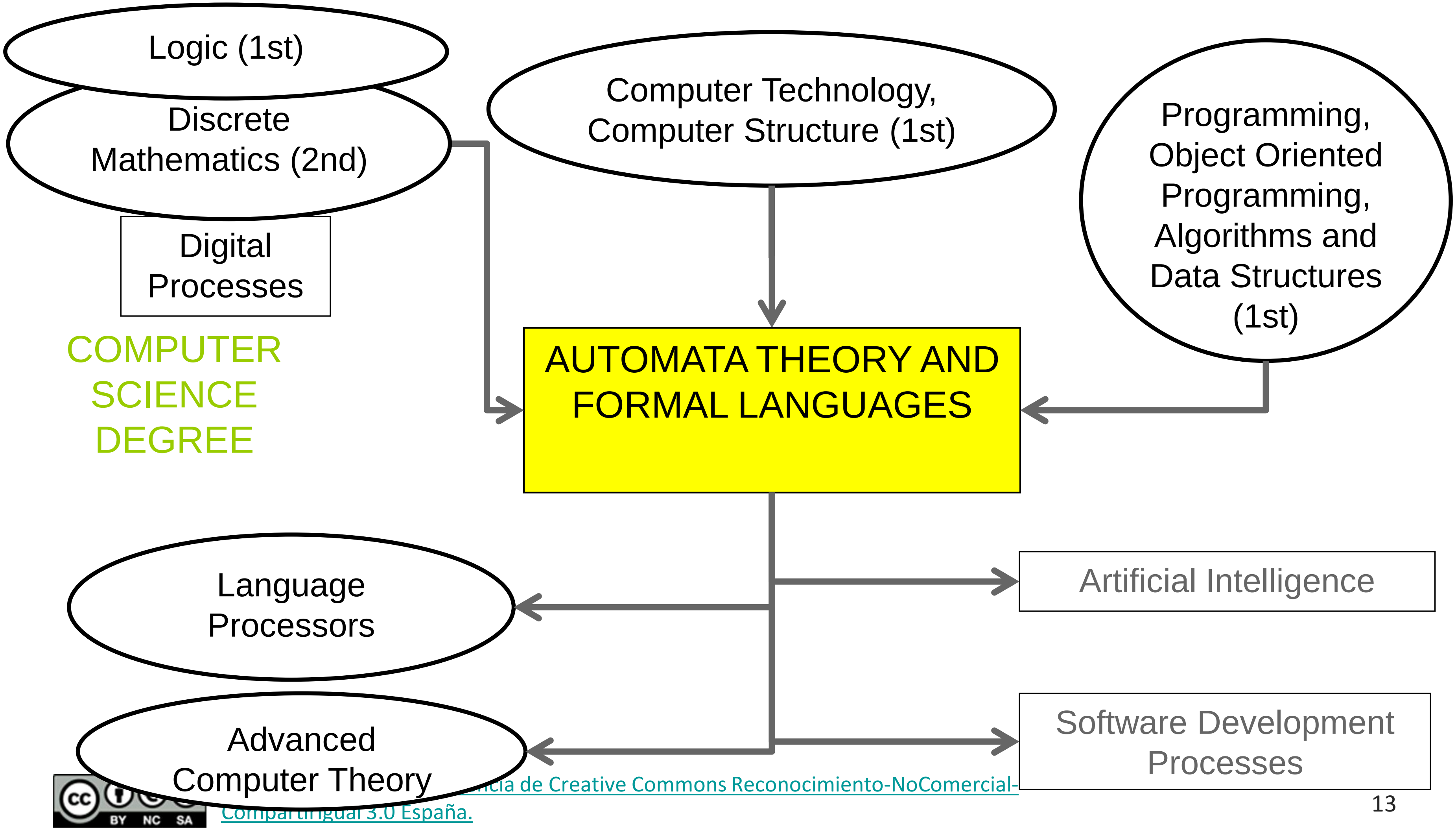
Neural Networks,

Complexity Study

....

Theoretical Computer Science

Relationship with other knowledge areas



Logic (1st)

Discrete Mathematics (2nd)

Digital Processes

COMPUTER SCIENCE DEGREE

Computer Technology, Computer Structure (1st)

Programming, Object Oriented Programming, Algorithms and Data Structures (1st)

AUTOMATA THEORY AND FORMAL LANGUAGES

Language Processors

Advanced Computer Theory

Artificial Intelligence

Software Development Processes

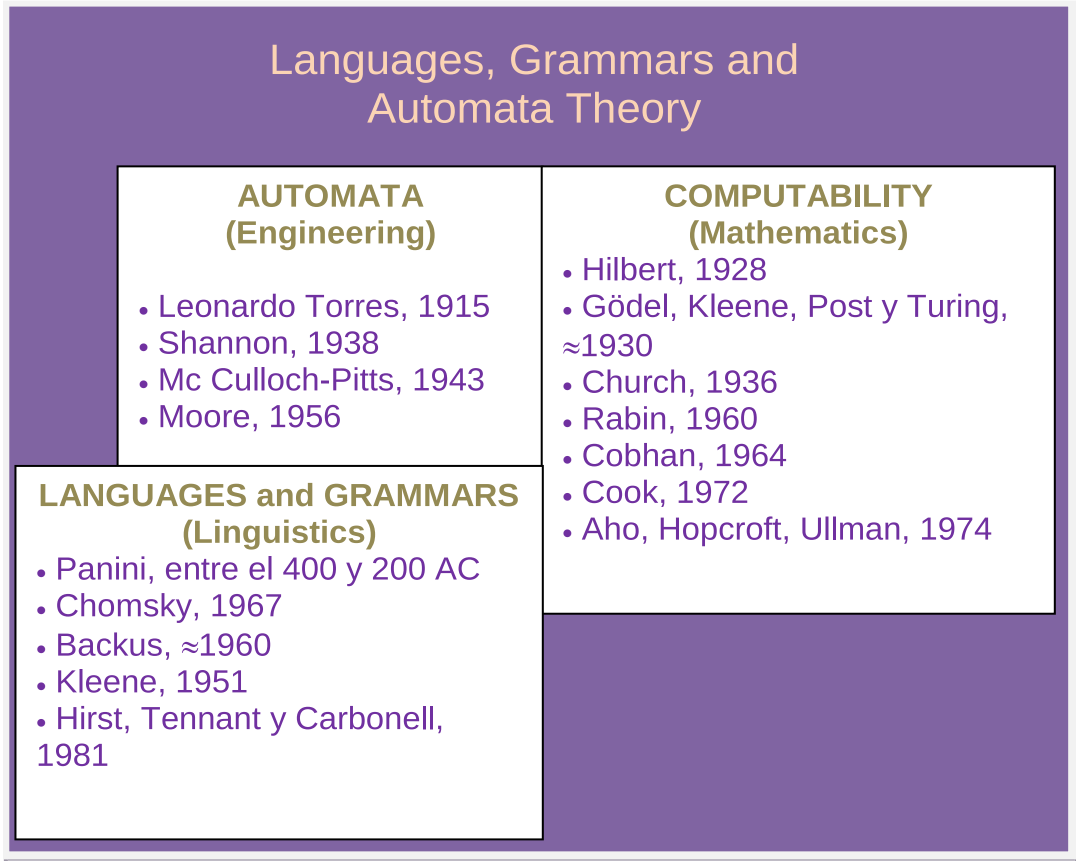
Relationship with other knowledge areas

- Computer Science:
  - Automata Theory and Formal Languages
    - Formal models (i.e. mathematicsbased on computation.
  - Computability Theory
    - Turing Machines, Algorithm. Church-Turing Tesis.
  - Complexity Theory.
- Basis of Theorethical Computer Science:
  - Computability.
  - Automata.
  - Languages and Formal Grammars.

OUTLINE

- Why Automata Theory? History and Origin
  - *Disciplines of Computing*
  - *What "Computer Science" is?*
  - *Can any problem be solved using a computer?*
  - *Which are the capabilities and limitations of a computer?*
- Relationship with other knowledge areas
- **Machines, languages and algorithms**

Basis of Theoretical Computer Science



Languages, Grammars and Automata Theory

AUTOMATA (Engineering)

- Leonardo Torres, 1915
- Shannon, 1938
- Mc Culloch-Pitts, 1943
- Moore, 1956

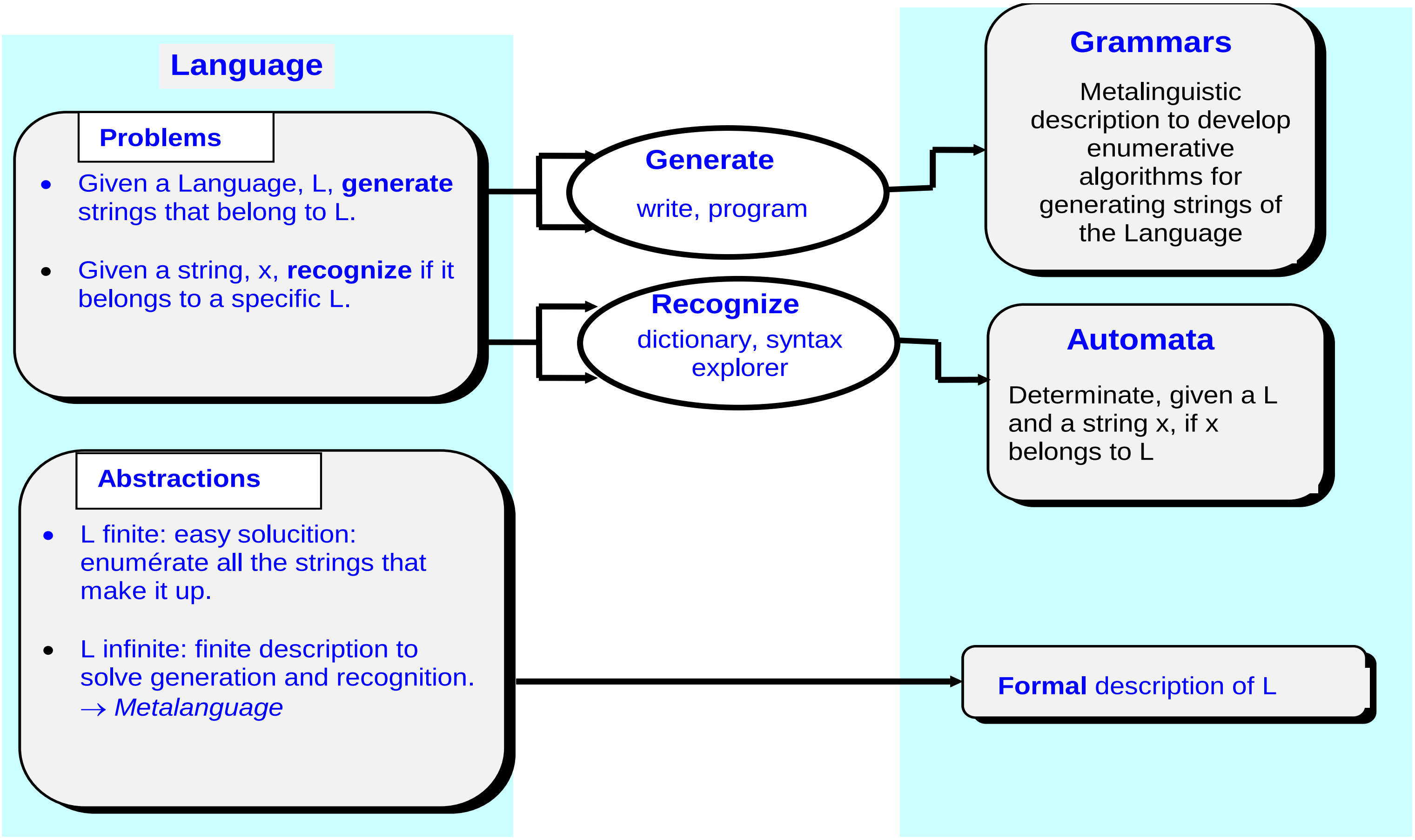
COMPUTABILITY (Mathematics)

- Hilbert, 1928
- Gödel, Kleene, Post y Turing, ≈1930
- Church, 1936
- Rabin, 1960
- Cobhan, 1964
- Cook, 1972
- Aho, Hopcroft, Ullman, 1974

LANGUAGES and GRAMMARS (Linguistics)

- Panini, entre el 400 y 200 AC
- Chomsky, 1967
- Backus, ≈1960
- Kleene, 1951
- Hirst, Tennant y Carbonell, 1981

Machines, languages and algorithms



Language

Problems

- Given a Language, L, **generate** strings that belong to L.
- Given a string, x, **recognize** if it belongs to a specific L.

Abstractions

- L finite: easy solucition: enumérate all the strings that make it up.
- L infinite: finite description to solve generation and recognition. → *Metalanguage*

Generate

write, program

Recognize

dictionary, syntax explorer

Grammars

Metalinguistic description to develop enumerative algorithms for generating strings of the Language

Automata

Determinate, given a L and a string x, if x belongs to L

Formal description of L

Computability

- Defined even before the existence and development of computers.
- What are the problems that can be resolved through automatic procedures?
  - Concept of computable function: function whose values can be automatically calculated using a well-established succession of steps (an algorithm) based on theoretical models (computing).
- Computability theory: answer the following questions:
  - what can computers do (without any restrictions)?
  - What are the inherent limitations in automatic calculation methods?

Automata

- Origins:
  - Create gadgets capable of carrying out daily and common tasks.
  - Machines to perform repetitive tasks that could match with a complex system.
  - Entertain their owners, by performing repetitive movements or emitting sounds.

- Arabs: Construction of automata using accurate calculations (mechanical watches, Astrology).
- Greek engineers contributed large knowledge, though their interest towards human knowledge more than practical applications.
- Fixed programs, that does not necessarily employed the notion of feedback.

Automata

- Mythology has many stories about automata, some a wild and fanciful, others may have been based on fact..
    - Prometheus was reputed to have made the first man and women on earth, with clay animated by fire and stolen from heaven.
    - Hephaestus God of all mechanical arts, also known as Vulcan, God of life was reputed to have made two female statues of pure gold which assisted and accompanied him wherever he went.
    - Hephaestus was also accredited with making Talus, a giant made from brass, which guarded Crete against intruders. Talus dispensed of his victims by heating up his body and hugging them to death. The Argonauts could only land on Crete after Talus was destroyed.
    - Pygmation King of Cyprus fell in love with a beautiful statue which he had made and marries her after Aphrodite brings her to life.

Automata

- 1500 B.C. Amenhotep constructed a statue of Memon, the King of Ethiopia, which emited sounds when it was illuminated by the rays of Sun at dawn.
- 500 B.C., King Tse in China invented a flying magpie made with wood and bamboo.
- 400 - 397 B.C., Taranto Archytar built a wooden young suspended a pivot which rotated with a supplier of water or steam.
- 300 - 270 B.C., Cresibio invented a water clock and a human body that moves with water.
- 220 - 200 B.C., Filon of Byzantium invented an aquatic automaton and a repetitive catapult.
- 206 B.C., a treasure of the first Emperor Chin Shih Hueng is found containing an Orchestra with mechanical dolls.

Automata

- 62 A.C., Hero of Alexandria wrote a famous book relating applications that can be solved using an automata: automatic theatre, animated birds, automatic trumpets, machines that work by means of coins.
- 335 A.C., Hsieh Fec built a vehicle made of wood with four wheels and a Buda figure.
- 700 A.C., Huang Kun built ships with animal, singers, dancers and musicians figures that were able to move.
- 770 aC., Yang Wu-Lien built a donkey that can extend his arms and ask for money, keeping it in a bag.
- 840 aC. Kaya prince, son of Kannu Emperor, built a doll able to move using water.
- 890 aC. , Han Chih Ho built a cat made of wood that can move and dance.

Automata

- Princep Bhoja wrote a book called el Samarangana-Sutradhara that includes comments about developing machine automata.
- Albertus Magnus (1204 - 1272) created a mechanical servant.
- Roger Bacon (1214 - 1294) built (after 7 years), a talking head.
- 1235: Villard d'Honnecourt wrote a book containing illustrations about developing human and animal automata.
- Clock in Strasbourg Cathedral , 1352 -1789.
- 1500: Leonardo Da Vinci built an automatic lion
- Salomón de Caus (1576-1626) developed ornamental fountains and pleasurable gardens, imitating the effects of nature.
- 1640: René Descartes invented an automata with a human figure.

Automata

- Jacques Vaucanson built his famous the duck, the best-known automaton; a duck made of copper, that drinks, eats, emit sounds, swims, and digest the food as a real duck.

- The Maillardet (end of the 18th century and beginning of the 19th) built a writer-painter that writes in English and French and draws landscapes. They also built a "magic" mechanism that answers questions and a bird that sings in a box.

- Robert Houdini made a doll that writes, a pastry chef, an acrobat, a man pointing a shotgun and a the trapeze artist.

- Thomas Alva Edison built in 1891a doll that speaks.

Automata

- Modern concept of automaton: **Leonardo Torres Quevedo**, in 1915:

  "the old automata imitated the aspect and movements of living beings, but the great interest in practice, and what we are looking for this type of devices is to try to achieve the results that gets a person instead of only imitating visible gestures, for thus replace a man with a machine..."

Automata

- Automata theory: application in diverse fields that handle concepts as:
  - Control,
  - Action,
  - Memory
  - Objects are controlled with symbols, words or phrases.
- Related to:
  - Communication Theory,
  - Logic of Sequential Circuits,
  - Computer networks and Coders,
  - Evolutionary and Reproductive Systems,
  - Pattern recognition,
  - Physiology of the nervous system structure,
  - Analysis of programming languages,
  - Automatic language translation,
  - Algebra Theory.

Automata

- In 1938, Shannon defined the foundations for the application of mathematical logic to Combinatorial and Sequential circuits (relays to solid-state electronic devices.
- In the following decades, Shannon ideas were considerably extended: formalization of a theory for sequential machines and finite automata.
- McCulloch-Pitts described in 1943 the artificial neuron: device designed to simulate a biological neuron activity logical calculations (finite automaton nervous activity model).
- The first rigorous automata survey was published by Moore in 1956.
- Prior to the development of the first computers, various methods for the synthesis of sequential circuits (Huffman 1954, Mealy 1955) had been studied.

Automata

- At the end of the 50s: interest about using automata in relation to languages,
- Most of the work on finite automata theory was conducted during the 1960s.
- Other models of automata have been developed in parallel:
  - 40 and 50s von Neumann formulated (among other best-known ideas) the general theory of automata, involving the cellular automata as systems capable of self-reproductive models.
  - Holland was based on these models to propose the first distributed architecture of computers in 1959.
  - In the 1970s and 80s became very popular two-dimensional cellular automata squares simulations, like J.H. Conway Game of Life.

Automata

- The concept of stochastic automaton (probabilistic) appeared in USA at the beginning of the 60s with the aim of improving the reliability of sequential circuits rather than its application in learning systems.
- Von Neumann had introduced the notion of probability in the formal McCulloch-Pitts neuron model to demonstrate theoretically that, thanks to redundancy, it is possible to synthesize reliable systems using unreliable components.
- However, none of these models considered the possibility of self-adapting structures.

Languages and Grammars

- It has its origins in a field away from the Informatics: Linguistics.
- Linguistics traditionally distinguish between particular grammar (concrete properties of languages such as frequency words, syntax rules, etc.) and universal grammar (general properties that can be applied to any human language), as was described by Chomsky in 1967.
- The American Structuralism School developed in the 50s some informal ideas about the universal grammar. For example, if a (natural) language is an innumerable set of phrases → description by means of a generative grammar or set of rules that underlie the composition of correct sentences and structural description for each phrase and explain how to compose such a phrase from the grammar.

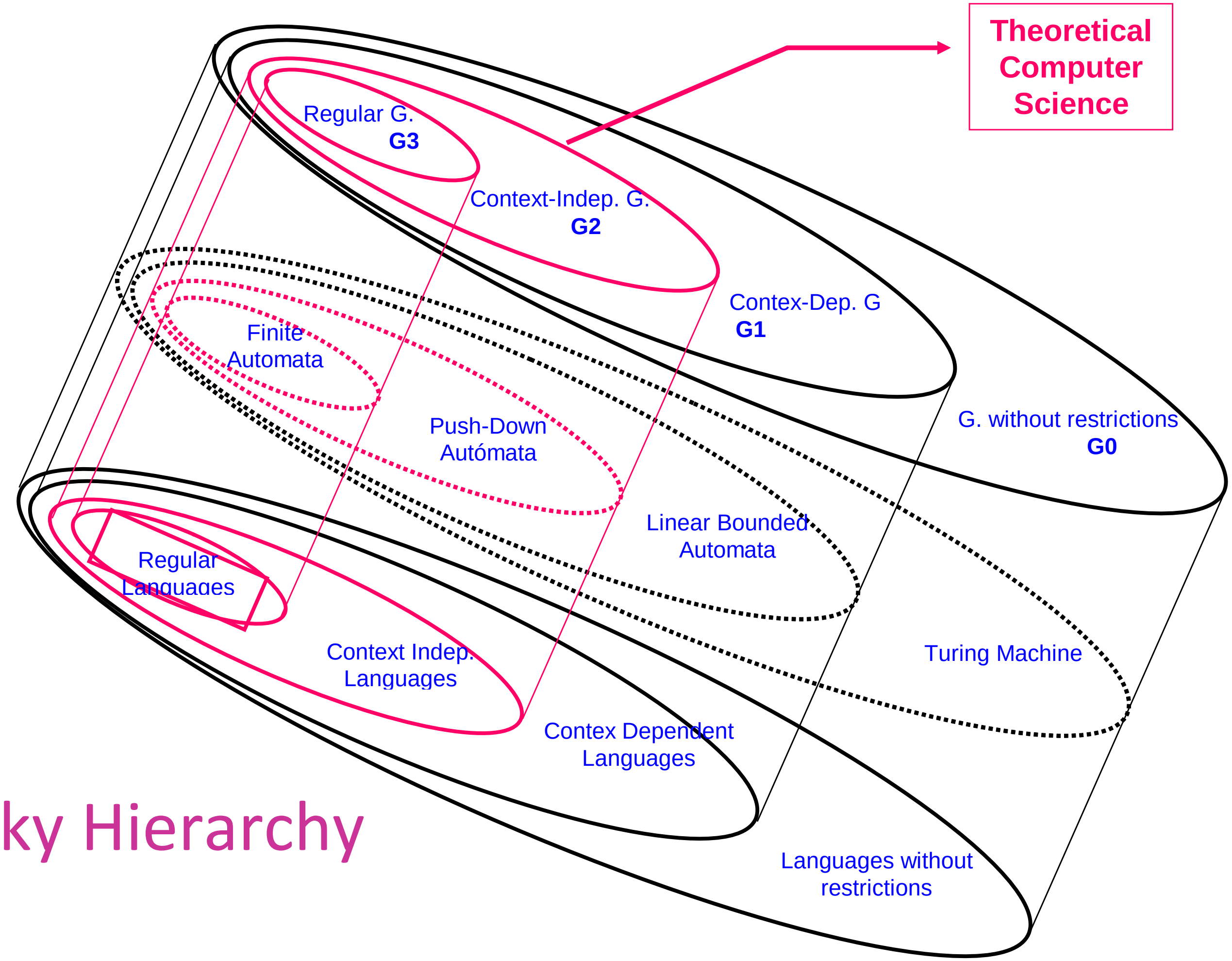
Languages and Grammars

- These concepts formalization was Chomsky work in 1956 (modern Linguistics): develop the mathematical basis for the theory of origin and nature of natural languages.
- For example, generative grammars explains the *creative character* of natural languages, i.e. the fact that there are recursive mechanisms that allow them to express a potentially infinite number of ideas, feelings, etc.
- The lack of a formalism to study these mechanisms had previously force certain behaviorist school linguistics to deny such property, and others, as Saussure, to think that it was something alien to the field of Linguistics.

Languages and Grammars

- The use of grammars to specify the syntax of programming languages emerged independently.
- While working in a draft of ALGOL60, John Backus adapted productions of Post. The resulting notation was a variant of independent context grammars.
- Linguist Panini designed an equivalent Syntactic notation to specify the grammar of Sanskrit between 400 BC and 200 BC
- A breakthrough in the description of regular languages were regular expressions proposed by Kleene in 1951 from McCulloch-Pitts work on artificial neuron. These neurons use a notation for strings input and output that were the basis for the development of regular expressions.
- The work of Kleene shows the equivalence between what he calls *two ways to define a same thing*, which are regular sets, i.e., regular expressions and sets specified by a finite automaton.

Machines, languages and algorithms



Chomsky Hierarchy