A bit of epistemology...

Scientific Methodology

Jean-Marc.Vincent@imag.fr
with a strong support from Arnaud Legrand and Nadine Mandran

Université Grenoble-Alpes and INRIA
Laboratoire d'Informatique de Grenoble
Équipe INRIA POLARIS

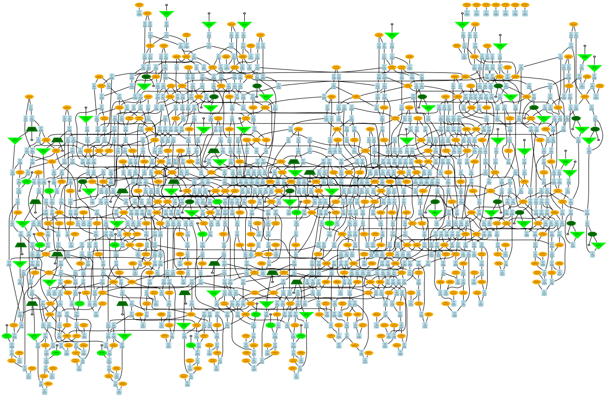
Master of Computer Science of Grenoble
December 2019

METHODOLOGY FOR SCIENTIFIC RESEARCH

UFR IM²AG
UNIVERSITÉ Grenoble Alpes

The Scheduling Problem : Parallel Computing

MOAIS/MESCAL Research Domain: HPC



Molecular Dynamics Simulation: Data/Communications Graph.

The Scheduling Problem : Parallel Computing

The Workload Generation Problem

MOAIS/MESCAL Research Domain: HPC

Part of the Grid5000 Experimental TestBed.

A Scheduling algorithm

Comparison of List Schedulers

1. Build a priority list of all tasks.
2. At each step of the scheduling:
   1. Find an available resource.
   2. Assign the highest priority task to it.

List Scheduling algorithms differ from each other on the strategy used to build their priority list.

A Scheduling algorithm

Strategies simulated

BottomLevel	longest path to sink
Sons	out-degree
Reverse	− out-degree
Random	uniform random choice

Experimental Design	
Sample Size:	1,000
Number of nodes:	100
Task Exec. Time:	1
Number of Proc.:	Varying

More than 1,500,000 simulations for this experiment.

Validation Methodology

Various Needs for Synthetic Workloads

Validation

- Unit Testing
- Random Testing

Performance Evaluation

- Structural Studies
- Quantitative Studies
- Expected Workloads

Validation Methodology

Workload Characterization

Uniform Generation of Random Graphs

Combinatorial Approach.

Specific Classes of Random Graphs

Graphs respecting a set of well known properties.

Traces / Collected Workloads

Identified instances from real/academic environments.

Validation Methodology

The Workload Generation Problem

Workload Characterization

Uniform Generation of Random Graphs

Even the count is unknown above 12 nodes.

Specific Classes of Random Graphs

Graphs respecting a set of well known properties.

Traces / Collected Workloads

Identified instances from real/academic environments.

Validation Methodology

Workload Characterization

Uniform Generation of Random Graphs

Even the count is unknown above 12 nodes.

Specific Classes of Random Graphs

Graphs respecting a set of well known properties.

Traces / Collected Workloads

Hard to generalize results.

Validation Methodology

Workload Characterization

Uniform Generation of Random Graphs

Even the count is unknown above 12 nodes.

Specific Classes of Random Graphs

Our focus.

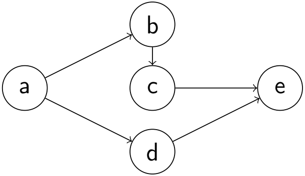
Traces / Collected Workloads

Hard to generalize results.

Validation Methodology

The Workload Generation Problem

Motivation: Simulation of Scheduling Algorithms



Input Characteristics: Directed Acyclic Graph

- Vertices are tasks to execute.
- Edges are precedence constraints or communications.
- Additional annotations for costs.

Results

Makespan Analysis

Average Completion Time on 4 Processors

	Sons		BottomLevel		Reverse		Random	
	avg.	sd.	avg.	sd.	avg.	sd.	avg.	sd.
GNP(100,0.25)	36.1	2.9	35.2	3.1	36.9	2.8	36.5	2.9
GNM(100,300)	25.1	0.4	25.0	0.0	27.2	0.9	26.3	0.8
FiFo(100,10,10)	27.9	1.4	27.9	1.4	29.4	1.6	28.6	1.5
Layer(100,10,0.5)	25.9	0.5	25.8	0.4	27.3	0.7	26.3	0.6
RandomOrders(100,2)	25.4	0.5	25.4	0.5	29.1	1.2	27.1	0.9
GNP(100,0.25)	35.0	3.2	35.0	3.2	35.0	3.2	35.0	3.2
GNM(100,300)	12.4	1.7	12.3	1.7	12.5	1.7	12.4	1.7
FiFo(100,10,10)	11.8	2.3	11.8	2.3	13.3	2.1	12.6	2.3
Layer(100,10,0.5)	10.2	0.4	10.2	0.4	10.2	0.4	10.2	0.4
RandomOrders(100,2)	16.7	1.7	16.7	1.7	16.7	1.7	16.7	1.7

Average Completion Time on 16 Processors

PROBLEM : CRITICAL PATH DENSITY

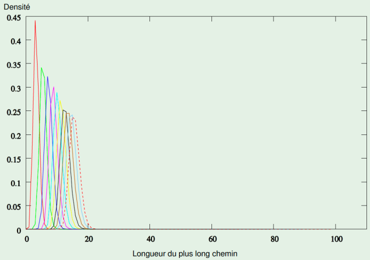


Fig.: Longueur du chemin critique pour Random Orders, $n \in \{10, 20, \dots, 100\}$ et $k = 2$.

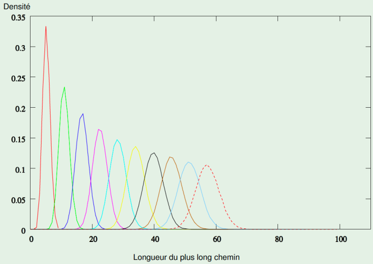


Fig.: Longueur du chemin critique pour $G(n, 0, 5)$, $n \in \{10, 20, \dots, 100\}$.

Moreover the average lenght of the critical path of a uniformly generated partial order is ~ 3 if n is large

METHODOLOGY FOR SCIENTIFIC RESEARCH

RESEARCH...

Brainstorming n°1

Who is science if she was an animal ?

UFR IM²AG
UNIVERSITÉ Grenoble Alpes

RESEARCH...

Brainstorming n°1

Who is science if she was an animal ?

Question n°1

In less than 5 lines give a definition of "Science".

About science...

Définition "Le Robert" (wikipedia)

Ce que l'on sait pour l'avoir appris, ce que l'on tient pour vrai au sens large. L'ensemble de connaissances, d'études d'une valeur universelle, caractérisées par un objet (domaine) et une méthode déterminés, et fondées sur des relations objectives vérifiables [sens restreint]

Définition Trésor de la Langue Française Informatisé

II. Ensemble structuré de connaissances qui se rapportent à des faits obéissant à des lois objectives (ou considérés comme tels) et dont la mise au point exige systématisation et méthode.

Dictionary of science and technology

science noun 1. the study of the physical and natural world and phenomena, especially by using systematic observation and experiment
2. a particular area of study or knowledge of the physical world
3. a systematically organized body of knowledge about a particular subject

New Oxford Dictionary

the intellectual and practical activity encompassing the systematic study of the structure and behavior of the physical and natural world through observation and experiment : the world of science and technology.
1.a particular area of this : veterinary science | the agricultural sciences.
2. a systematically organized body of knowledge on a particular subject : the science of criminology.
3. archaic knowledge of any kind.

Research in Computer Science

Brainstorming n°2

Give 5 examples of scientific facts (in computer science)

Give 5 examples of non scientific facts

RESEARCH IN COMPUTER SCIENCE

Brainstorming n°2

Give 5 examples of scientific facts (in computer science)

Give 5 examples of non scientific facts

Question n°2

In less than 5 lines give the definition of a scientific fact.

Scientific Fact

A scientific fact is an **hypothesis** that have been confirmed by a **specific** experience.

ABOUT SCIENCE...

Question n°3

Step 1

Write in less than 5 lines a scientific hypothesis and propose an experiment to validate it

About Science...

Question n°3

Step 1

Write in less than 5 lines a scientific hypothesis and propose an experiment to validate it

Step 2

Switch your hypothesis with your neighbor.
Propose an experiment to invalidate the hypothesis.

METHODOLOGY FOR SCIENTIFIC RESEARCH

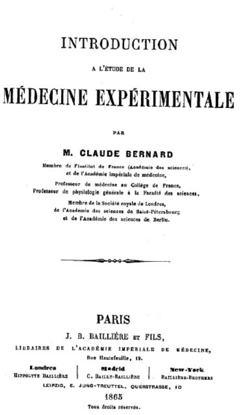
CLAUDE BERNARD 1813-1878

3 steps of the scientific method :

1. observation of the reality is possible without premises ;
2. formulation of an hypothesis (theory) from scientist creativity ;
3. experimental verification by confrontation of the hypothesis with the reality (which is always true).

Inductivism (reasoning from the particular case to the general situation) : "The best theory is the one check by the more numbers of facts."

CLAUDE BERNARD 1813-1878



INTRODUCTION

A L'ÉTUDE DE LA

MÉDECINE EXPÉRIMENTALE

PAR

M. CLAUDE BERNARD

Membre de l'Institut de France (Académie des sciences),
et de l'Académie impériale de médecine,

Professeur de médecine au Collége de France,
Professeur de physiologie générale à la Faculté des sciences.

Membre de la Société royale de Londres,
de l'Académie des sciences de Saint-Pétersbourg
et de l'Académie des sciences de Berlin.

PARIS

J. B. BAILLIÈRE ET FILS,

LIBRAIRES DE L'ACADÉMIE IMPÉRIALE DE MÉDECINE,

Rue Hautefeuille, 19.

Londres — HIPPOLYTE BAILLIÈRE | Madrid — C. BAILLY-BAILLIÈRE | New-York — BAILLIÈRE-BROTHERS

LEIPZIG, E. JUNG-TREUTTEL, QUERSTRASSE, 10

1865

Tous droits réservés.

Wikipedia

Electronic French version

KARL POPPER 1902-1994

- Criteria to separate science and non-science :
  Is scientific a theory that could be falsifiable, that could be submitted by empirical falsification = refutable by facts
- Asymmetry between verification and falsification. It is an epistemology logical and normative. Theories should be clearly formulated, as precisely as possible, can't be suppressed without a 'good reason' (falsification, or theory with "superior degree of falsifiability"), can't be immunized.
- The non-ended play of science.
  - World 1 : the world of physical objects and events, including biological entities
  - World 2 : the world of mental objects and events
  - World 3 : objective knowledge.

Karl Popper's text

In the following text K.R Popper try to answer the questions : "*When a theory get a scientific status ?*" "*Does it exist a criteria to assert the nature or the scientific status of a theory ?*"

1. It is easy to obtain confirmations, or verifications, for nearly every theory-if we look for confirmations.
2. Confirmations should count only if they are the result of risky predictions ; that is to say, if, unenlightened by the theory in question, we should have expected an event which was incompatible with the theory–an event which would have refuted the theory.
3. Every 'good' scientific theory is a prohibition : it forbids certain things to happen. The more a theory forbids, the better it is.
4. A theory which is not refutable by any conceivable event is nonscientific. Irrefutability is not a virtue of a theory (as people often think) but a vice.
5. Every genuine test of a theory is an attempt to falsify it, or to refute it. Testability is falsifiability ; but there are degrees of testability : some theories are more testable, more exposed to refutation, than others ; they take, as it were, greater risks.

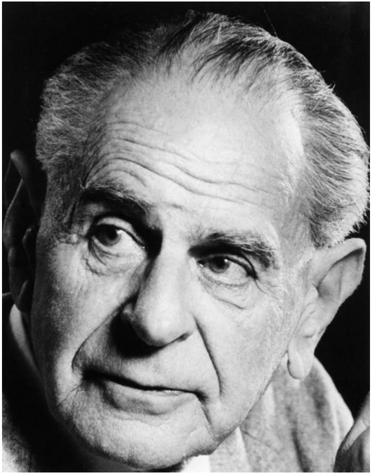
UFR IM²AG
Université Grenoble Alpes

6. Confirming evidence should not count except when it is the result of a genuine test of the theory ; and this means that it can be presented as a serious but unsuccessful attempt to falsify the theory. (I now speak in such cases of 'corroborating evidence'.)
7. Some genuinely testable theories, when found to be false, are still upheld by their admirers–for example by introducing ad hoc some auxiliary assumption, or by re-interpreting the theory ad hoc in such a way that it escapes refutation. Such a procedure is always possible, but it rescues the theory from refutation only at the price of destroying, or at least lowering, its scientific status. (I later described such a rescuing operation as a 'conventionalist twist' or a 'conventionalist stratagem'.)

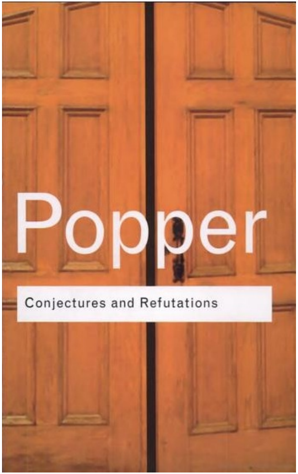
One can sum up all this by saying that the criterion of the scientific status of a theory is its falsifiability, or refutability, or testability.
K.R. Popper, Conjectures and refutations.

Thanks to C. Grasland

Karl Popper 1902-1994



Wikipedia



Electronic version

UFR IM²AG
Université Grenoble Alpes

Science : a dynamical process

- **Phase 1** - It exists only once and is the pre-paradigm phase, in which there is no consensus on any particular theory, though the research being carried out can be considered scientific in nature. This phase is characterized by several incompatible and incomplete theories. If the actors in the pre-paradigm community eventually gravitate to one of these conceptual frameworks and ultimately to a widespread consensus on the appropriate choice of methods, terminology and on the kinds of experiment that are likely to contribute to increased insights.
- **Phase 2** - Normal Science, begins, in which puzzles are solved within the context of the dominant paradigm. As long as there is consensus within the discipline, normal science continues. Over time, progress in normal science may reveal anomalies, facts that are difficult to explain within the context of the existing paradigm. While usually these anomalies are resolved, in some cases they may accumulate to the point where normal science becomes difficult and where weaknesses in the old paradigm are revealed.
- **Phase 3** - This phase is a crisis. Crises are often resolved within the context of normal science. However, after significant efforts of normal science within a paradigm fail, science may enter the next phase.
- **Phase 4** - Scientific revolution is the phase in which the underlying assumptions of the field are reexamined and a new paradigm is established.
- **Phase 5** - Post-Revolution, the new paradigm's dominance is established and so scientists return to normal science, solving puzzles within the new paradigm.[4]

A science may go through these cycles repeatedly, though Kuhn notes that it is a good thing for science that such shifts do not occur often or easily. (source wikipedia)

Thomas Kuhn 1922-1996

Pre-science

1. debate on the bases ;
2. no selection of facts ;
3. no scientific domain.

Normal Science

existence of a paradigm, a matrix for the domain :

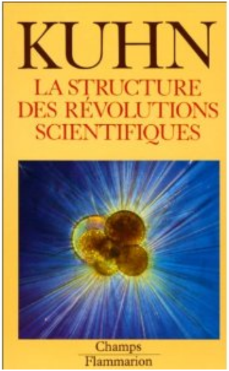
1. common language for the "tribe"
2. shared believes (ontologies, metaphors and analogies) ;
3. shared values (values, methodological, epistemological norms) ;
4. socialization examples. Normal science check that the paradigm is right, never contradictory facts

Scientific revolution

strong anomalies : fundamental questioning of the basic paradigms at a psychological, sociological, political level. Incommensurability between the old and the new paradigm.

Thomas Kuhn 1922-1996

Paradigms

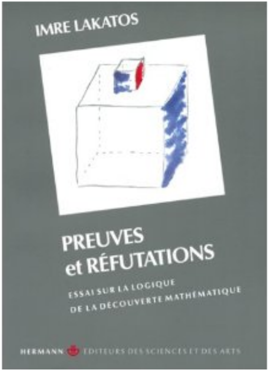


Electronic version

Imre Lakatos 1922-1974

- **Research Program : hard core** values et fundamental believes fondamentales ontologic and methodologic (ideology of the group), never questioned ("negative heuristic").
- **Protecting belt** : theories confirming observed facts and protecting the hard core from critics. We falsify at the protecting belt level, never at the hard core level. We evaluate series of theories rather than falsifying a particular one (as Popper did).
- **Main Science :** characterized by several concurrent research programs.
- **Progressive Program :** progress at the theoretical level (increase coherence) and at the empirical level (new facts).
- **Degenerated Program** : no progress at the theoretical (no improvements) / empirical (no unpredictable facts) level

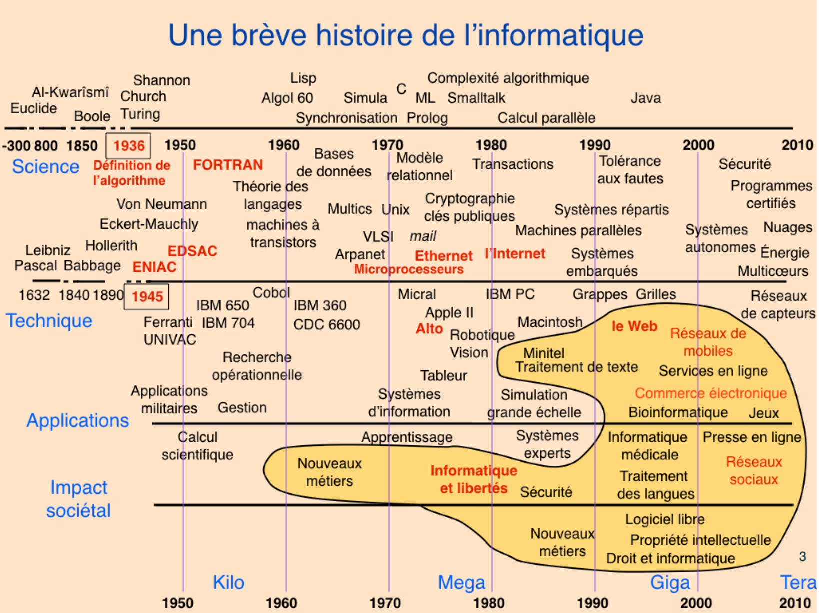
Imre Lakatos 1922-1974



Wikipedia

Electronic version

METHODOLOGY FOR SCIENTIFIC RESEARCH



(source S. Krakowiak 2016)

Computer Science

Computing

- A **science**
  Science of artificial... but not only
- A **technology**, an industry
  Hardware, software, network, services,...
- **Applications**
  increasing area
- **Social impact**
  Numerical/Information society

Computer Science

Computing

- A **science**
  Science of artificial... but not only
- A **technology**, an industry
  Hardware, software, network, services,...
- **Applications**
  increasing area
- **Social impact**
  Numerical/Information society

Computer Science

- **Concepts : representation of the object**
  Information
  Representation, communication, compression,...
  Algorithm
  Operative process
  Programming **Language**
  link between levels of abstraction
  Architecture (Computing Engine)
  abstraction of the physical world
  Human in the loop
- **Methods**
  Back and forth between theory and experimentation
  Automatic abstraction transform
  Self-generated tools
- **Human organization**

METHODOLOGY FOR SCIENTIFIC RESEARCH

UFR IM²AG
UNIVERSITÉ Grenoble Alpes

SYNTHESIS

Scientific Method

Falsifiability is the logical possibility that an assertion can be shown false by an observation or a physical experiment. [Popper 1930]

SYNTHESIS

Scientific Method

Falsifiability is the logical possibility that an assertion can be shown false by an observation or a physical experiment. [Popper 1930]

Modeling (abstract representation) comes before experimenting

SYNTHESIS

Scientific Method

Falsifiability is the logical possibility that an assertion can be shown false by an observation or a physical experiment. [Popper 1930]

Modeling (abstract representation) comes before experimenting

Modeling principles [J-Y LB]

- (Occam :) if two models explain some observations equally well, the simplest one is preferable
- (Dijkstra :) It is when you cannot remove a single piece that your design is complete.
- (Common Sense :) Use the adequate level of sophistication.

Synthesis

Scientific Method

Falsifiability is the logical possibility that an assertion can be shown false by an observation or a physical experiment. [Popper 1930]

Modeling (abstract representation) comes before experimenting

Modeling principles [J-Y LB]

- (Occam :) if two models explain some observations equally well, the simplest one is preferable
- (Dijkstra :) It is when you cannot remove a single piece that your design is complete.
- (Common Sense :) Use the adequate level of sophistication.

Science is a Social Phenomena

- collaborative construction of knowledge
- dynamic evolution of knowledge

...now it's your turn ...

UFR IM²AG
Université Grenoble Alpes

To go further

Ethique / Intégrité / Déontologie:

→ *les trois piliers d'une science responsable*

Ethique de la recherche	Intégrité scientifique	Déontologie du fonctionnaire
Les grandes questions que posent les progrès de la science et leurs répercussions sociétales	Les règles qui gouvernent la pratique de la recherche	Le contrôle des liens d'intérêts & cumuls d'activité des fonctionnaires
Dimension culturelle: doit se discuter en permanence puis s'impose	Dimension universelle: s'impose comme un code professionnel de « droit souple »	Loi Le Pors 1983 rév. 2016: *"Le fonctionnaire exerce ses fonctions avec dignité, impartialité, intégrité et probité"*
→ *Tous les chercheurs*	→ *Tous les chercheurs*	→ *Chercheurs publics*
Des comités	**Des référents chercheurs**	**Des référents juristes**

by Olivier Le Gall Inra Bordeaux

METHODOLOGY FOR SCIENTIFIC RESEARCH

UFR IM²AG
UNIVERSITÉ Grenoble Alpes

Références I

Bernard, C. (1865), *Introduction à l'étude de la médecine expérimentale*, J.-B. Baillère et fils.

Carnap, R. (1995), *An Introduction to the Philosophy of Science*, Dover Publications.

Chalmers, A. F. (1990), *Qu'est-ce que la science ?*, Le Livre de Poche.

Chalmers, A. F. (2013), *What Is This Thing Called Science ?*, Hackett Publishing Company, Inc.

Godfrey-Smith, P. (2003), *Theory and Reality : An Introduction to the Philosophy of Science (Science and Its Conceptual Foundations series)*, University Of Chicago Press.

Kosso, P. (2011), *A Summary of Scientific Method*, Springer.

Kuhn, T. S. (1996), *The Structure of Scientific Revolutions*, The University of Chicago Press.

Kuhn, T.-S. (2008), *La structure des révolutions scientifiques*, Flammarion.

Lakatos, I. (1976), *Proofs and Refutations : The Logic of Mathematical Discovery*, Cambridge University Press.

Lakatos, I. (1984), *Preuves et Réfutations : essai sur la logique de la découverte mathématique*, Hermann.

Popper, K. (2002), *Conjectures and Refutations : The Growth of Scientific Knowledge*, Routledge.