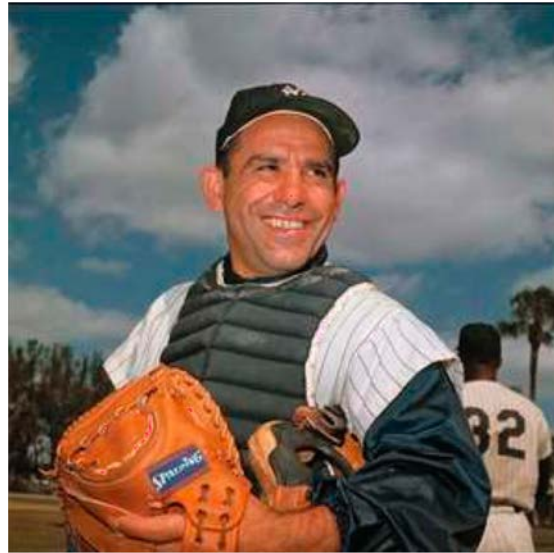


Yogi Berra et ses yogismes



- In theory there is no difference between theory and practice. In practice there is.
- What Time Is It? You Mean Now?
- The towels were so thick at this hotel that I could barely close my suitcase
- Baseball is 90% mental and the other half is physical
- Even Napoleon had his Watergate

En théorie, la théorie et la pratique,
c'est pareil. En pratique, c'est pas vrai.

Yogi Berra

Gérard Berry

Professeur au Collège de France

Chaire Algorithmes, machines et langages

Académie des sciences, Académie des technologies

<http://www.college-de-france.fr/site/gerard-berry>

Cours donné à l'Inria Bordeaux, 13/12/2018



COLLÈGE
DE FRANCE
— 1530 —



Agenda

1. Où la pratique reste confuse malgré une théorie parfaite
2. Où la pratique déborde la théorie
3. Où théorie et pratique convergent et divergent alternativement
4. Où théorie et pratique se marient harmonieusement

Agenda

1. Où la pratique reste confuse malgré une théorie parfaite
 - 1a. Le loto à Antibes
 - 1b. *Le Monty Hall Problem*

1. Où la pratique reste confuse
malgré une théorie parfaite :

*1a. Le Loto à Antibes,
ou les fausses intuitions probabilistes*

Si j'ai une chance sur un million de réussir,
il y a une chance sur deux que j'y arrive

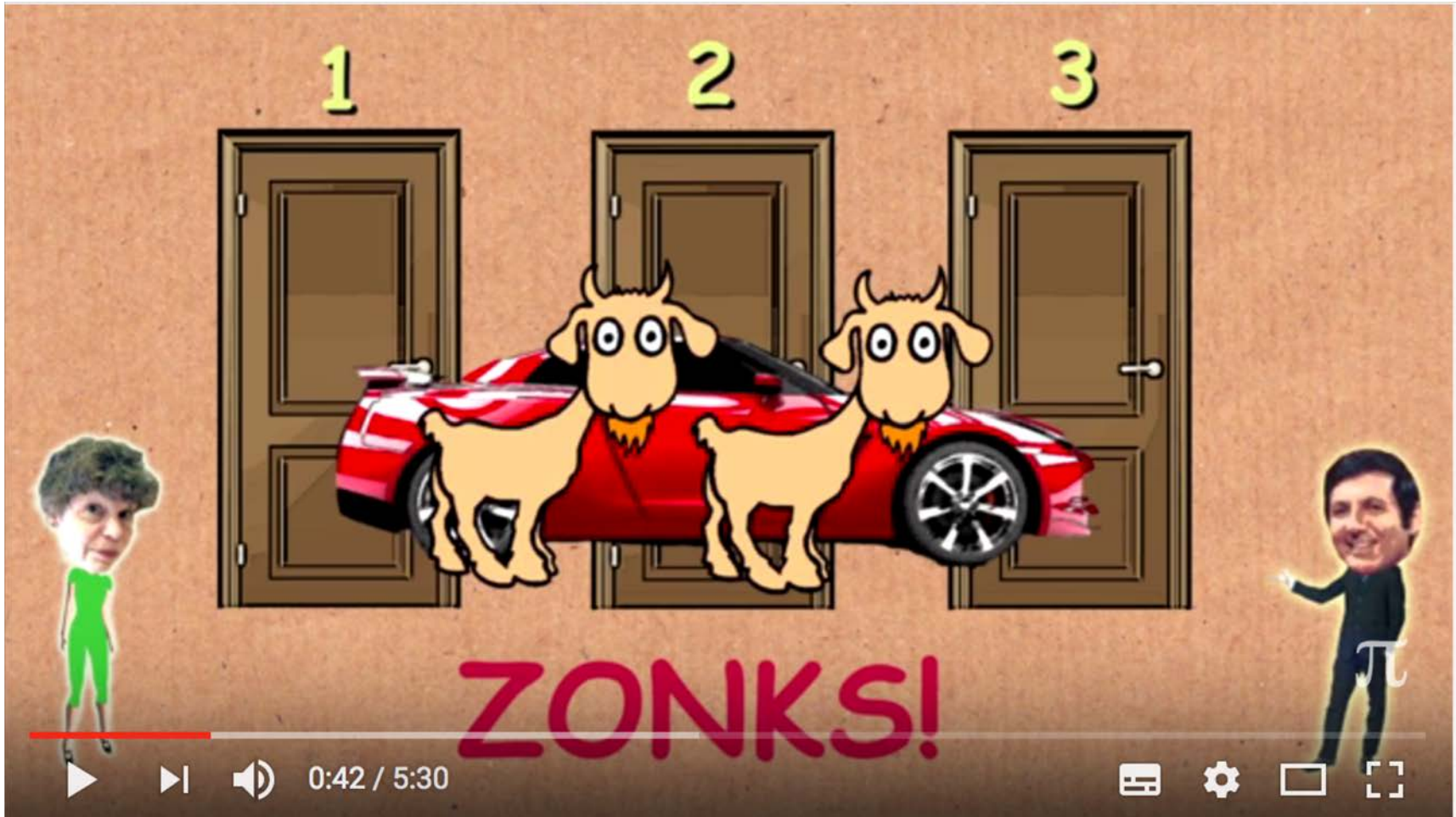
Comme la fusée a une chance sur un million
de marcher, ratons vite les 999 999 premiers essais
Le Professeur Shadoko

Au tabac d'Antibes, vers 1978

- Ma grille : 1 2 3 4 5 6
- Monsieur, je refuse de valider votre grille !
- Mais pourquoi donc ?
- Parce qu'ici les gens viennent pas pour perdre !
- Mais une grille ou une autre c'est pareil, non ?
- Mais enfin, tous vos numéros sont serrés sur la même ligne ! Il faut *ré-par-tir* !
- ...
- Vous faites quoi comme métier ?
- Je suis mathématicien
- Ça m'étonne pas ! raison de plus !

The Monty Hall Problem

(lire l'excellent livre de Jason Rosenhouse)



The Monty Hall Problem *(lire l'excellent livre de Jason Rosenhouse)*



Changer son choix, ou le maintenir ?

Ma probabilité de gain au départ était $1/3$



1. Ma probabilité est passée à $1/2$ car il n'y a plus que deux portes, cela que je change ou pas
2. Ma probabilité devient $1/2$ si je ne change pas, **mais** passe à $2/3$ si je change, donc je change

Une histoire étonnante !



- En 1990, Marilyn von Savant (Guinness record de QI) traite la question dans *Parade* en répondant à la question d'un lecteur, et donne la bonne réponse :
- Il faut changer

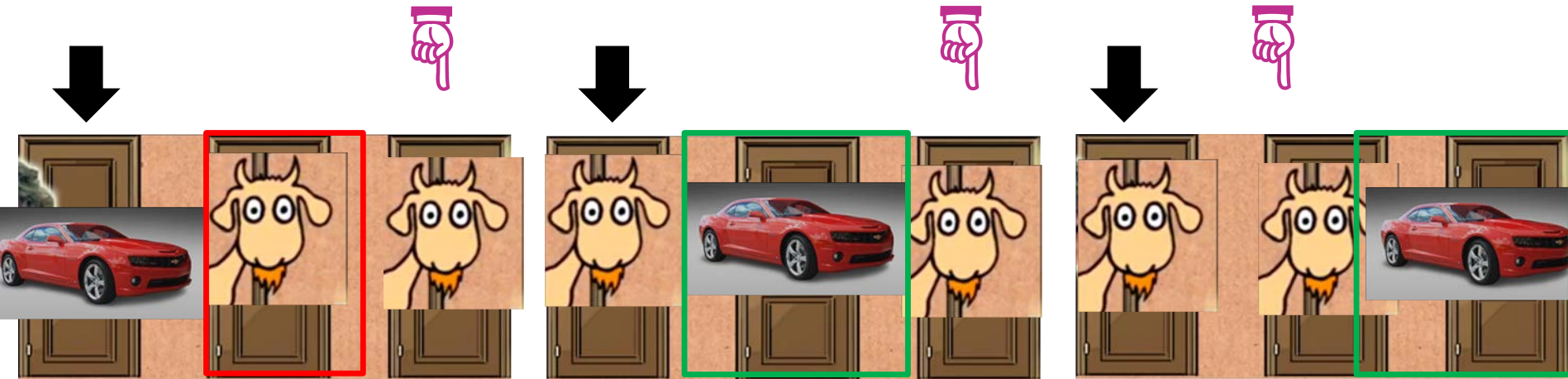
I'm receiving thousands of letters nearly all insisting I'm wrong... Of the letters from the general public, 92% are against my answer; and of letters from University, 65% are against my answer.


You are **utterly incorrect** about the game-show question, and I hope this controversy will call some public national attention to **the serious national crisis in mathematical education**. If you can admit your error, you will have contributed constructively towards the solution of a deplorable situation. **How many irate mathematicians are needed to get you to change your mind?**



Même l'immense
Paul Erdős ne voulait pas y croire !

Mais oui, il faut changer ! La preuve :



 2 fois sur 3

Avantage personnel :
je n'ai pas vu que c'était un problème de probas !

Comment M. von Savant s'en est tirée

Faire faire l'expérience par des centaines de classes,
sur des centaines de tirages.

Ils ont tous trouvé $2/3$!

(idem pour Erdős)

Our class, with unbridled enthusiasm, is proud to announce that our data supports your position. Thank you so much for your faith in America's educators to solve this.



Agenda

- 2. Où la pratique déborde la théorie
 - 2a. La logique Booléenne et ses algorithmes
 - 2b. La loi de Moore
 - 2c. Les fausses affirmations médiatiques
 - 2d. La sécurité informatique

2a. Où la pratique déborde la théorie: la logique Booléenne et ses algorithmes

Calcul logique de base, mais aussi super-solveur
de problèmes combinatoires finis
voir cours SAT et SMT des 16 et 23 mars 2016

Half of the lies they say about me aren't true

You should always go to other people's
funerals, otherwise, they won't come to yours.

Yogi Berra

Le problème SAT (cf. cours du 16/03/2016)

- Formule Booléenne :

- variables x, y à valeur 0 ou 1 (vrai ou faux, ...)
- opérateurs et (\wedge), ou (\vee), non (\neg), implique (\implies), ...
- littéraux : x, \bar{x} pour $\neg x$

- CNF = forme normale conjonctive,

$$(x \vee y \vee \bar{z}) \wedge (\bar{y} \vee z \vee \bar{t} \vee u) \wedge \bar{v}$$

clause

- SAT : la formule peut-elle être rendue vraie par une assignation de 0 ou 1 aux variables ?

- $x = 1, z = 1, v = 0$ (ou $x \wedge z \wedge \bar{v}$)

- $z = 0, u = 1, v = 0$ (ou $\bar{z} \wedge u \wedge \bar{v}$)

- $y = 1, t = 0, v = 0$ (ou $y \wedge \bar{t} \wedge \bar{v}$)

cube

...

Cook 1971 : la NP-complétude

1. Problème NP : facile de vérifier une proposition de solution en temps polynomial, mais dur de trouver une solution (polynomial ou exponentiel ?)
2. De très nombreux problèmes combinatoires importants, dits **NP-complets**, sont équivalents : si on sait résoudre l'un en temps polynomial, cela vaut pour les autres.

SAT (Cook, 1971), **pavage du plan** (Levin, 1973)

21 problèmes variés (Karp, 1972)

Maintenant : **des milliers de problèmes** dans de nombreux domaines : graphes, optimisation, ordonnancement, réseaux, logique, algèbre,...

https://fr.wikipedia.org/wiki/Liste_de_problèmes_NP-complets

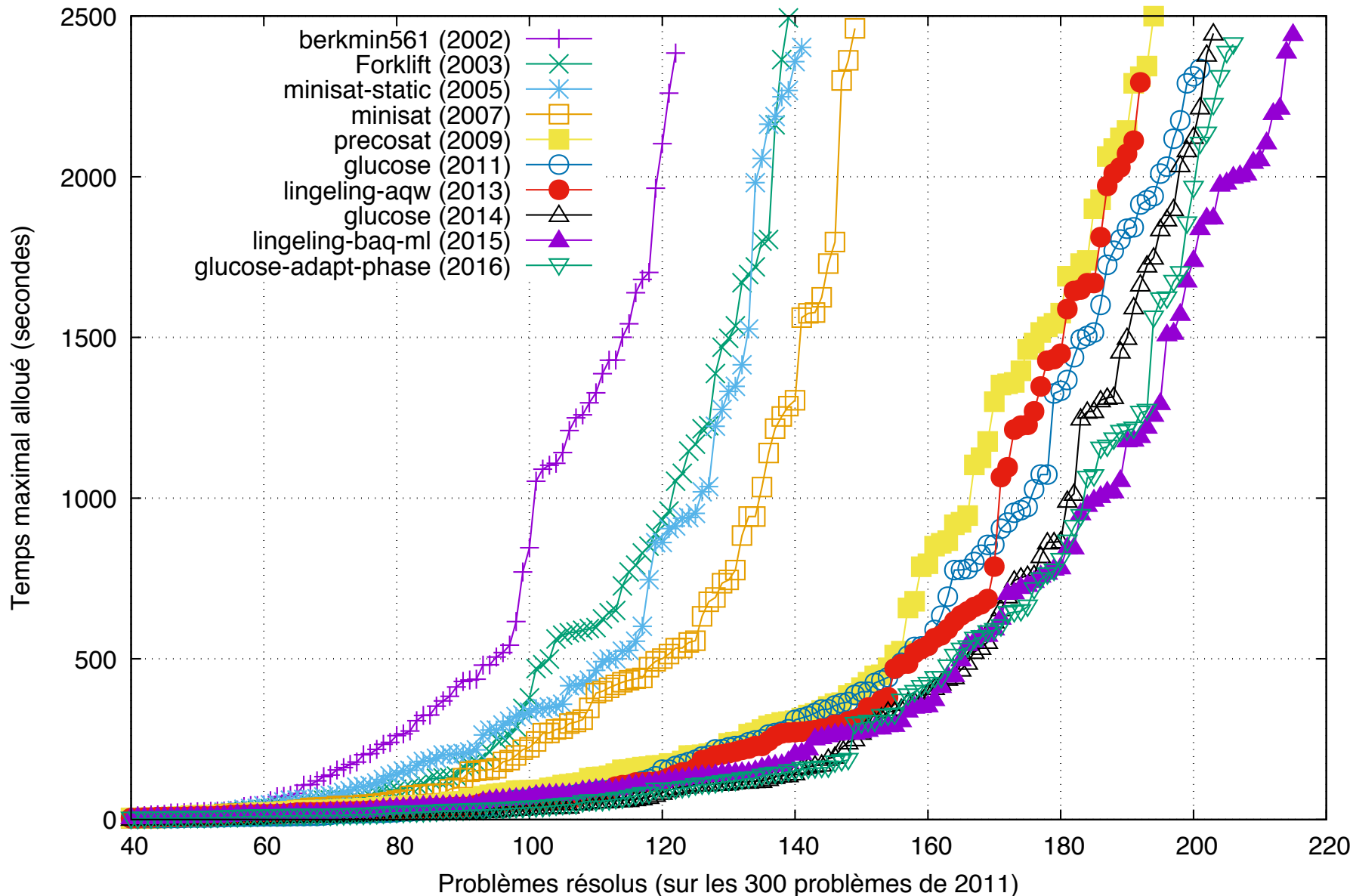
Théorie vs. pratique

Conjecture : dans les pires cas, résoudre SAT est exponentiel dans le nombre de variables

exemple : 14 trous et 15 pigeons \Rightarrow 2 dans le même trou ?

- **Pratique** : heuristiques résolvant SAT dans les cas pratiques, par exemple **industriels** ou **mathématiques**
 - 1985 : **DPLL**, exploration / backtrack
 - 1986 : **BDDs** (Bryant), forme normale compacte
 - 1996 : **GRASP (CDCL)**, apprentissage pour mieux explorer
 - 2001 : **CHAFF** : structures de données magiques
 - **Maintenant** : progrès constants, très nombreuses applications, **mais on ne sait toujours pas pourquoi ça marche !**
(cf. *Deep Learning*, cours de Stéphane Mallat)

Des progrès considérables en 15 ans !



Mathématiques : triplets pythagoriciens

$$a, b, c \in \mathbb{N}. a^2 + b^2 = c^2$$

Problème : peut-on partitionner \mathbb{N} en deux sous-ensembles disjoints tels qu'aucun ne contient de triplets pythagoriciens ?

Lemme : si c'est faux pour le sous-ensemble $\{1, \dots, n > 0\}$, alors c'est faux pour tous les sous-ensembles plus grands et pour \mathbb{N}

Réponse : vrai pour $n \leq 7824$, faux pour $n \geq 7825$

Preuve : pour $n > 0$, construire une formule booléenne $F(n)$ satisfiable ssi elle résout le problème pour n , et monter $SAT(F(7824))$ et $UNSAT(F(7825))$

Solving and Verifying the Boolean Pythagorean Triples problem via Cubes and Conquer

M. Heule, O. Kullmann et V. Marek, 2016

Construction de la formule pour $\{1, \dots, n\}$

- Une variable x_i pour chaque $i \leq n$.
- Pour chaque partition (P, N) de $\{1, \dots, n\}$, le littéral x_i code $i \in P$ et le littéral \bar{x}_i code $i \in N$
- Pour chaque triplet pythagoricien $a, b, c < n$, on ajoute deux clauses interdisant a, b, c tous trois dans P ou dans N :

$$(x_a \vee x_b \vee x_c) \wedge (\bar{x}_a \vee \bar{x}_b \vee \bar{x}_c)$$

- Résultat pour 7825 : 6494 variables, 19844 clauses
- Réduction : $F(7825)$: 3745 variables, 14672 clauses

Calcul en 3 phases

- *Cube splitting* : bien choisir une variable x_i (heuristiques fines) et résoudre séparément $x_i \wedge F(x_i \leftarrow 1)$ et $\overline{x_i} \wedge F(x_i \leftarrow 0)$, itérer pour couper avec 10^6 cubes du type $x_i \wedge x_j \wedge \overline{x_k} \dots$ couvrants, c'est à dire de v global 1
- Traiter les 10^6 formules obtenues **en parallèle** (mélange de CDCL incrémental et de cube splitting)
- **Une seule est SAT pour 7824**, facile de construire toutes les solutions et de les propager à la formule globale
- Pour 7825, construire des **preuves logiques UNSAT** (subtil, **Glucose 3.0**, Audebert & Simon), puis les vérifier indépendamment (facile mais long), et assembler et vérifier le tout (pas facile)

Coût du calcul

Calculateur parallèle, 800 cœurs :

35 000 heure CPU pour SAT / UNSAT

16000 heures CPU pour la validation des preuves UNSAT

Une contradiction explicite pour 7825 :

$$5180^2 + 5865^2 = 7825^2 \quad 625^2 + 7800^2 = 7825^2$$

par 7824, 5180 et 5865 doivent être dans le même ensemble,
alors que 625 et 7800 doivent être dans l'autre

Taille de la preuve originale :

200 téra-octets ($1,6 \cdot 10^{15}$ bits) !

~~*Two-hundred-terabyte proof is the largest ever*~~

~~Evelyn Lamb, Nature, 26 mai 2016~~

La taille au dessus : Schur(5), Heule 2017

Théorème (Schur) : pour tout entier k , on peut colorier les entiers avec k couleurs (partitionner les entiers) de façon telle qu'aucune couleur ne colorie trois nombres a, b, c tels que $a + b = c$

Problème : pour tout entier k , calculer $S(k)$ maximal tel que la propriété soit vraie pour $\{1, \dots, S(k)\}$

Réponses : $S(1) = 1, S(2) = 4, S(3) = 16, S(4) = 44,$

$$S(5) = 160$$

Schur Number Five

Marin J.H. Heule, [arXiv.org>cs>arXiv:1711.08076](https://arxiv.org/abs/1711.08076)

La preuve SAT



- Trois outils : **March_cu** pour cube splitting
Glucose 3.0 pour SAT / UNSAT
Vérifieur des preuve UNSAT en ACL2
- Coût pour 160 → SAT : **14 ans CPU** (en 3 jours)
161 → UNSAT : **36 ans CPU**
Lonestar 5 (Texas), 50 x 24 cores bi-threads
- Preuve : **2,1 Péta-octets** (comprimée puis simplifiée)
- Pour 160 : **2 447 113 088 colorations possibles**
 - Cube splitting : 10 330 615 cubes, seulement **961 SAT**
- Pour 161 : **UNSAT** pour ces 961 cubes !



Peut-on trouver une preuve vraiment mathématique ?

2b. Où la pratique déborde la théorie: La « Loi » de Gordon Moore

~~La puissance des ordinateurs va doubler tous les 18 mois~~

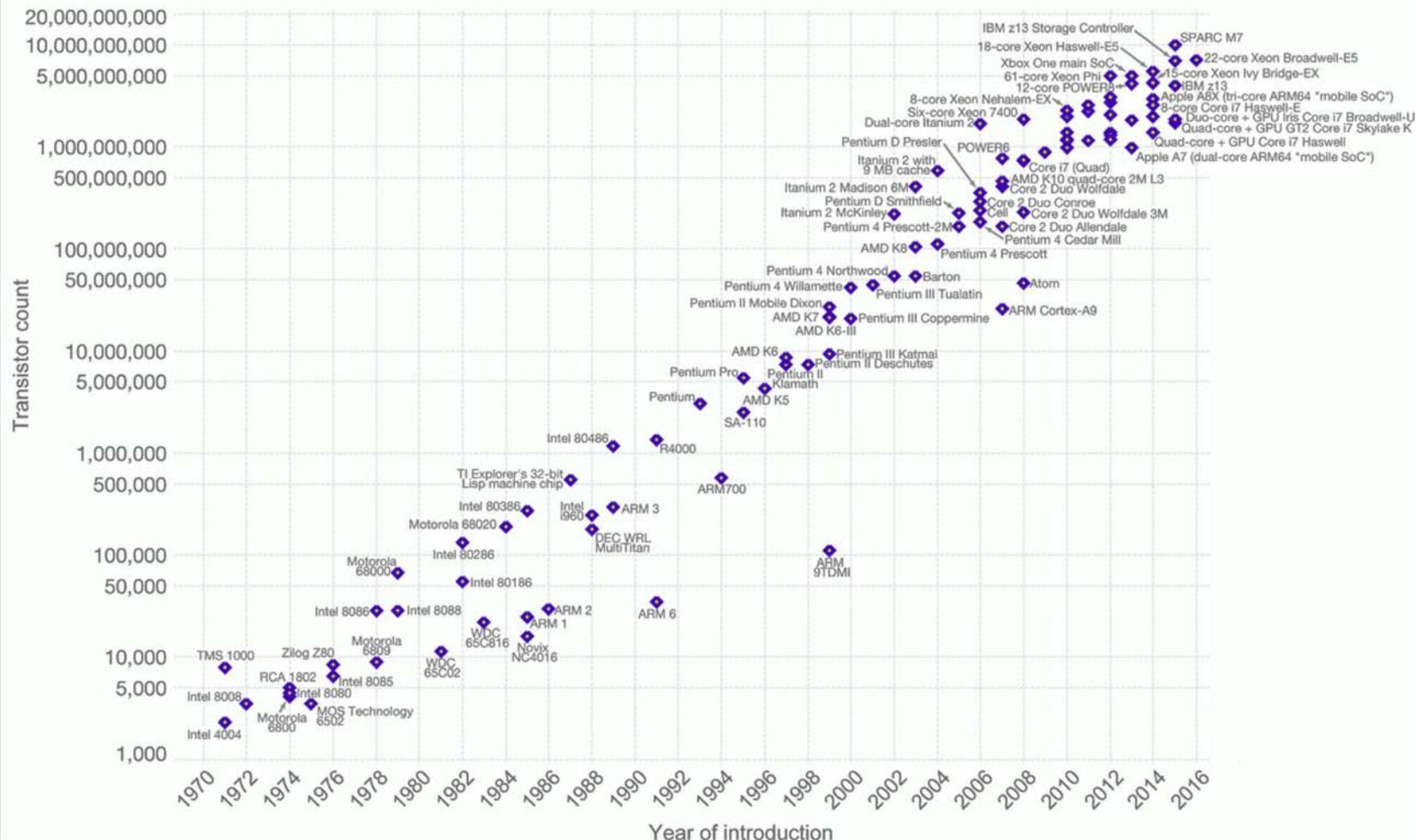
 Le nombre de composants électroniques par unité de surface double tous les 2 ans 

La loi de Moore

Moore's Law – The number of transistors on integrated circuit chips (1971-2016)



Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore's law.



Data source: Wikipedia (https://en.wikipedia.org/wiki/Transistor_count)

The data visualization is available at [OurWorldinData.org](https://ourworldindata.org). There you find more visualizations and research on this topic.

Licensed under CC-BY-SA by the author Max Roser.

Pratique contre fausse théorie

Pas une loi physique, mais une **décision pratique concertée** de l'ensemble des industries concernées :
semiconducteurs, lasers, chimie, CAO, robotique, etc.
Appliquée de façon étonnante depuis 1971 (47 ans !)

Loi complémentaire : Le nombre de « théoriciens »
« bien informés » présidant la fin de la loi de Moore
par atteinte des « limites de la physique »
double tous les 2 ans

C'est sûr, « ils » toucheront la limite à
50 → 20 → 15 → 10 nm



L'avancée continue toujours

- Apple / TSMC, iPhone A12 Z : **7 nm**, 9 milliards de transistors
- 2019 : **5 nm en production** (Samsung)
- 2022 : **3 nm en chantier** (TSMC)
- Plusieurs équipes ont construit des transistors à **1 atome** (0,35 nm) !

Mais la loi de Moore ralentit

- Le mur de la chaleur
- La difficulté croissante de la CAO électronique
- Le prix démesuré des nouvelles usines
 - Prévission : usine TSMC 3nm, 2022 : \$ 20 milliards
Chine : > \$ 30 milliards
- Réduction associée du nombre de fournisseurs
 - ST Microelectronics : Abandon à 28 nm ?
 - Août 2018 : Global Foundries (ex. AMD) abandonne le 7 nm pour des raisons de coût





Enfin de la place pour d'autres approches !
machines neuromorphes, memristors, graphène, ...
→ nouvelle recherche excitante

2c. Belles affirmations, mais fausses ? (quand les médias se lâchent)

L'ordinateur quantique sera
infiniment plus rapide que l'ordinateur classique,
et calculera beaucoup plus de choses

1. Pour les calculs classiques, est au mieux exponentiellement meilleur (ce qui est bien), ou \sqrt{n} meilleur
2. Il calcule les même fonctions, mais aussi l'aléatoire, aussi greffable à l'ordinateur classique (générateurs quantiques)
3. Les problèmes technologiques restent considérables

Ceci ne vaut pas pour la simulation de phénomènes quantiques

2c. Belles affirmations, mais fausses ? (Quands les commentateurs se lâchent)

Grâce à Internet, l'accès aux réseaux et l'informatique distribuée sont devenus **ubiquitaires**

On sait que c'est faux **dès qu'on habite à la campagne !**

- **téléphone** en bout de ligne, ADSL 512 K, lignes non maintenues
- **portable** ne passant pas, ou trop cher pour Internet
- **plans numériques** répétés, mais peu d'effet

Il faudra un investissement énorme et pas seulement urbain pour que ça devienne plus ou moins vrai (5G?)

2d. Une communication chiffrée est sûre

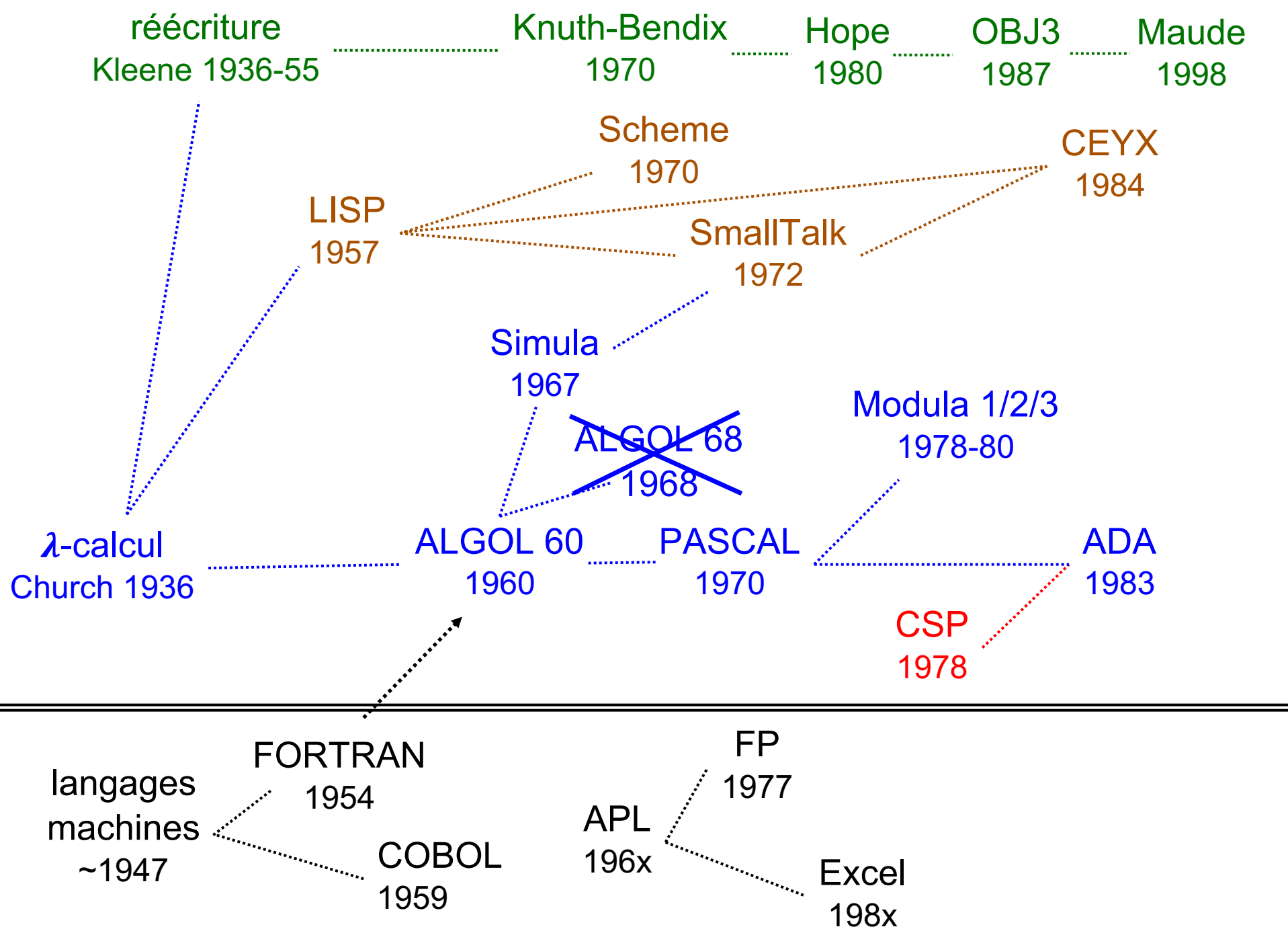
- Le **déchiffrement brutal** n'est pas toujours impossible (cf. NSA pour vpn, téléphone, etc.)
- Attaques par **canaux cachés** (Adi Shamir sur RSA, voitures)
- Souvent, les communications sont cassables par les **faiblesses des normes de sécurité** et **les bugs des protocoles de sécurité** plus que par le déchiffrement brutal
- **Freak** (Microsoft, Inria, etc.), *man in the middle* pour forcer un chiffrage faible maintenu pour parler à de vieux appareils
- **Logjam** (CNRS, Inria) : trop peu de groupes Diffie Hellmann
- 12 / 2017 : **attaques sur les protocoles 5G** (ETH, Nancy, etc.)

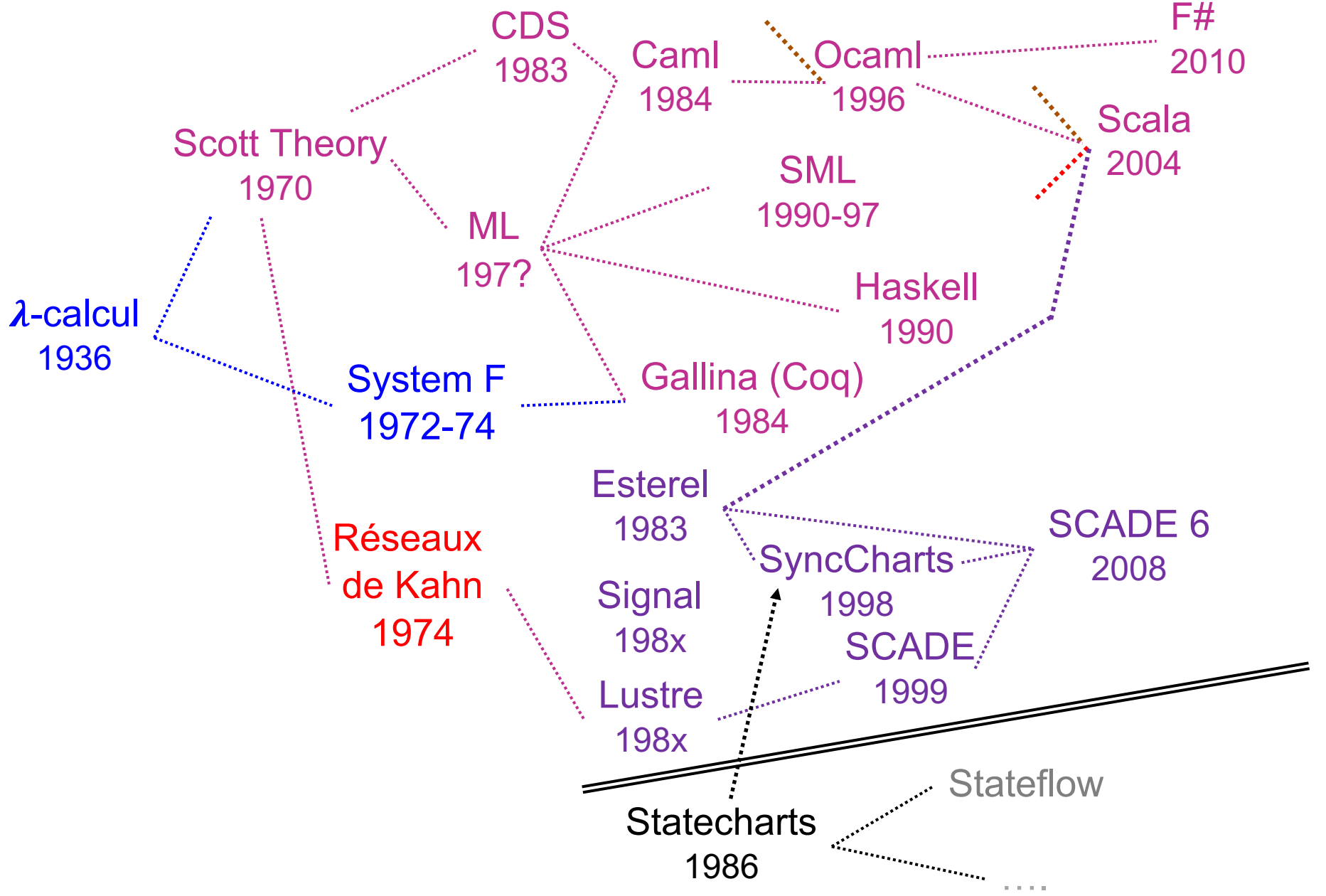
Voir mon cours sécurité du 13 février 2019
avec un séminaire de Guillaume Poupard, DG de l'ANSSI

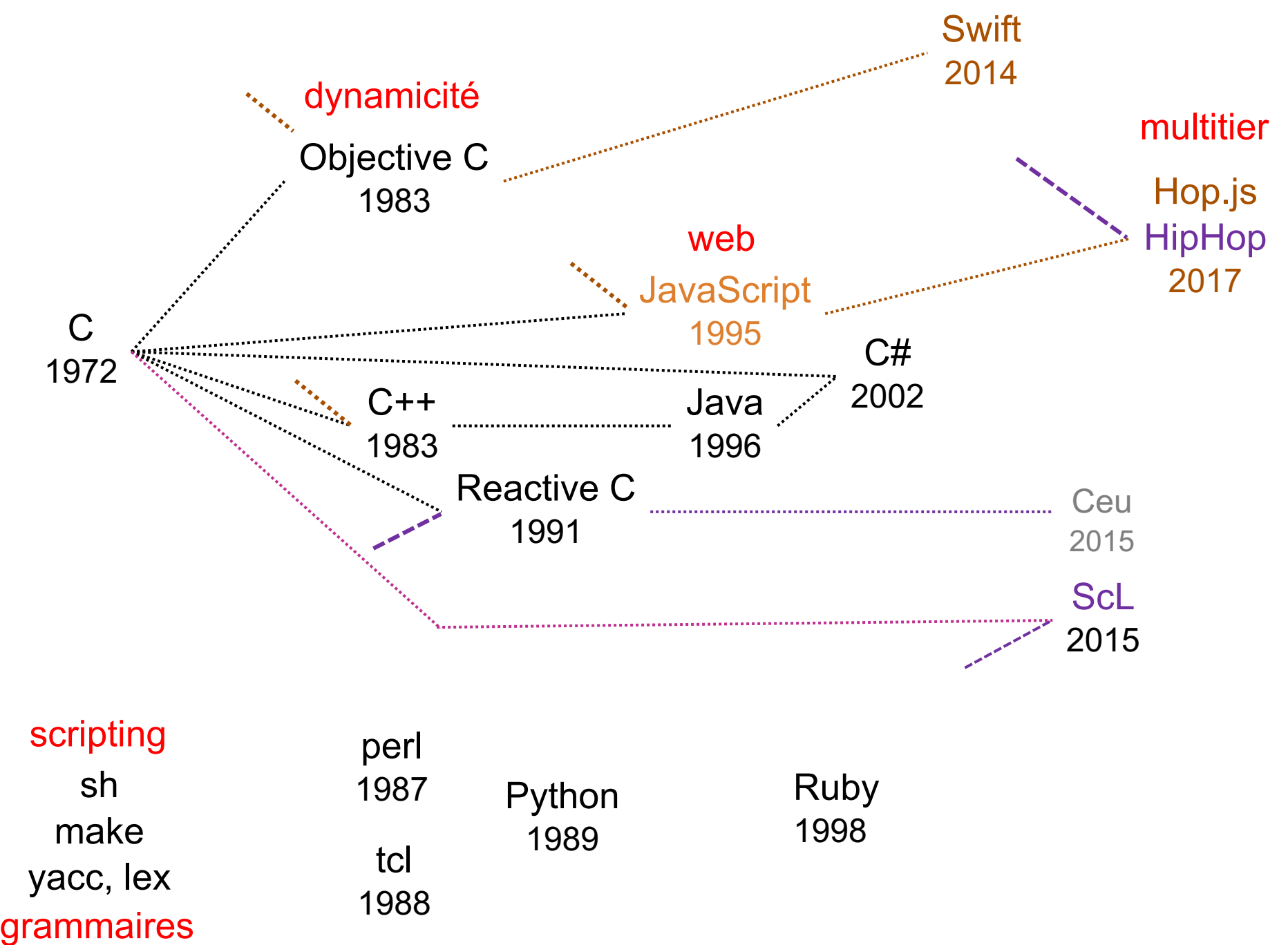
Agenda

1. Où théorie et pratique convergent et divergent
alternativement

La saga des langages de programmation



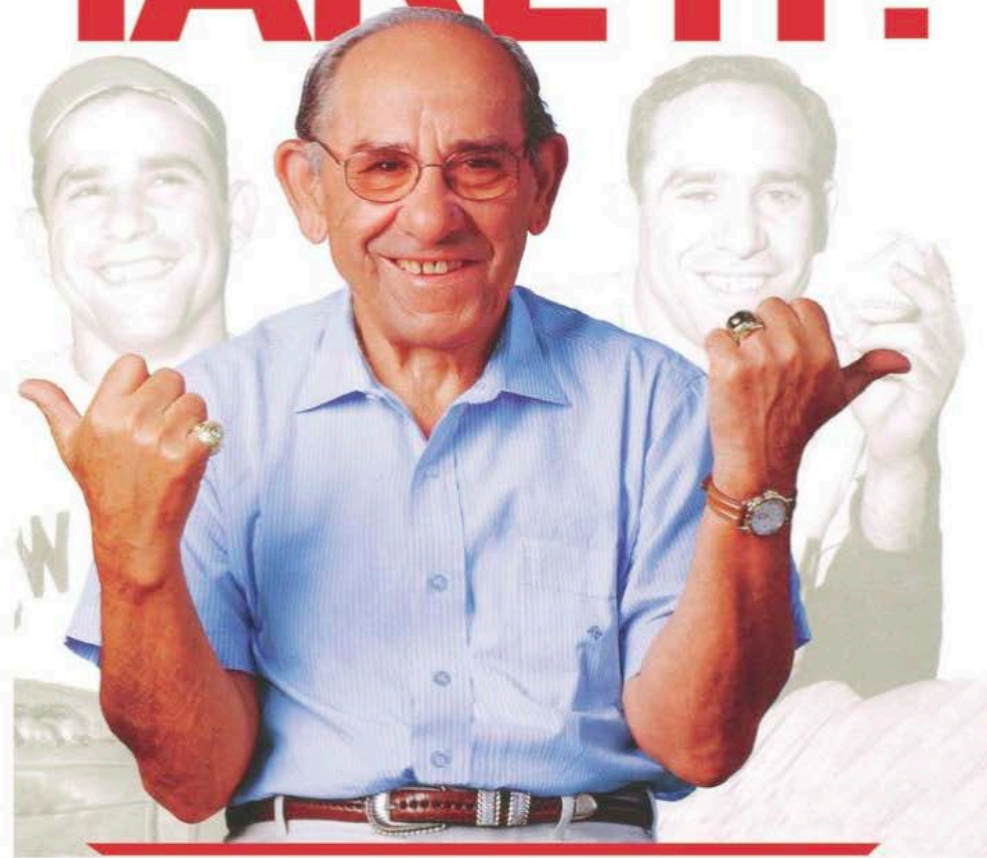




NEW YORK TIMES BESTSELLER

When You Come to a Fork in the Road,

TAKE IT!



Inspiration and Wisdom from One of Baseball's Greatest Heroes

Y  **GI BERRA**

with DAVE KAPLAN

Agenda

- 4. Où théorie et pratique se marient harmonieusement
 - 4a. Programmation et vérification
 - 4b. Algorithmes randomisés
 - 4c. Informatique et sciences, simulation

4a. Programmation et vérification

- Langages typés avec sémantique claire
 - généraux : [Caml](#), [Haskell](#), [Scala](#),...
 - spécifiques : [Esterel](#), [Lustre](#), [Signal](#), [ScL](#), [Zélus](#),...
 - mathématiques : [Mathematica](#)
- Analyse de programmes par interprétation abstraite
 - D. Scott (1970), P. & R. Cousot (1977)
 - Industrialisation par [Polyspace](#) (Matlab), [AbsInt](#)
 - Utilisation industrielle : Airbus, automobile, nucléaire, etc.
- Systèmes logiques pour l'assistance à la preuve
 - classiques : [HOL](#), [HOL Light](#), [Isabelle](#)
 - Curry-Howard : [Coq](#)
 - Recherche sur lisibilité / d'automatisation : [Mizar](#), [Why3](#), ...

4b. les algorithmes randomisés

La randomisation bien dirigée est un bon moyen
d'explorer des espaces complexes
et elle peut facilement se paralléliser

Ethernet, du protocole standard à Kapetanakis

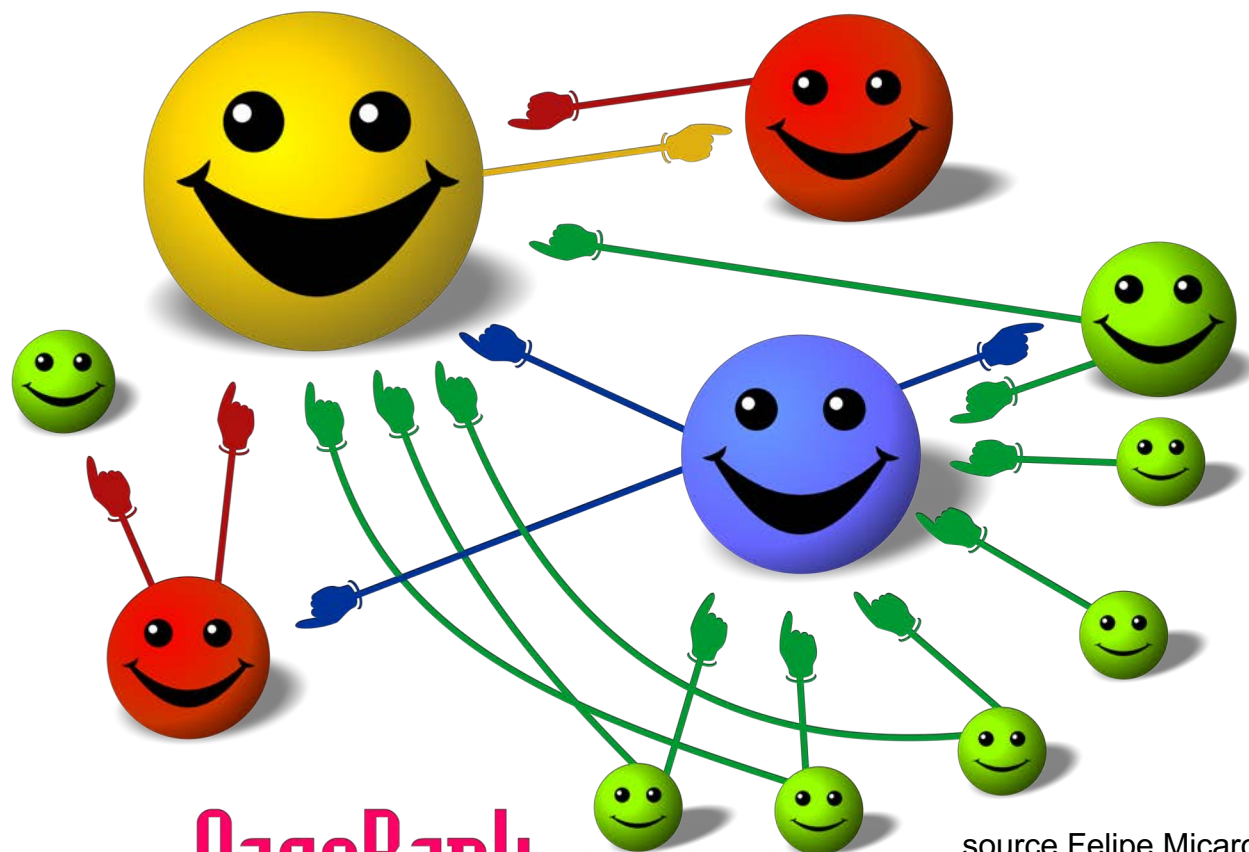
PageRank de Google

Algorithmes géométriques

Analyse de très grands graphes, etc.

Voir les cours de Jean-Daniel Boissonnat (2016-2017)
et Claire Mathieu (2017-2018)
sur la chaire Informatique et sciences numériques

PageRank : la popularité des pages web



PageRank

source Felipe Micaroni Lalli, CC BY-SA 2.5,
<https://commons.wikimedia.org/w/index.php?curid=2776582>

$$P(A) = \frac{1 - d}{N} + d \times \sum_{i=1}^n \frac{P(B_i)}{C(B_i)}$$

En cadeau : mon algorithme randomisé pour trancher l'indécision

Si vous hésitez entre **A** et **B**, tirez au sort,
et rejouez autant que nécessaire



4c. L'informatique et les sciences : mêler informatique et théories scientifiques

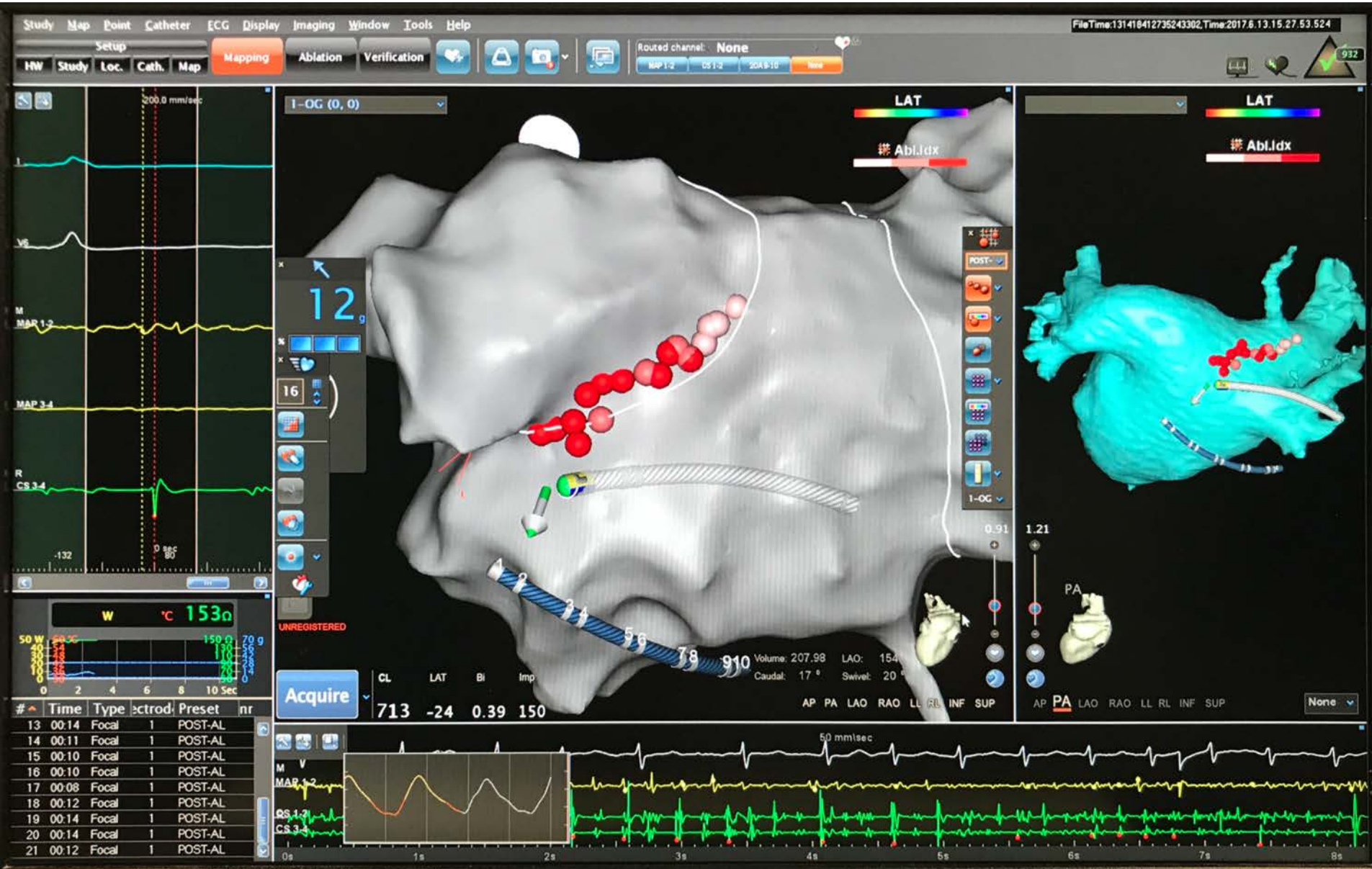
- Le traitement d'images et la réalité virtuelle
 - imagerie médicale
- La simulation des phénomènes physiques ou biologiques
 - équations : maths appliquées, calcul numérique
 - autres modèles : programmation par contraintes et algèbres de processus en biologie, systèmes de Lindenmaier en botanique, etc.
- La conception et la réalisation des nouveaux instruments scientifiques

Médecine : intervenir par l'image



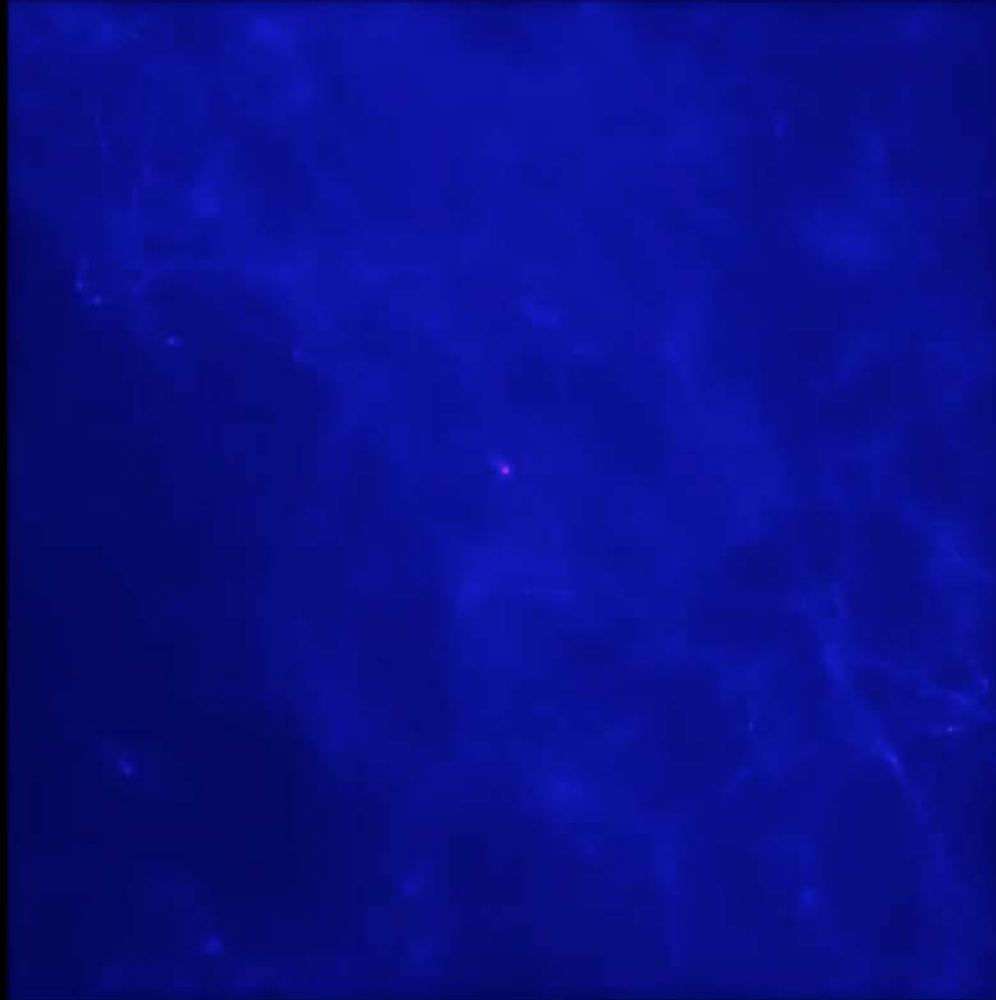
**Radiologie interventionnelle
(voir Colloques du 02/05/18 et 23/04/19)**

Traitement de fibrillations de l'oreillette



Collisions de galaxies

$a = 0.121$ $z = 7.3$ $t = -13.0$ Gyr



Source F. Combes, Collège de France

Simulation des dunes de sable par automates cellulaires

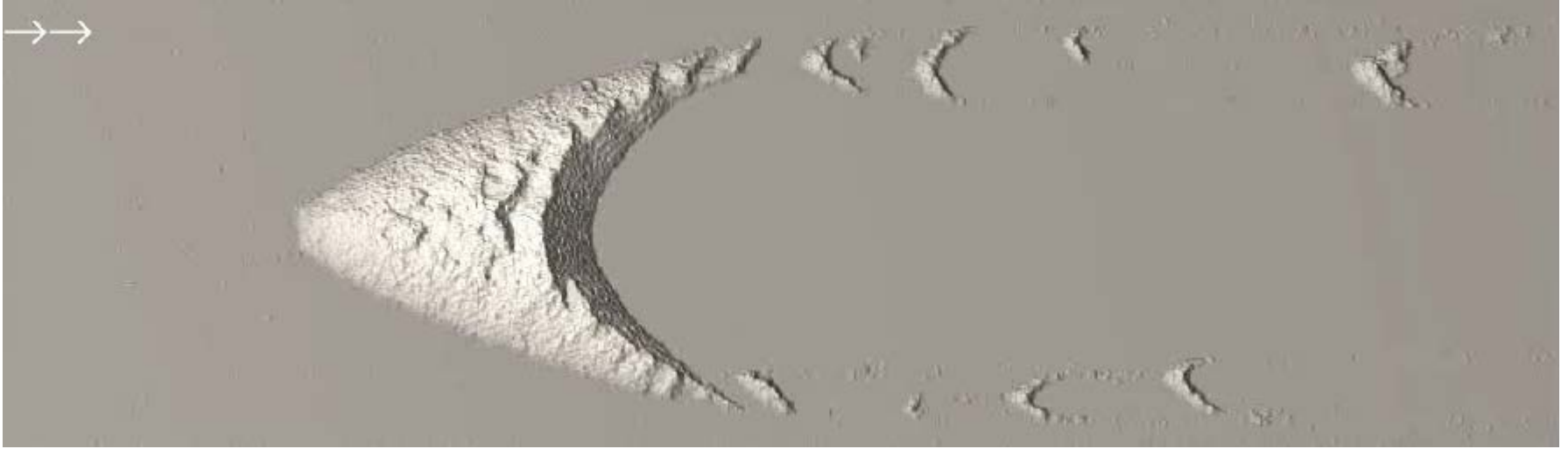


Cf. cours du 28/01/2015 et [Clément Narteau](#), IPGP,
Sculptures éoliennes au sein des mers de sable

Colloque « Arts et sciences, de nouveaux domaines pour l'informatique »,

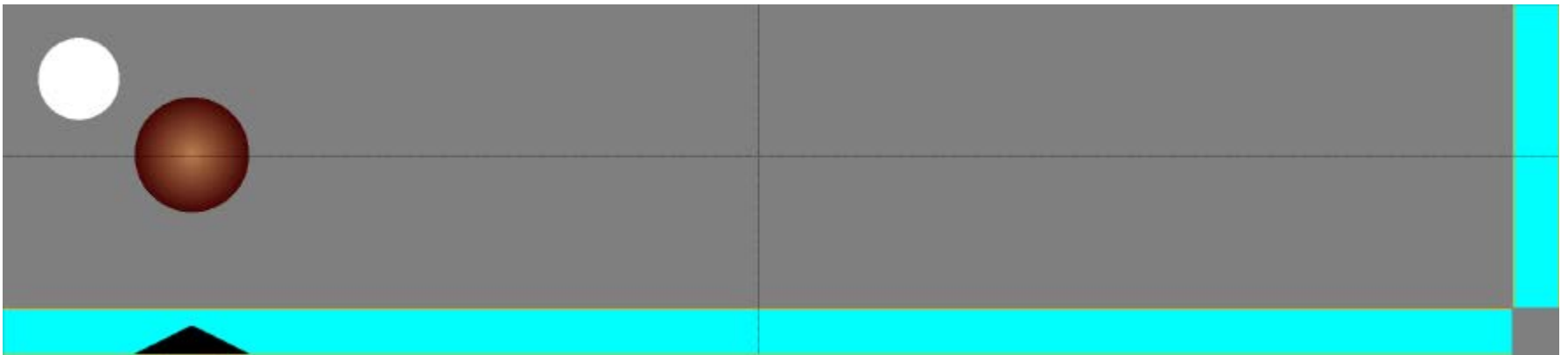
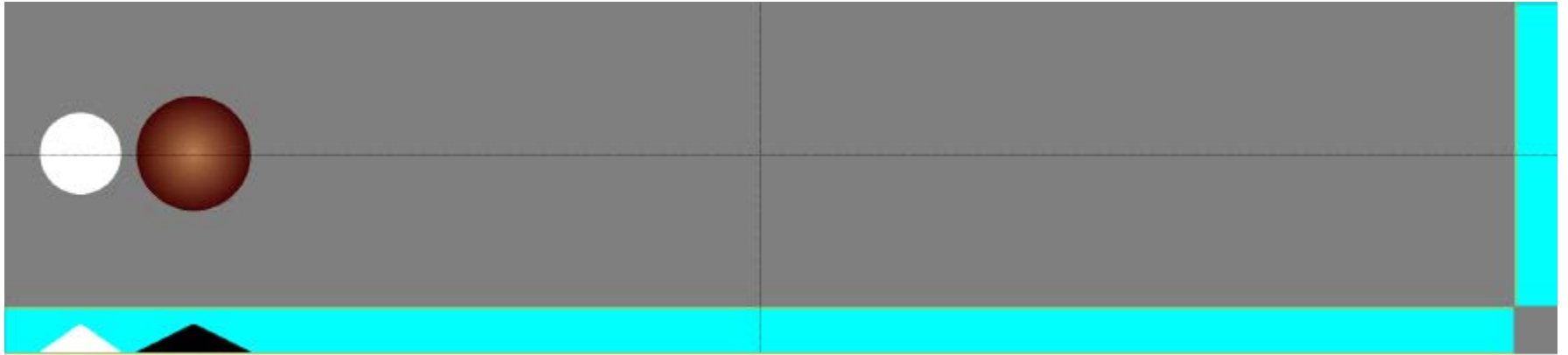
<https://www.college-de-france.fr/site/gerard-berry/symposium-2016-05-27-14h00.htm>

Simulation des dunes de sable par automates cellulaires

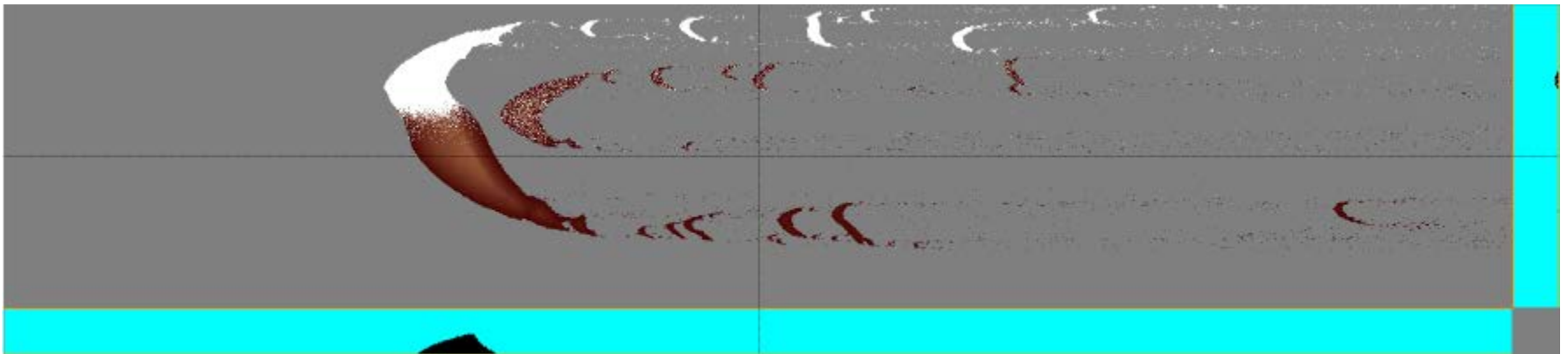
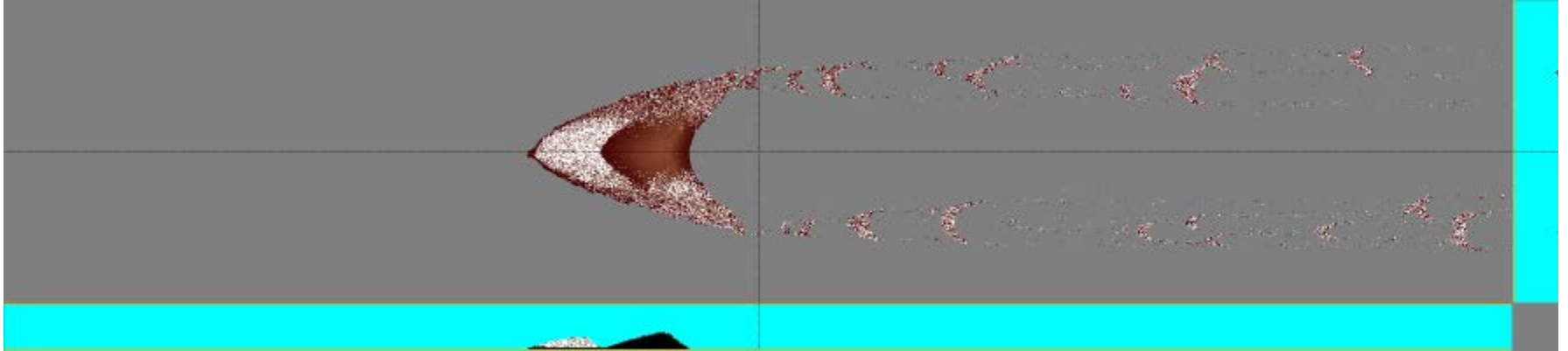


*Source Clément Narteau
Institut de Physique du Globe de Paris*

Collision frontale ou latérale



Collision frontale ou latérale



*Source Clément Narteau
Institut de Physique du Globe de Paris*

Au fond, la simulation, c'est quoi ?

- Remplacer matière, énergie et ondes par la seule **information**
- Remplacer les lois de la nature par leur **équivalent algorithmique**
- Remplacer le temps physique par le temps de calcul
 - **simulations rapides de phénomènes lents**
 - **simulations lentes de phénomène rapides**
 - **simulations en temps réel**

Il faut bien sûr de la matière et de l'énergie et des ondes pour simuler en machine, mais celles si, sont **universelles** et sans rapport avec celles du phénomène simulé !

On ne trouve pas de pétrole en forant la carte...

Conclusion

The best practice is inspired by theory

Donald Knuth

En théorie, tout va bien, quand je
serai grand, j'irai vivre là-bas !

*Un enfant qui a tout compris
... mais qui risque de beaucoup rater !*

Yogi Berra logicien

- Never answer an anonymous letter
- I really didn't say everything I said
- If you ask me a question I don't know, I'm not going to answer
- No one goes there nowadays, it's too crowded
- You've got to be very careful if you don't know where you are going, because you might not get there.

S'il vous plaît, coupez ma pizza en 4 parce que je ne n'ai pas assez faim pour en manger 6 parts