

La scienza come ignoranza degli esperti ed il governo del numero

“Science is the belief in the ignorance of experts”

(R. Feynman)

Giuseppe Longo

CNRS et Ecole Normale Supérieure, Paris, *and*

Dept. of Immunology, Tufts Univ., Boston

ww.di.ens.fr/users/longo

Computers' Networks: a fantastic opportunity and challenge

Facing an alternative:

- Correlate **diversity** and increase **adaptivity** and **variability**, by the “flexibility” and richness of networks

or

- Construct **uniformity** by a **Global Mean Field**

Computers' Networks: a fantastic opportunity and challenge

Facing an alternative:

- Correlate **diversity** and increase **adaptivity** and **variability**, by the “flexibility” and richness of networks

or

- Construct **uniformity** by a **Global Mean Field**

In Physics:

In **Mean Field Theory**, spin moves in the average field produced by all other spins.

In a four dimensional field, **no more singularities**

Next: *1 - Information without Knowledge (pure Data)*

2 - “Governance” without Government

*or, the wrong use of networks %
3*

Knowledge Construction

In Science, thinking differently, the novelty ...

and *critical thinking*:

Awareness of the **limits** of knowledge construction, always an **interpretation**, a perspective (a friction on, canalized by “reality”)

Knowledge Construction

In Science, thinking differently, the novelty

and *critical thinking*:

Awareness of the **limits** of knowledge construction, always an **interpretation**, a perspective (a friction on, canalized by “reality”)

The relevance of **Negative Results** (Pythagoras, Riemann, Poincaré ...)

“Science is the belief in the ignorance of experts”

Enabling the debate and change in perspective in the *interpretation* of the “real” in Science

and in the *government* of Society vs *governance* %

Knowledge and Government vs Information and Governance

Governance vs Government, beyond interpretation:

The **automatisms** of governance (by numbers) vs
government as **interpretation** in
law making, governing, judging

Knowledge and Government vs Information and Governance

Governance vs Government, beyond interpretation:

The **automatisms** of governance (by numbers) vs
government as **interpretation** in
law making, governing, judging

Follow the rule (equilibrium economics, a geodetics ...

Merkel-Sarkozy: “automatic”)

“Objective”, with no responsibility

(A. Supiot, “La gouvernance par le nombre”)

Economic Governance: follow the geodetics,
“There is no alternative”

Walras' equilibrium equations ...

Economic Governance: follow the geodetics,
“There is no alternative”

Walras' equilibrium equations ... Poincaré's letter, 1901:

- “Economic satisfaction, an arbitrary function ... relative individuals ... to eliminate in computations”
- “In celestial mechanics, one may neglect the friction of planets, while your economic individuals as perfectly egoistic with perfect knowledge ... “quelques reserves” ”

A paradigmatic **scientism** still governing us (2008 ... Fokker-Planck)

From the *rule* and the *optimal* paths to **Governance by Objective Data:**

Content: towards Big Data

- Big Data, reality and myth
- Ramsey Theory, the limits
- Motivations from Cancer Biology (Soto Lab, Boston)
- Hints to alternative paths in Biology (theories and interpretations)
- Bibliometrics (a few words)

Big Data and their Potentialities

IBM [18 : "What is Big Data ?"] estimates that

« Every day, we create **2.5 quintillion bytes of data** – so much that 90% of the data in the world today has been created in the last two years alone. »

Fantastic tool for knowledge and science!

Big Data and their Potentialities

IBM [18 : "What is Big Data ?"] estimates that

« Every day, we create **2.5 quintillion bytes of data** – so much that 90% of the data in the world today has been created in the last two years alone. »

Fantastic tool for knowledge and science!

After **Greek observation and theorizing:**

Experimental method (Galileo),
Mathematics for Physics (Newton)

Big Data and their Potentialities

IBM [18 : "What is Big Data ?"] estimates that

« Every day, we create **2.5 quintillion bytes of data** – so much that 90% of the data in the world today has been created in the last two years alone. »

Fantastic tool for knowledge and science!

After Greek **observation** and **theorizing**:

Experimental method (Galileo),

Mathematics for Physics (Newton)

Immense Databases (if soundly used ...)

E.g. *statistics* or extensive *use* of the immense data bases on, e.g.

Biological Rhythms (cardiac, metabolic ...) in

Longo G., Montévil M., **Perspectives on Organisms: Biological Time, Symmetries and Singularities**, *Springer*, Berlin, 2014.

Data as a Result of a Theoretical Decision

Numbers are not already in physical/biological/social processes

They are the result of a **choice of observables** (1) and of a difficult procedure of **measuring** (2)

(theory, reference syst., dimension, metrics, tools for measurement)

These are all an **interpretation** of reality: *a compressed theory*

Science begins when (1) and (2) are made **explicit** (and discussed)

Data as a Result of a Theoretical Decision

Numbers are not already in physical/biological/social processes

They are the result of a **choice** of **observables** (1) and of a difficult procedure of **measuring** (2)

(theory, reference syst., dimension, metrics, tools for measurement)

These are all an **interpretation** of reality: *a compressed theory*

Science begins when (1) and (2) are made **explicit** (and discussed)

Scientism, some properties:

- data as **absolutes**
- projecting the **latest machine** on the world (brain, DNA ...)
- make intelligible and *govern* the world by **optimization methods**

Association “G. Cardano” <http://cardano.visions-des-sciences.eu>

“**Science** against **Scientism**”

Big Data analysis as « The End of Science »

Big Data analysis as « The End of Science »

C. Anderson, 2008: « The End of Theory: The Data Deluge Makes the Scientific Method Obsolete »

« Correlation is enough We can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms **find patterns** where science cannot ».

« with enough data, the numbers speak for themselves ... **Correlation supersedes causation**, and science can advance even without coherent models, unified theories. »

The largest the best ... *Independently* of any analysis of the “meaning” or “content” (no *interpretation*), prediction and **rules for action are provided by the data mining** (*NSF project*).

References are in:

Calude C., Longo G. The Deluge of Spurious Correlations in Big Data, 2016
(<http://www.di.ens.fr/users/longo/download.html>).

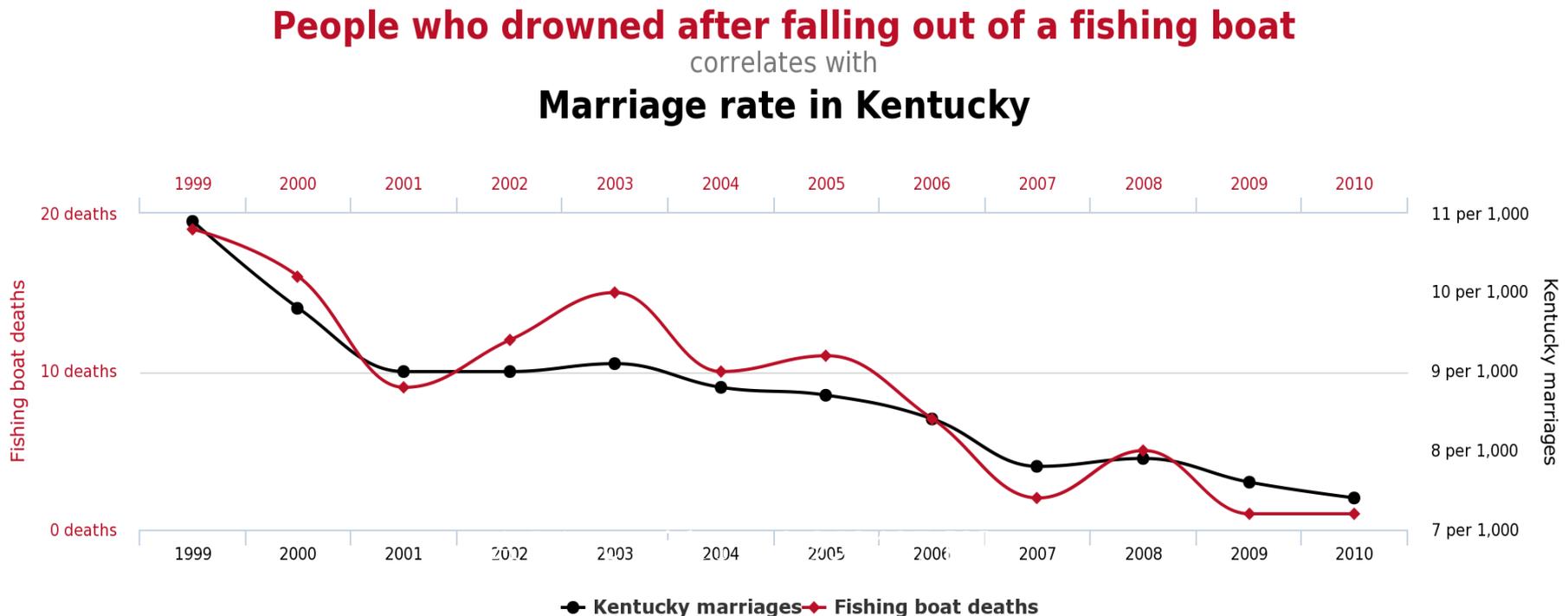
Just looking at Big Data

M. Hayden (former CIA dir): **"We kill people based on metadata"**
(2014 : <https://www.youtube.com/watch?v=UdQiz0Vavmc>)

Just looking at Big Data

M. Hayden (former CIA dir): "**We kill people based on metadata**"
(2014 : <https://www.youtube.com/watch?v=UdQiz0Vavmc>)

An **empirical** response: A large collection of spurious correlations :
<http://www.tylervigen.com/spurious-correlations>, 2015.



Use Mathematics to fight the Big Data Folies

PART II: By some use of “Ramsey theory” (born in the 1920's), prove:

Informal: "Given **any arbitrary correlation** on sets of data, **there exists** a large enough number (size) such that **any data set of that size or more**, realises that type of correlation."

PART III: Since this large enough **data set is arbitrary**, it may have been obtained by a **random** generator of digits or numbers (series of dice throws or quantum spins measurements).

Note: it is exactly the **size of the data** that allows our result: the more data, the more arbitrary, meaningless and useless (for future action) correlations will be found in them. *How large?*

Calude C., Longo G. The Deluge of Spurious Correlations in Big Data, 2016
(<http://www.di.ens.fr/users/lengo/download.html>)

**"Colored" Van der Waerden and
Ramsey Theorems
(Finite Combinatorics)**

"Colored" Van der Waerden

Finite Van der Waerden theorem

(for sequences of digits or colors):

For all integers c and k there is an integer γ such that all strings, made out of c digits or colors, of length more than γ contain an arithmetic progression with k occurrences of the same digit or color, i.e. a monochromatic arithmetic progression of length k .

"Colored" Ramsey Theorems

Finite Ramsey theorem

(for n -ary relations $[A]^n$ or n -subsets of a set A):

For all integers s, n, c there is an integer γ such that for every finite set A containing more than γ elements and for every partition $P : [A]^n \rightarrow \{1, 2, \dots, c\}$ there exists a subset B of A containing s elements whose n -sets are monochromatic, i.e. $P(x)$ has the same value (color) for every x in $[B]^n$.

Later: How large is γ ?

Hints to applications of Ramsey Theorems

Let A be a relational database. Fix s, n, c ,
a **correlation of variables** in A
is a set B of size s (e.g. the number of years and quantities)
whose n -ary relations $(a_1, a_2 \dots a_n)$
form the correlation (for the *given criteria* or colors c)

When the **correlation applies**, all elements are given **the same color**,
out of c (are *monochromatic*).

Hints to applications of Ramsey Theorems

Let A be a relational database. Fix s, n, c ,
a **correlation of variables** in A
is a set B of size s (e.g. the number of years and quantities)
whose n -ary relations $(a_1, a_2 \dots a_n)$
form the correlation (for the *given criteria* or colors c)

When the **correlation applies**, all elements are given **the same color**,
out of c (are *monochromatic*).

Then by **Ramsey theorem** one has that:
given any “correlation”, i.e. any s, n and c , there always exists a large
enough number γ such that any set A of size greater than γ , in *any*
way “P” it colored, contains a set B of size s whose subsets of n
elements (n -ary relations) are all correlated – that is, monochromatic.

Since A is arbitrary, it may be generated by a **random** process ...
Intuition: Gaz, Clouds ...

Ramsey: How large is γ ?

Let $c = 2$ and $\gamma = R(s,n)$ the Ramsey number of s, n and 2
(i.e. given s and n , any set A of cardinality $R(s,n)$ contains a
subset B of cardinality s , with $[B]n$ monochromatic)

Immensely large if $\text{card}(B) > \min(B)$ [Paris-Harrington, 1978; Longo, 1981]

Upper and lower bounds have been computed for $R(s,s)$: these are
exponentials compatible with today's size of Big Data
[Erdos, Szkeres, 1947; Szemeredy, 1980; Conlon, 2009]

For $n = s$: A (corrected) exponential upperbound:

$$R(s, s) \leq [1 + o(1)] \frac{4^{s-1}}{\sqrt{\pi s}}.$$

An exponential lower bound,

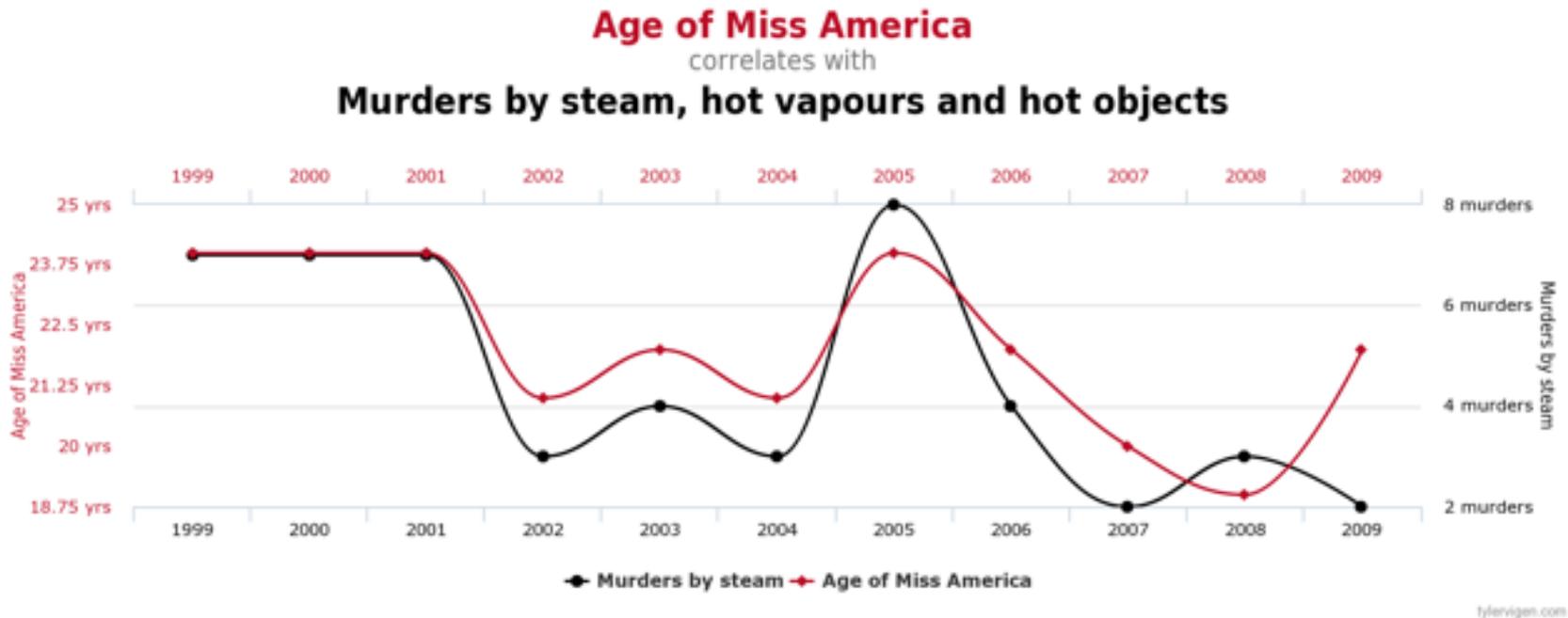
$$R(s, s) \geq [1 + o(1)] \frac{s}{\sqrt{2e}} 2^{s/2},$$

What is a "spurious" correlation ?

Theory dependent definition ...

What is a "spurious" correlation ?

Theory dependent definition ...



Potatoes ...

How many "spurious" correlations ?

"Spurious" a relative notion

Spurious? It depends on the available theories ...

Define: a correlation is "**spurious**" when it belongs to a set A *produced by a random process*

A very strong (restrictive) definition

But then, what "**random**" means, for sets of numbers ?

Randomness for sequences of numbers ...%

Randomness for sequences of numbers

Algorithmic Information Theory

(Kolmogorof, 1960; Martin-Löf, 1965; Chaitin, 1970; Calude, 2002):

Martin-Löf's **randomness** for *infinite sequences of numbers*

Corroborated by *asymptotic* correlations to Physical Randomness:

PhD Theses: M. Hoyrup, C. Rojas (2008), A. Abbott (2015, with C.)

Randomness for sequences of numbers

Algorithmic Information Theory

(Kolmogorof, 1960; Martin-Löf, 1965; Chaitin, 1970; Calude, 2002):

Martin-Löf's **randomness** for *infinite sequences of numbers*

Corroborated by *asymptotic* correlations to Physical Randomness:

PhD Theses: M. Hoyrup, C. Rojas (2008), A. Abbott (2015, with C.)

Definition (Kolmogorof): A **finite** sequence is **incompressible** if there is no program *shorter than the sequence* that generates it.

Easy **extension to n-ary relations**, i.e. valid for both VdW and Ramsey frames:

- any finite set is computationally isomorphic to a sequence with a "low cost" of coding

How many "random" A ?

Random finite sequence, an approximation: **incompressible**

- In **practice** (compression algorithms) : The best average rate of algorithmic compressibility is of about 86.4%
- The **probability that a binary string x** of (relatively short) length 2048 is reduced by 13.6%
is smaller than 10^{-82}
(10^{82} is the number of hydrogen atoms in the Universe)

*In other words, for large n, very few strings of length n are compressible, that is **not algorithmic random***

*Or **most large sets of numbers are algorithmically random***

Possible Objections to our Approach

- 1 - In Ramsey Theory, the size of γ is huge ... - *not so large*
- 2 - Very Large Databases detect **average** in Biology

Possible Objections to our Approach

1 - In Ramsey Theory, the size of γ is huge ... - *not so large*

2 - Very Large Databases detect **average** in Biology - *very bad*:

A heritage from *Statistical Physics* (averaging out, Central Limit Theor... Avogadro) in “**noise biology**”, see [Bravi, Longo, 2015]

Unsuitable in Biology (and historical sciences) ... e.g. its origin:

From **DNA** changes to “hopeful monsters” extremely **rare** event
scan the time of evolution [Buiatti, Longo, 2013], [Longo, 2017]

- Possible bridge: **Large Deviation Theory** [Vulpiani et al, 2014]
- **Mesoscopic** Level [Giuliani, 2013]

Variability, diversity ...

SUMMARY ON BIG DATA

The larger the set of data:

the larger is

The Deluge of Spurious Correlations in Big Data

Yet, the better for sound statistical analyses ... %

Big Data for Statistical Analyses (S. Huang, A. Giuliani)

Current Statistical Analysis are **hypothesis** (thus, **theory**) driven:

- *research hypotheses* (with alternatives),
- *null hypothesis* (no sense relationship between two data sets)
- give *probability's thresholds*

Big Data for Statistical Analyses (S. Huang, A. Giuliani)

Current Statistical Analysis are **hypothesis** (thus, **theory**) driven:

- *research hypotheses* (with alternatives),
- *null hypothesis* (no sense relationship between two data sets)
- give *probability's thresholds*

Typically:

- A comparison is **statistically significant** if the relationship between the data sets would be an unlikely realization of the *null hypothesis* given a **threshold probability**—the significance level.
- The process of distinguishing between the *null hypothesis* and the **alternative hypothesis** is aided by identifying two conceptual types of errors (type I & type II), and by specifying parametric limits on e.g. how much type 1 error will be permitted.

Note : **type I** error is the incorrect rejection of a true null hypothesis (a "**false positive**"), while a **type II** error is the failure to reject a false null hypothesis (a "**false negative**").

Big Data, a fantastic challenge for science

Greek observations and speculation further enriched by

Experimental method (Galileo),
Mathematics (Descartes, Newton),
Big Data (if soundly used ...)

Big Data, a fantastic challenge for science

Greek observations and speculation further enriched by

Experimental method (Galileo),
Mathematics (Descartes, Newton),
Big Data (if soundly used ...)

The relevance of **Negative results**:

Presence of randomness in **Big Data** *is correlated to*

Concrete **Unprovability** (1978 - ...):

1 - Ramsey-Paris-Harrington (add fix point $\text{Card}(A) > \min(A)$;
proof by ϵ_0 or **geometric well-ordering**)

2 - Kruskal-Friedman (γ_0 or geometric well-ordering)
a form of **Feynman's awareness of ignorance**

No way to **predict, act** and **govern** by (un-interpreted) Big Data

Knowledge and Government: debate, interpretation and choice ...

A major primary confusion

Intelligence is “**elaboration of information**”

A major primary confusion

Intelligence is “**elaboration of information**”

No, intelligence is *also*

“**imagination of configuration of sense**”

Elaboration of information *vs* imagination of configuration of sense

Interpolating **stars**: projecting meaning ...

Elaboration of information vs imagination of configuration of sense

Interpolating **stars**: projecting meaning ...

Drawing a **border**:



Lascaux, -18,000 years: just borders

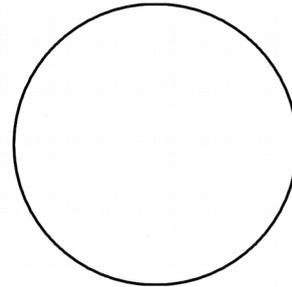
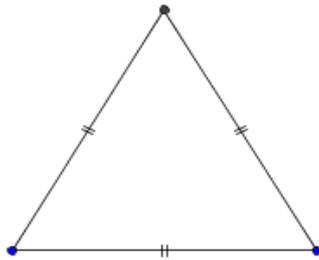
The origin of geometry

The origin of geometry

Euclid's definition β :

“A *line* is a length with *no thickness*”

Just a contours: all Euclid's figures are "**just borders**”:

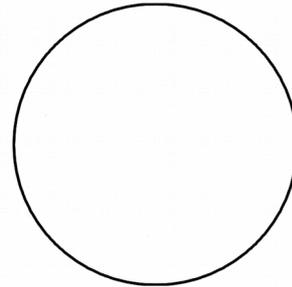
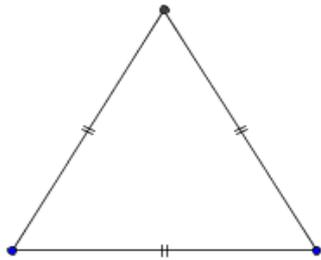


The origin of geometry

Euclid's definition β :

“A *line* is a length with *no thickness*”

Just a contours: all Euclid's figures are "**just borders**”:



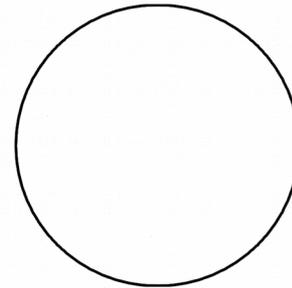
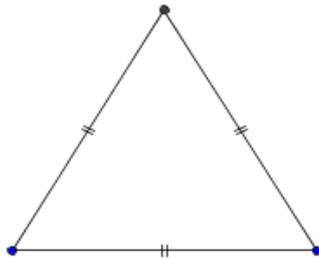
In Mathematics, today: **imagine** a “sheaf on a site”

The origin of geometry

Euclid's definition β :

“A *line* is a length with *no thickness*”

Just a contours: all Euclid's figures are "**just borders**":



Foundations of Mathematics:

Beyond the “annex of a Philosophy of Language”
Towards a component of a Philosophy of Nature

Today's challenges in Biology

Motivations for Big Data Analysis: The case of Cancer

Collaboration, since 2008 with

C. Sonnenschein, A. Soto

Department of Integrative Physiology and Pathobiology

Tufts University School of Medicine, Boston

Cancer Biology: Tissue Organization Field Theory (**TOFT**)

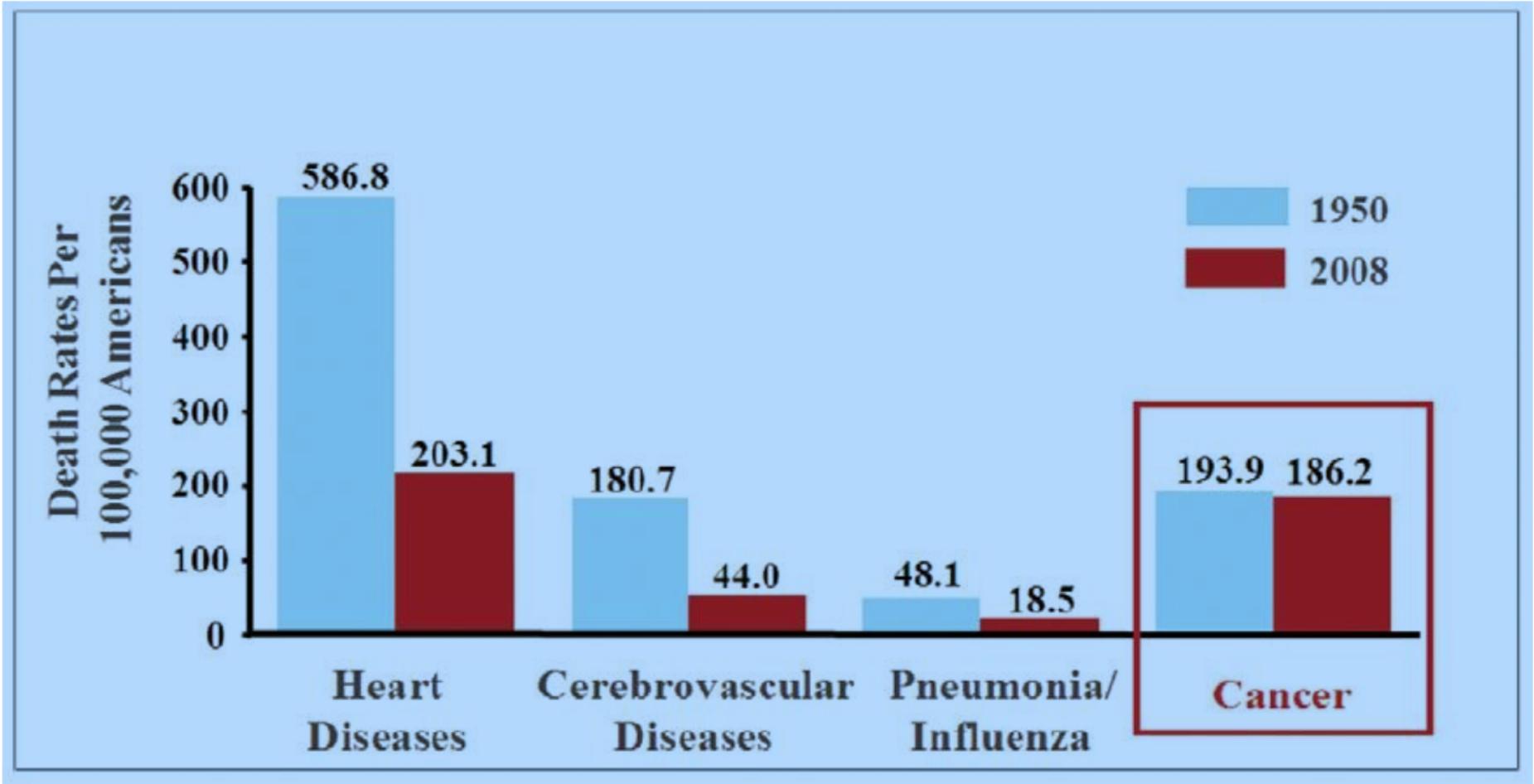
M. Montévil,

Former PhD student, then joint post-doc U. Boston and ENS, Paris (+4 people)

Soto A., Longo G. (eds.) **From the century of the genome to the century of the organism: New theoretical approaches**, *Special issue of Progress in Biophysics and Molecular Biology*, Vol. 122, Issue 1, Elsevier, 2016

G. Longo *Information and Causality: Mathematical Reflections on Cancer Biology*, to appear, 2018

Some Data on Cancer (USA)



While almost doubling the **therapeutical success**,
thus doubling incidence

“Doses” of Chemicals in the Ecosystem

***Unnoticed* Endocrine disruptors:**

82,000 artificial molecules produced in the XX century

(FDA Rep Congress, 2008)

Do not worry: **small doses** *and* **not stereo-specific**

“Doses” of Chemicals in the Ecosystem

***Unnoticed* Endocrine disruptors:**

82,000 artificial molecules produced in the XX century

(FDA Rep Congress, 2008)

Do not worry: **small doses** *and* **not stereo-specific**

The myth of the “genetic program”:

In order to carry **information**, **de-program** the genetic program
molecular interactions ***necessarely stereospecific*** (**key-lock**) ...

[Monod, 1970], [Maynard-Smyth, 1999]

“Doses” of Chemicals in the Ecosystem

Unnoticed Endocrine disruptors:

82,000 artificial molecules produced in the XX century

(FDA Rep Congress, 2008)

Do not worry: **small doses** *and* **not stereo-specific**

The myth of the “genetic program”:

In order to carry **information**, **de-program** the genetic program
molecular interactions *necessarely stereospecific* (**key-lock**) ...

[Monod, 1970], [Maynard-Smyth, 1999]

No:

- **non-linear** effects, **low** chemical affinities
- **varying** association/dissociation constants, **contextual**
- to be given in **probabilities**, depending on the **context** [Elowitz, 2002]

Some Data on Endocrine Disruptors

- **endocrine target organs, cancer general increase (1994 – 2012) :**
breast 26% ; testis 56% ; prostate 94%
thyroid cancer (+285% in 30 years, till 2012)

S. De Coster, N. van Larebeke, Endocrine-disrupting chemicals, **J. Environ. Public Health** 2012.

N. Howlander, et al, SEER Cancer Statistics Review, 1975–2012,
National Cancer Institute.

- **The case of asbestos (Maltoni '70s; Huang, 2011)**

Some Data on Endocrine Disruptors

- **endocrine target organs**, cancer general increase (1994 – 2012) :
breast 26% ; *testis* 56% ; *prostate* 94%
thyroid cancer (+285% in 30 years, till 2012)

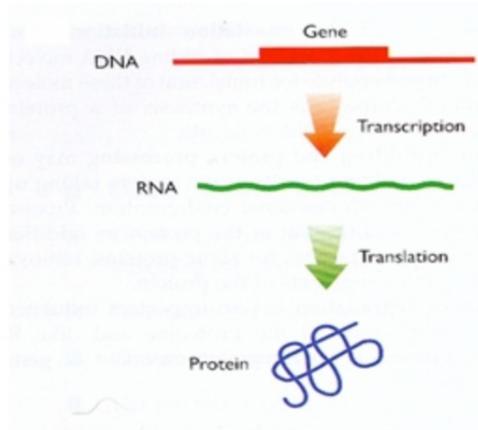
S. De Coster, N. van Larebeke, Endocrine-disrupting chemicals, **J. Environ. Public Health** 2012.

N. Howlander, et al, SEER Cancer Statistics Review, 1975–2012,
National Cancer Institute.

- **The case of asbestos** (Maltoni '70s; Huang, 2011)
- **Lowering by 50% (!) of human spermatozoa density** since 1950's
E. Diamanti-Kandarakis et al. *Endocrine-disrupting chemicals: an Endocrine Society scientific statement.* **Endocr Rev** 30:293-342, 2009
N. Skakkebaek, *Sperm counts, testicular cancer, environment,* **BMJ**, 2017
- **GMOs: children of the Central Dogma: programming the plant ...**
(Buiatti, 2000 ...)

The Central Dogma still resisting! The genotype completely “drives” the phenotype

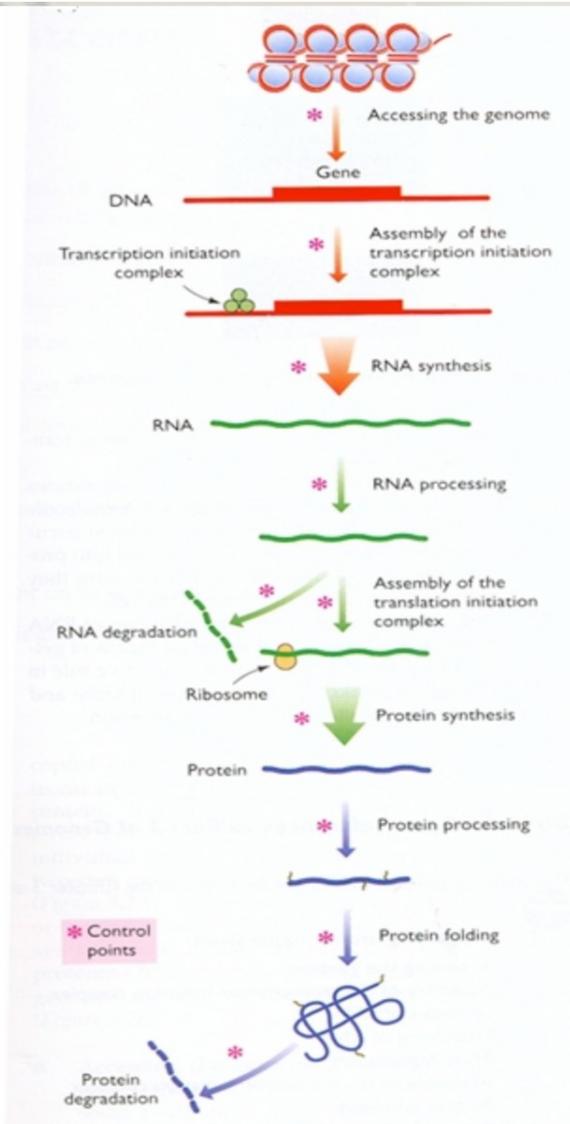
En 1960



(d'après Brown, 2002, in *genomes*)

En 2000

Transcription
Traduction



Cancer: search for the **onco-gene**, **proto-onco-gene**, **onco-suppressor-gene**

The GMO's: a direct consequence of the *Central Dogma*

Remember: the **completeness of the DNA coding** of an organism
« the organism: a mere vehicle ... », « once the DNA fully decoded
... on a CD-rom... this is a man, this is me» (Collins, Gilbert, Guyon,...)

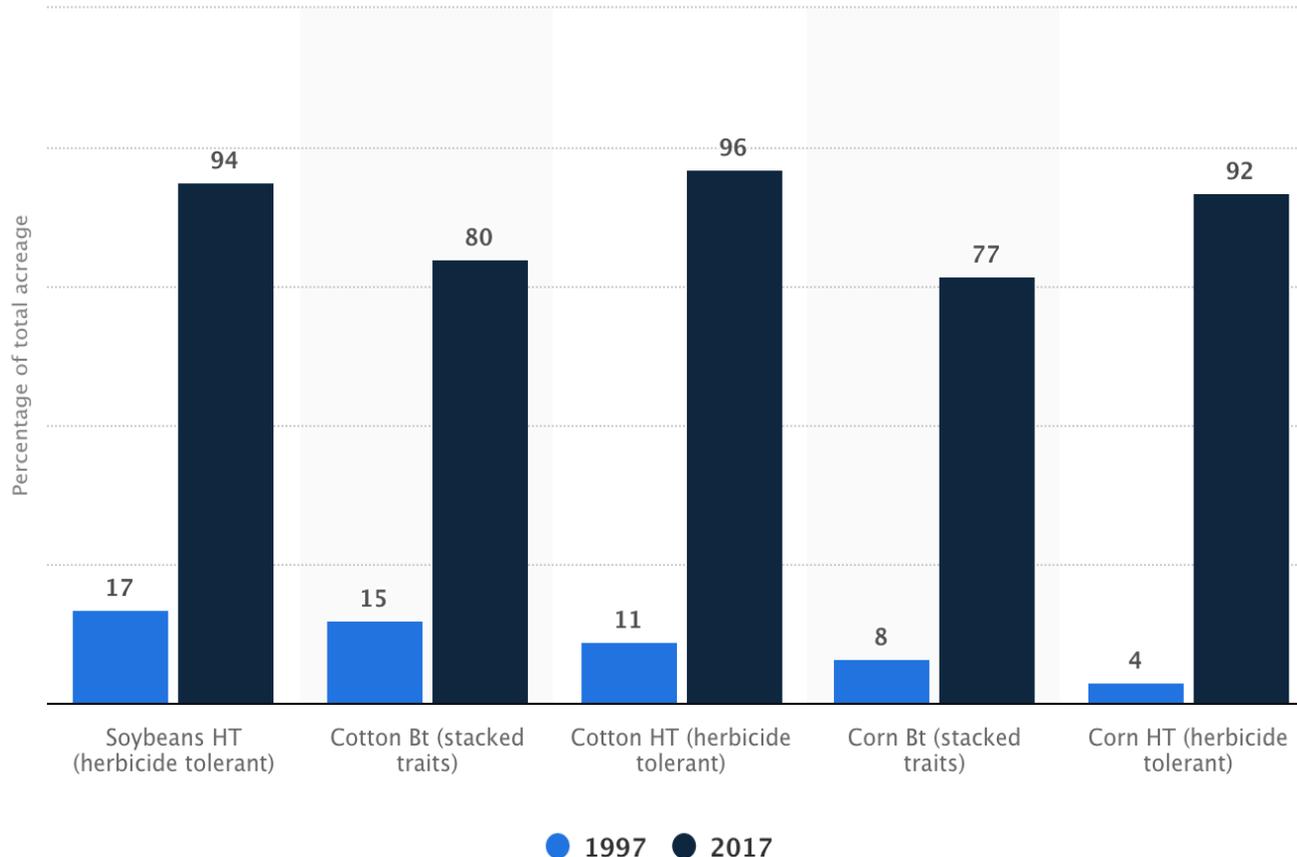
The GMO's: a direct consequence of the *Central Dogma*

Remember: the **completeness of the DNA coding** of an organism
« the organism: a mere vehicle ... », « once the DNA fully decoded
... on a CD-rom... this is a man, this is me» (Collins, Gilbert, Guyon,...)

- Indirect consequences of **pesticides resistance** (absorption, transfer ...)
- *Major modifications of microbial flora and fauna* (fungi, roots, soil)
 - G.A. Kowalchuk et al., 2003. Assessing responses of soil microorganisms to GM plants. **Trends in Ecology and Evolution** 18, 403–410.
 - M. Castaldini, et al, 2005, Impact of Bt Corn on Rhizospheric and Soil Eubacterial Communities and on Beneficial Mycorrhizal Symbiosis in Experimental Microcosms, **Applied and Environmental Microbiology**, 71: 6719-29
 - M. A. Badri et al., 2009, Unintended molecular interactions in transgenic plants expressing clinically useful proteins..., **Proteomics**, 9: 746–756.

The GMO's: a direct consequence of the *Central Dogma*

Percentage of genetically modified crops in the U.S. in 1997 and 2017, percent of total acreage)



DOWNLOAD SETTING



DESCRIPTION SOURCE

This statistic shows the genetically modified in t 2017, by type, as a perc each crop. In 2017, som crops were herbicide to

Codings, codings ... and formal rules

The origin of coding

Codings, codings ... and formal rules

Schrödinger, 1944, *part 1*:

« In calling the structure of the chromosomes a *code-script*, we mean that the all-penetrating mind, **once conceived by Laplace**... could tell from their structure how the egg would develop... . »

Turing 1950: “my DSM is **Laplacian**” (determination implies predictability)

Coding and digital information: the *governance* by exact/integer numbers and rules

Schrödinger’s *right consequences* of his principles!

Today, the code-script has been fully decoded...

Codings, codings ... and formal rules

Schrödinger, 1944, *part 1*:

« In calling the structure of the chromosomes a *code-script*, we mean that the all-penetrating mind, **once conceived by Laplace**... could tell from their structure how the egg would develop... . »

Turing 1950: “my DSM is **Laplacian**” (determination implies predictability)

Coding and digital information: the *governance* by exact/integer numbers and rules

Schrödinger’s *right consequences* of his principles!

Today, the code-script has been fully decoded...

Schrödinger, 1944, *part 2*: morphogenesis and Gibbs free-energy

The Human DNA “decoding”, 2000-01

Robert A. Weinberg,

a major promotor of the *Somatic Mutation Theory* (SMT) of cancer :

Co-author of a "classic" synthesis:

Hanahan D and Weinberg RA. *The hallmarks of cancer*. **Cell**,
100, 57–70, 2000. *(20,000 citations by 2010)*

« ... **cancer biology and treatment ... will become a science** with a conceptual structure and logical coherence that rivals that of chemistry or physics »

The cancer is **clonal**; the **onco-gene** *or* **proto-onco-gene**, *or* **onco-suppressor-gene**, on the ground of the Central Dogma ... %

DNA decoding, 2000-01 (fantastic technological achievement)

F. Collins, 2001: « we have grasped the **code written by God** »

C. Venter, 2001: the "decoder" of the human genome

A. von Eschenbach, director Nat. Cancer Inst. 2003: "to eliminate the suffering and death from **cancer**, and to do so by 2015"

Diagnosis and prognosis within two or three years ... *NO WAY!*

DNA decoding, 2000-01 (fantastic technological achievement)

F. Collins, 2001: « we have grasped the **code written by God** »

C. Venter, 2001: the "decoder" of the human genome

A. von Eschenbach, director Nat. Cancer Inst. 2003: "to eliminate the suffering and death from **cancer**, and to do so by 2015"

Diagnosis and prognosis within two or three years ... *NO WAY!*

C. Venter, interview for the *Spiegel*, July 29, 2010:

Title: « **We have learned nothing from the genome** »

« ... phonies ... the ill-founded belief that those who know the DNA sequence also know every aspect of life. This nonsense ... »

<http://www.spiegel.de/international/world/spiegel-interview-with-craig-venter-we-have-learned-nothing-from-the-genome-a-709174.html>

Yet, we did learn a lot: *the case of cancer* ..

"Coming Full Circle – from endless complexity to simplicity and back again" by R. A. **Weinberg**, MIT Center for Molecular Oncology,
Cell 157, March 27, 2014

Capitulation

« **Half a century** of cancer research had generated an enormous body of observations about the behavior of the disease, but there were **essentially no insights** into how the disease *begins and progresses* to its life-threatening conclusions. »

« ... essentially incoherent phenomena that constituted "cancer research [at the molecular level]" ... **one should never, ever confuse cancer research with science** »

... the story ...

Cancer and the DNA decoding

From the massive DNA decoding of cells in cancer tissues:

- 1 - Gene-expression signatures for **benign** and **malignant** cancer may coexist in the same tumor.
- 2 - DNA sequencing does not help in distinguishing a **primary** from a **metastatic** cancer (80 % of letal cancer).

(Imielinski et al., 2012 ; Gerlinger, 2012 ; ...)

G. Longo. Mathematical Reflections on Cancer Biology, *in print*, 2018.

« most human carcinogens are **not mutagenic** » (!) (Weinberg, 2014)

Cancer and the DNA decoding

From the massive DNA decoding of cells in cancer tissues:

- 1 - Gene-expression signatures for **benign** and **malignant** cancer may coexist in the same tumor.
- 2 - DNA sequencing does not help in distinguishing a **primary** from a **metastatic** cancer (80 % of letal cancer).

(Imielinski et al., 2012 ; Gerlinger, 2012 ; ...)

G. Longo. Mathematical Reflections on Cancer Biology, *in print*, 2018.

« most human carcinogens are **not mutagenic** » (!) (Weinberg, 2014)

See also: R. Gatenby “*Of cancer and cave fish*”, **Nature**, 2011

E. Jablonka, M. Lamb, 2008 ; M. West-Eberhard, 2003

Cells isolated from cancers revert to normalcy when placed in a normal microenvironment (Maffini et al. 2005; Hendrix et al. 2007; Bussard et al. 2010)

Cancer and Big Data

Since « ... one should never, ever confuse cancer research with science » ... « **myriads of unexpected mutations** » (Weinberg, 2014)

Let's then predict and act on the grounds of Dig Data !

Purely **Big Data** Driven *cancer research*: all -omics, *predict and act*:

Cancer Institute, Oregon Health & Science Univ. & Intel, 2016 :

<http://www.informationweek.com/big-data/big-data-analytics/can-big-data-help-cure-cancer-/d/d-id/1326295>

Many Biology University Labs & **IBM**, 2016:

<http://www.businessinsider.in/IBMs-Watson-can-now-do-in-minutes-what-takes-cancer-doctors-weeks/articleshow/47168413.cms>

Cancer and Big Data

Since « ... one should never, ever confuse cancer research with science » ... « **myriads of unexpected mutations** » (Weinberg, 2014)

Let's then predict and act on the grounds of Dig Data !

Purely **Big Data** Driven *cancer research*: all -omics, *predict and act*:

Cancer Institute, Oregon Health & Science Univ. & Intel, 2016 :

<http://www.informationweek.com/big-data/big-data-analytics/can-big-data-help-cure-cancer-/d/d-id/1326295>

Many Biology University Labs & **IBM**, 2016:

<http://www.businessinsider.in/IBMs-Watson-can-now-do-in-minutes-what-takes-cancer-doctors-weeks/articleshow/47168413.cms>

Microsoft 2016 <http://news.microsoft.com/stories/computingcancer/>
“we debug the DNA and *it is a solved problem*” (A bias on research funds)

Alternative theories:

Sonnenschein C., Soto A.M. **The society of cells: cancer and control of cell proliferation.** Springer, 1999.

Tissue Organisation Field Theory (TOFT)

Main (and explicit) *theoretical assumption*:

Darwin's first principle:

Alternative theories:

Sonnenschein C., Soto A.M. **The society of cells: cancer and control of cell proliferation.** Springer, 1999.

Tissue Organisation Field Theory (TOFT)

Main (and explicit) *theoretical assumption*:

Darwin's first principle:

Cells' default state is *reproduction with variation*

massively controlled, in an organism (tissue structure, cells' exchanges, hormones' system ... the ecosystem)

An analysis of ontogenesis (and **DNA's** key role):

Soto A., Longo G. eds., *From the century of the genome to the century of the organism: New theoretical approaches.* **Prog. Biophys. Mol. Biol.**, 122, 2016. 71

Biology: the organismal level

Starting points:

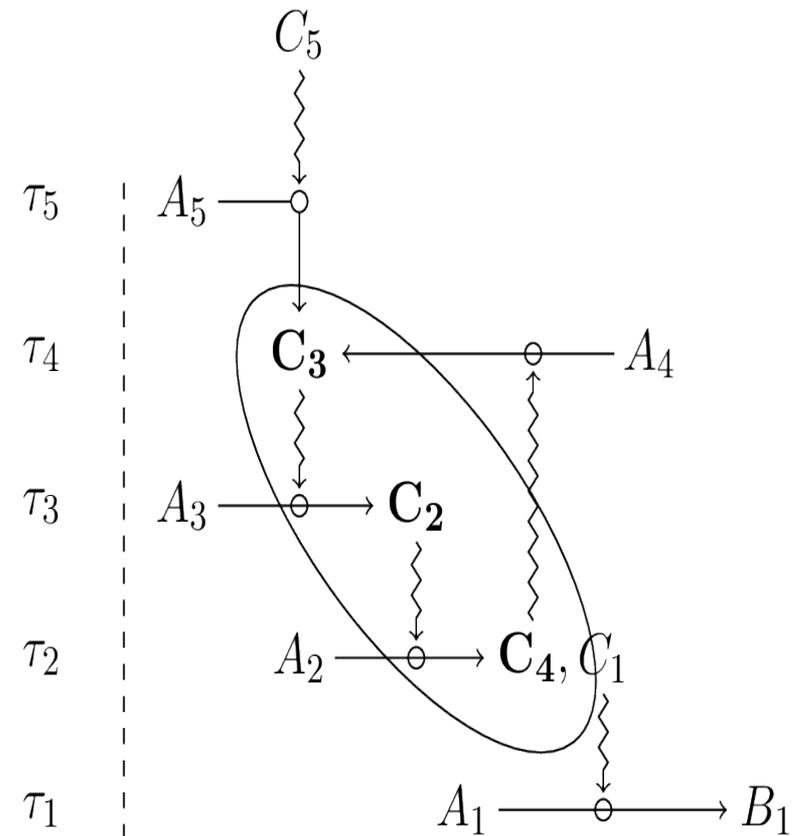
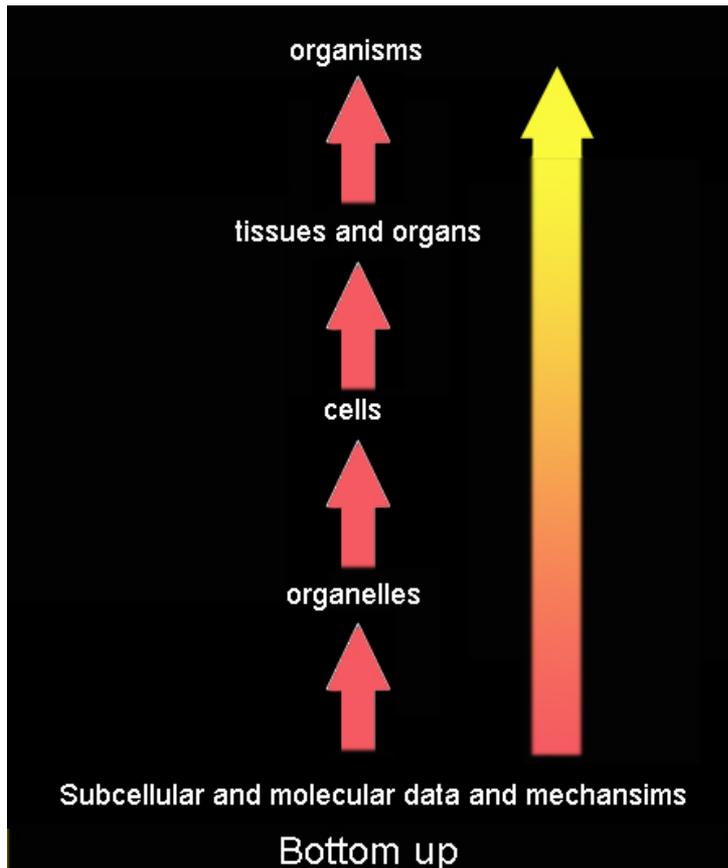
Darwinian principles:

“Reproduction with variation” (*and motility*)

“Selection” (*enablement*)

Soto A., Longo G. (eds.) **From the century of the genome to the century of the organism: New theoretical approaches**, *Special issue of Progress in Biophysics and Molecular Biology*, Vol. 122, Issue 1, Elsevier, 2016.

Move from *Central Dogma* to *Closure of Constraints*



M. Montévil, M. Mossio, 2015

Soto A., Longo G. (eds.) 2016

Back to Science *and* Democracy

Science *and* Democracy

Bibliometrics:

The identification of democracy with the **majority vote** (1), actually the *governance by the “audience”*, disregarding the **division of powers** (2) and the formation of **alternative views** (3)

(2) including “tenure” (*cf.* the Director of *Science*)

(3) and novelty (*cf.* Negative Results)

Science is cannot be “governed” by the “vote of a **global majority**” by “audience” on Earth, the **number** of quotations, in the short term (the *impact factor* concerns two or five years old publications)

Bibliometrics reinforces dominant fashions, kills diversity and critical thinking, discourages the formation of *small communities in emerging domains*

Evaluating Scientific Work

Science, even within a “school”, is the **new path** opened by a minority, which may become an “**occasional**” majority (a jury) voting against a global, uniform opinion

Counting: a gauge that may become a target, is no longer a gauge
Measure “*Quality*” ... a matter of *interpretation* and *ethics*

See:

2008, ENS-INRIA: <http://www.di.ens.fr/users/longo/files/Data/lettre-bibliometrie.pdf>

J.P.A. Ioannidis ‘Why most Published Research Findings are False’ **PLOS Medicine**, 2005

2009, MSCS Editorial: **bibliometrics and the curators of orthodoxy**

G. Longo. *Science, Problem Solving and Bibliometrics*. Invited Lecture, Academia Europaea Conference on "Use and Abuse of Bibliometrics", Stockholm, May 2013. Proceedings, Wim Blockmans et al. (eds), Portland Press, 2014. (*downloadable*)
(Fantoni ed il 10% ...)

ROARS: <https://www.roars.it/online/complexita-scienza-e-democrazia/>

Conclusion

Governance (by the rule, by the numbers...) is not *Government*

Information is not *Knowledge Construction*

Some references (*downloadable*: Google: Giuseppe Longo Paris)

Bravi B., Longo G. **Biology, from Noise to Functional Randomness.** *in* Springer LNCS 9252, pp 3-34, 2015

Calude C., Longo G. **The Deluge of Spurious Correlations in Big Data,** *in* **Foundations of Science**, 1-18, March, 2016

Calude C., Longo G. **Classical, Quantum and Biological Randomness as Relative Unpredictability.** Spec issue, **Natural Computing**, Springer, 2016.

Soto A., Longo G. (eds.) **From the century of the genome to the century of the organism: New theoretical approaches,** *Special issue of Progress in Biophysics and Molecular Biology*, Vol. 122, Issue 1, Elsevier, 2016.

Longo G. **Information and Causality: Mathematical Reflections on Cancer Biology, Organisms, J. Biology,** *to appear*, 2018.

Some references (*downloadable*: Google: Giuseppe Longo Paris)

- Bailly F., Longo G. **Mathematics and the Natural Sciences. The Physical Singularity of Life.** *Imperial College Press*, London, 2011 (français: Hermann, 2006).
- Longo G., Montévil M., **Perspectives on Organisms: Biological Time, Symmetries and Singularities,** *Springer*, Berlin, 2014.

Downloadable from <http://www.di.ens.fr/users/longo/>

Buiatti M., Longo G. *Randomness and Multilevel Interactions in Biology, Theory of Biosciences*, vol. 132, n. 3:139-158, 2013.

Longo G., P. A. Miquel, C. Sonnenschein, A. Soto. *Is Information a proper observable for biological organization? Progress in Biophysics and Molecular Biology*, Vol. 109, Issue 3, pp. 108-114, August 2012.

Longo G., Montévil M., *Models vs. Simulations: a comparison by their Theoretical Symmetries.* Invited paper, *Springer Handbook of Model-Based Science*, *downloadable*.

“The key-lock, hand-glove paradigms in molecular biology ... exact transmission and elaboration of biological information”

(Stanford Encyclopedia, 2016)

Against evidence:

"Even more radically ... proteins never do fold into a particular shape, but rather remain unstructured or “disordered” ... In mammals, about 75% of signaling proteins and half of *all* proteins are thought to contain long, disordered regions, while about 25% of all proteins are predicted to be “fully disordered” ... Many of these intrinsically unstructured proteins are involved in regulatory processes, and are often at the center of large protein interaction networks"

Jörg Gsponer and M. Madan Babu, “The Rules of Disorder Or Why Disorder Rules,” *Progress in Biophysics and Molecular Biology* (2009).

« ... 210 tumors found **no mutations** in 73 tumors ... 183 lung adenoma carcinomas, **only 6% of tumors had mutations** assigned to all six classic **hallmarks** of cancer. » (Imielinski et al., 2012)

« in different regions of the **same tumor** region, 63 to 69% of all somatic mutations **not detectable** ... Gene-expression signatures of **good and poor prognosis** detected» ...

Molecular « **intratumor heterogeneity**, associated with heterogeneous protein function, may foster tumor adaptation and **therapeutic failure** through Darwinian selection » ***Not clonal!***
(Gerlinger et al. *Engl J Med* 366;10, march 8, 2012)

« ... 210 tumors found **no mutations** in 73 tumors ... 183 lung adenoma carcinomas, **only 6% of tumors had mutations** assigned to all six classic **hallmarks** of cancer. » (Imielinski et al., 2012)

« in different regions of the **same tumor** region, 63 to 69% of all somatic mutations **not detectable** ... Gene-expression signatures of **good and poor prognosis** detected» ...

Molecular « **intratumor heterogeneity**, associated with heterogeneous protein function, may foster tumor adaptation and **therapeutic failure** through Darwinian selection » ***Not clonal!***
(Gerlinger et al. *Engl J Med* 366;10, march 8, 2012)

Cells isolated from cancers revert to normalcy when placed in a normal microenvironment

(Maffini et al. 2005; Hendrix et al. 2007; Bussard et al. 2010)

Economy beyond Networks as Global “mean fields”

An alternative use of Networks:
Platforms for **collaborative economy**
“La pleine commune” (IdF)

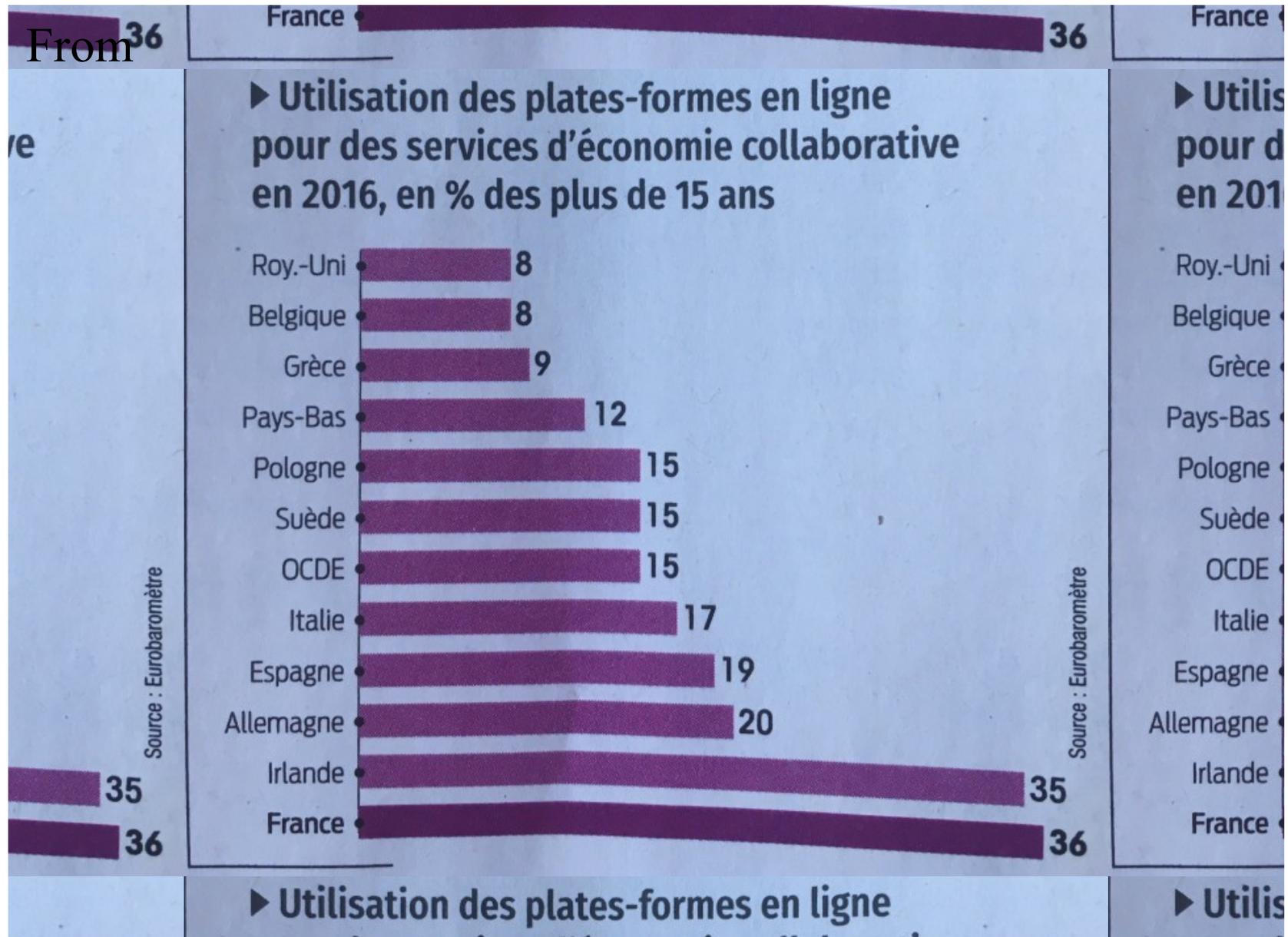
Beyond Networks as Global “mean fields”

Project: La **Pleine Commune** (9 cities north of Paris)

From the **global** macro-economic **governance** to
“local structures of economic (and cultural) coherence”

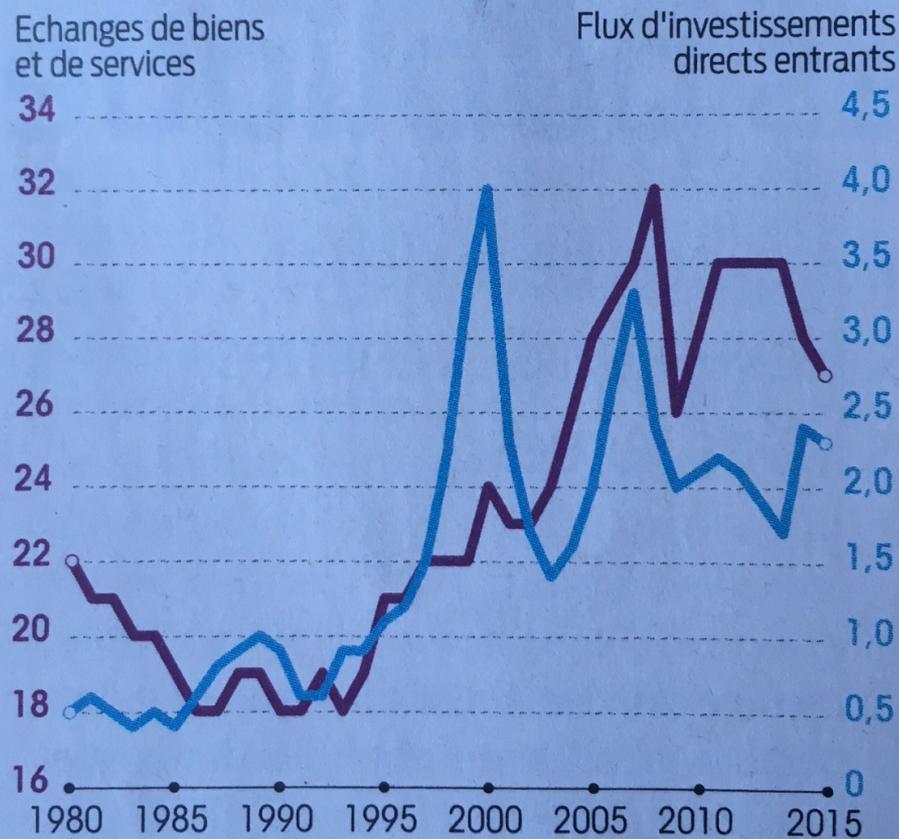
- In between *macro* and *micro*:
meso-networks of collaborative economics (internet platforms)
- Five PhD Thesis in Economics, Informatics, Sociology ...
- Adjusting **collaborative social dynamics** to **technical progress** and *viceversa*

Use of internet platforms for collaborative economics



Le coup d'arrêt de la mondialisation

► Echanges commerciaux et flux entrants
d'investissements directs, en % du PIB mondial



Sources : FMI et Cnucead

The IMF (FMI) and UNICED attribute the slow-down of Globalization to **collaborative economics** (2016 data)

Some references (on stochasticity)

Elowitz, MB, Levine, AJ, Siggia, E & Swain, PS 2002 “Stochastic Gene Expression in a Single Cell”. *Science*, 297.

Kuznetsov, VA, Knott, GD & Bonner, RF 2002 “General statistics of stochastic process of gene expression in eukaryotic cells”. *Genetics*, 161(3):1321–1332.

Paldi, A 2003 “Stochastic gene expression during cell differentiation: order from disorder?” *Cell Mol. Life Sci.*, 60, 1775-1779.

Giuliani, A 2010 “Collective motions and specific effectors: a statistical mechanics perspective on biological regulation” *BMC Genomics*, 11(suppl 1):S2.

Fromion, P, Leoncini, E & Robert, P 2013 “Stochastic gene expression in cells: A point process approach”. *SIAM Journal on Applied Mathematics*, 73(1):195–211.

Marinov, G.K., Williams, B.A., McCue, K., Schroth, G.P., Gertz, J., Myers, R.M. & Wold, B.J. 2014 “From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing”. *Genome Res.* 24, 496-510.

Some more references on Cancer

- Baker, S 2014 “*Recognizing Paradigm Instability in Theories of Carcinogenesis*”, **British Journal of Medicine & Medical Research**, 4(5): 1149-1163.
- Brossel R, Yahi A, David S, Moreno L, Guinebretière J-M (2016) *Mechanical Signals Inhibit Growth of a Grafted Tumor In Vivo*. **PLoS ONE** 11(4)
- Rong Chen et al. (2016) *Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases*, **Nature biotechnology**
- Paska A. et al. (2015) *Repulsive cues combined with physical barriers and cell–cell adhesion determine progenitor cell positioning during organogenesis*, **Nature Comm.**, DOI: 10.1038/ncomms11288
- Karp, RD et al., (1973) “*Tumorigenesis by Millipore filters in mice: histology and ultrastructure of tissue reactions as related to pore size*”. **J Natl Cancer Inst.** 51(4): 1275- 85
- Cai, L., Friedman, N. & Xie, X. S. (2006) *Stochastic protein expression in individual cells at the single molecule level*. **Nature** 440, 358–362 .