Science mathématique et science informatique:

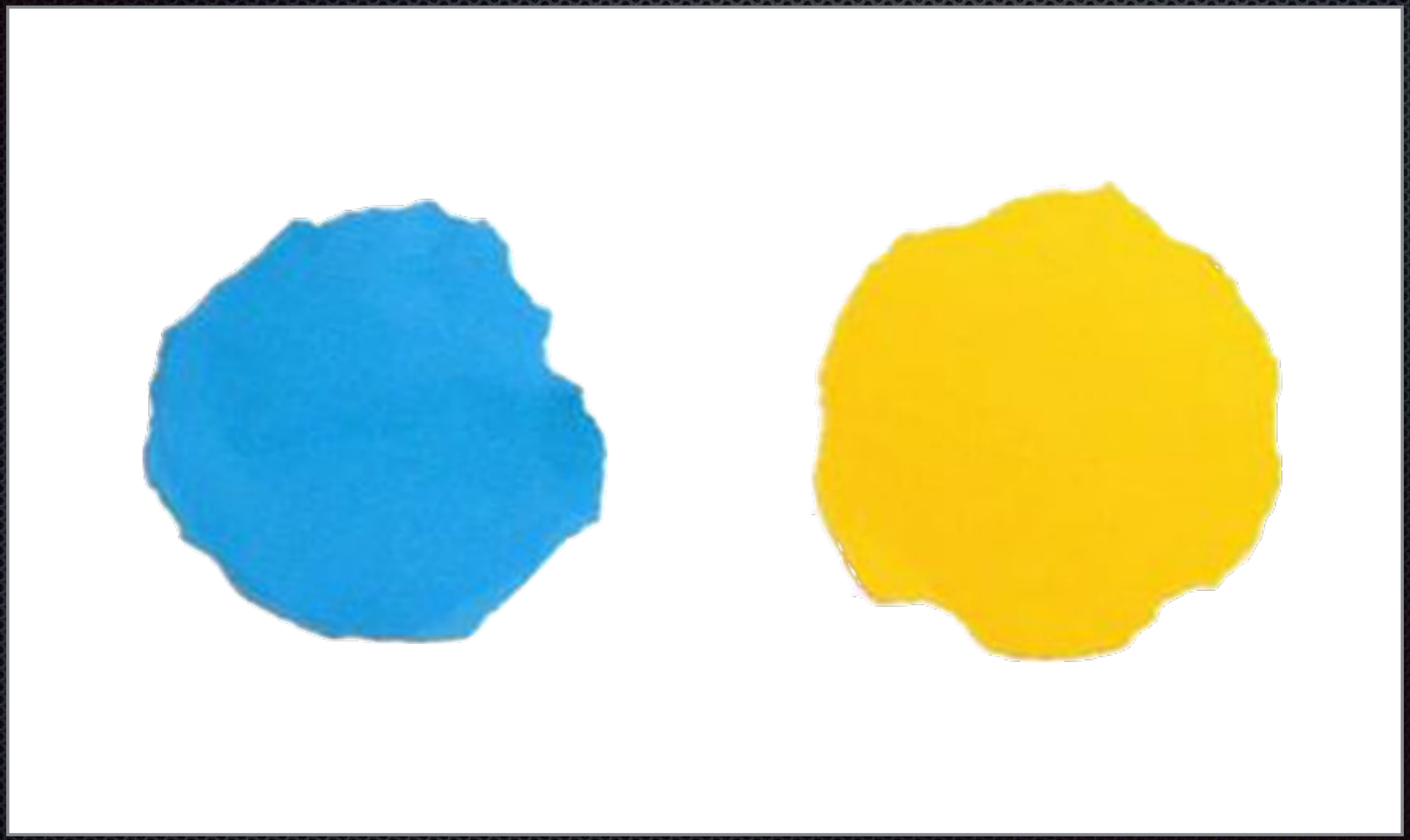
Un couple d'avenir!

Luc Bougé, ENS Rennes

Journées Algorithmique et numérique au collège
IREM, Lille, 8 février 2016

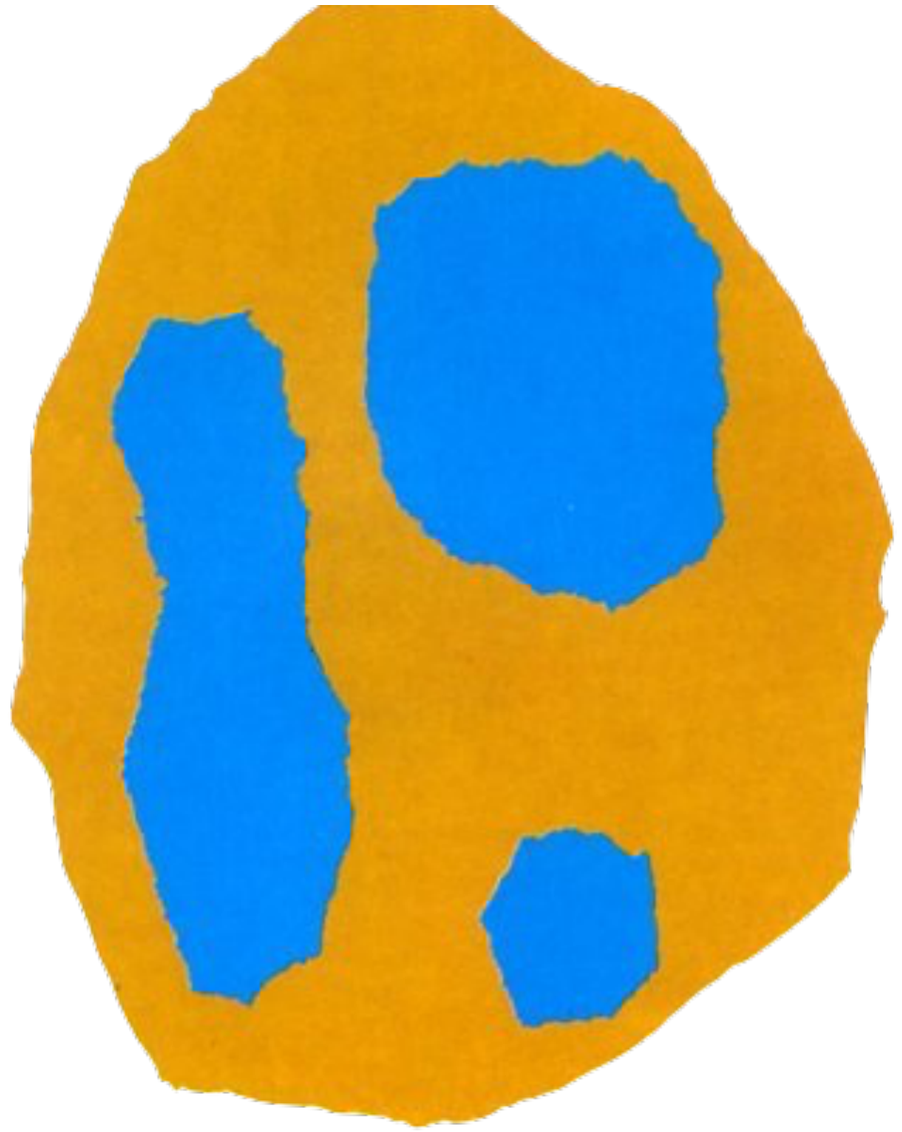

Petit-Bleu et Petit-Jaune
Leo Lionni

Un couple d'amis…

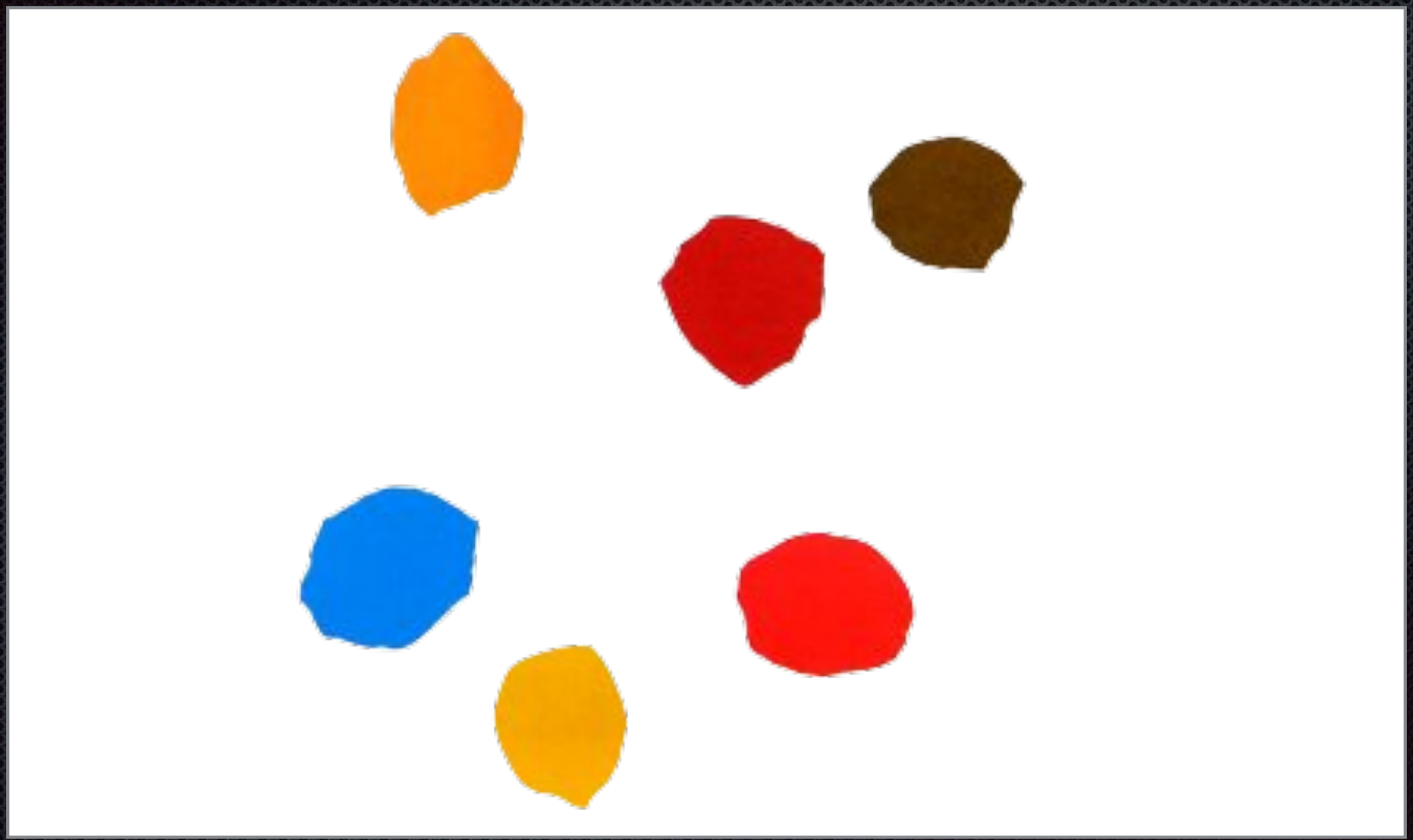


Mais où est Petit Jaune?

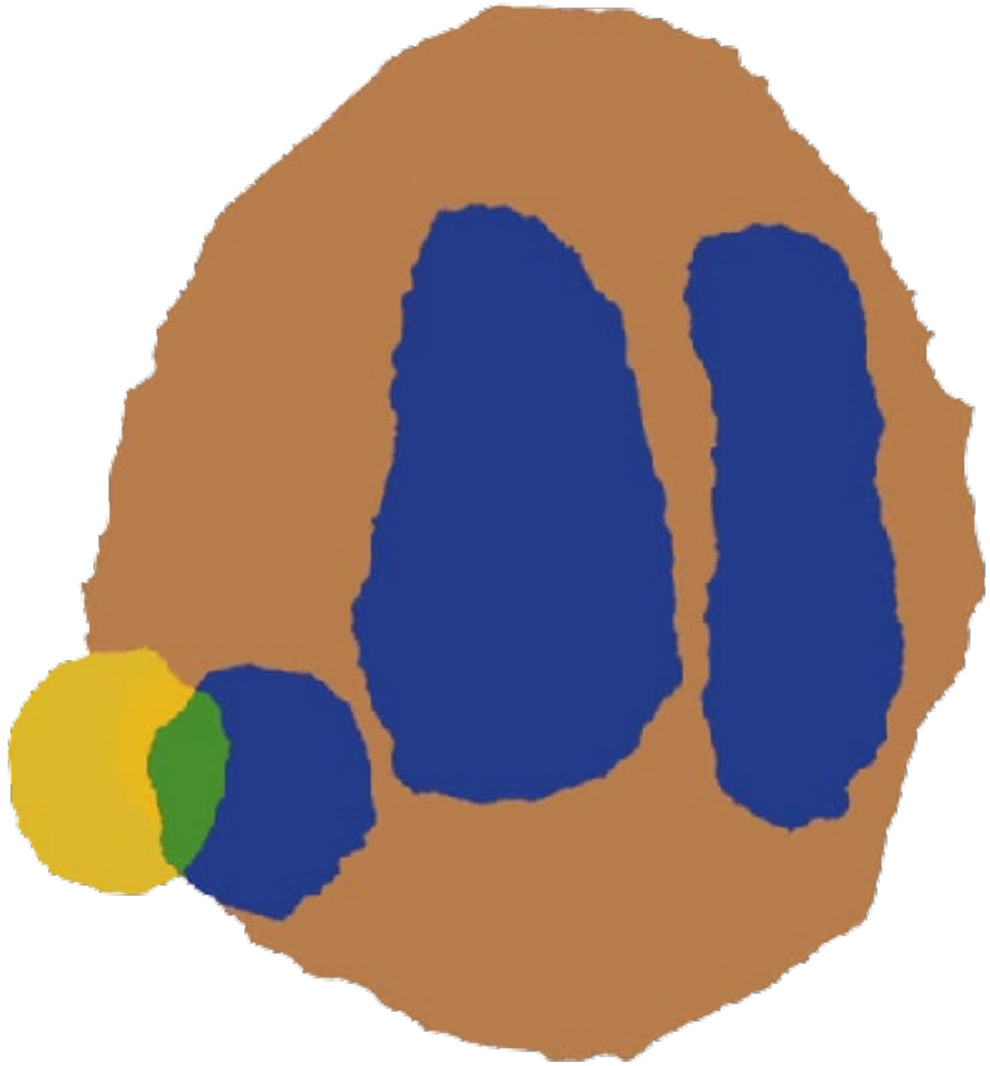
Demander à Papa Bleu et Maman Bleu…



Demander aux copains…



Petit jaune est retrouvé!



Et il y a une grande joie pour tous!

Sommaire

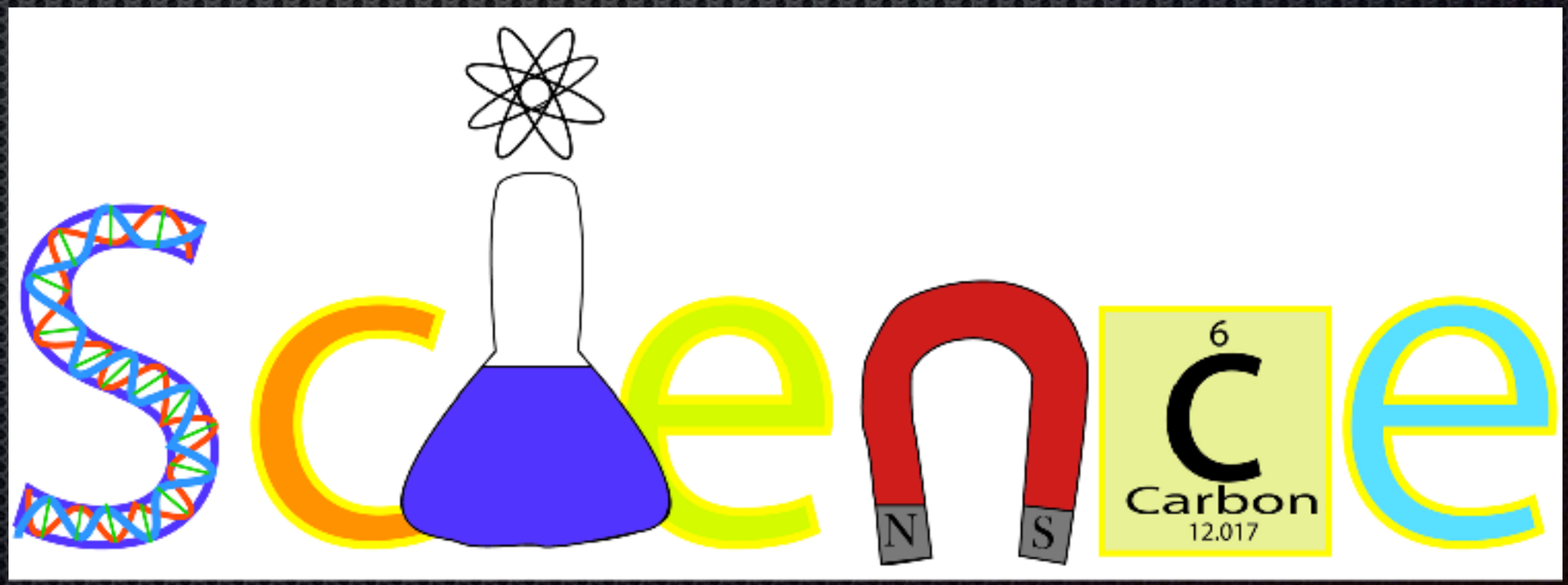
- Science mathématique et science informatique
- Explorer les frontières
  - Structures algèbriques
  - Méthodes numériques
  - Logique et preuves assistées
  - Apprentissage automatique
- Conclusion: un couple d'avenir!

Science

- Observer un objet
- Proposer un modèle
- Prédire un comportement
- Vérifier cette prédiction
- Améliorer le modèle

Erasme: 1466 – 1536
Desiderius Erasmus Roterodamus

Science mathématique
Science informatique



Méthode scientifique

Make Observations
What do I see in nature? This can be from one's own experiences, thoughts, or reading.

Think of Interesting Questions
Why does that pattern occur?

Formulate Hypotheses
What are the general causes of the phenomenon I am wondering about?

Develop Testable Predictions
If my hypotesis is correct, then I expect a, b, c,...

Gather Data to Test Predictions
Relevant data can come from the literature, new observations, or formal experiments. Thorough testing requires replication to verify results.

Develop General Theories
General theories must be consistent with most or all available data and with other current theories.

Refine, Alter, Expand, or Reject Hypotheses

https://en.wikipedia.org/wiki/Scientific_method
Johannes Kepler, 1571 — 1630

Science mathématique

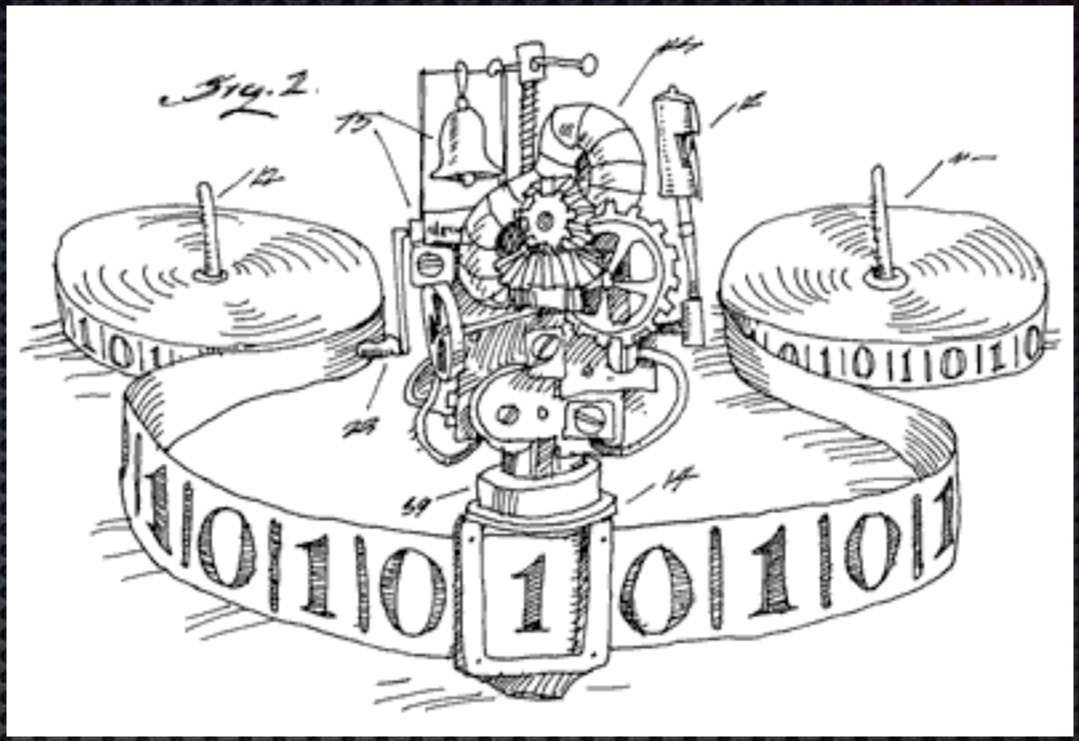
$$\frac{P \Rightarrow Q, P}{Q}$$

- Objets: structures formelles
- Modèle ≈ objet
- Méthode: preuve par inférence
  - Axiomes à choisir: ZFC?
- Prédiction: conjecture
- Vérification: preuve
  - Corollaires, applications
- Amélioration: raffinement de la structure

Science mathématique

- Raisonnement fini
  - Nombre fini de symbole
  - Temps fini
  - Nombre fini de règles d'inférence
- Manipulation implicite d'objet infinis
  - Non définissables explicitement: π
- Pas de notion de coût
- Pas d'exigence de constructibilité
  - Si je peux déduire une contradiction en supposant que non P est vrai, alors j'ai démontré P

Science informatique



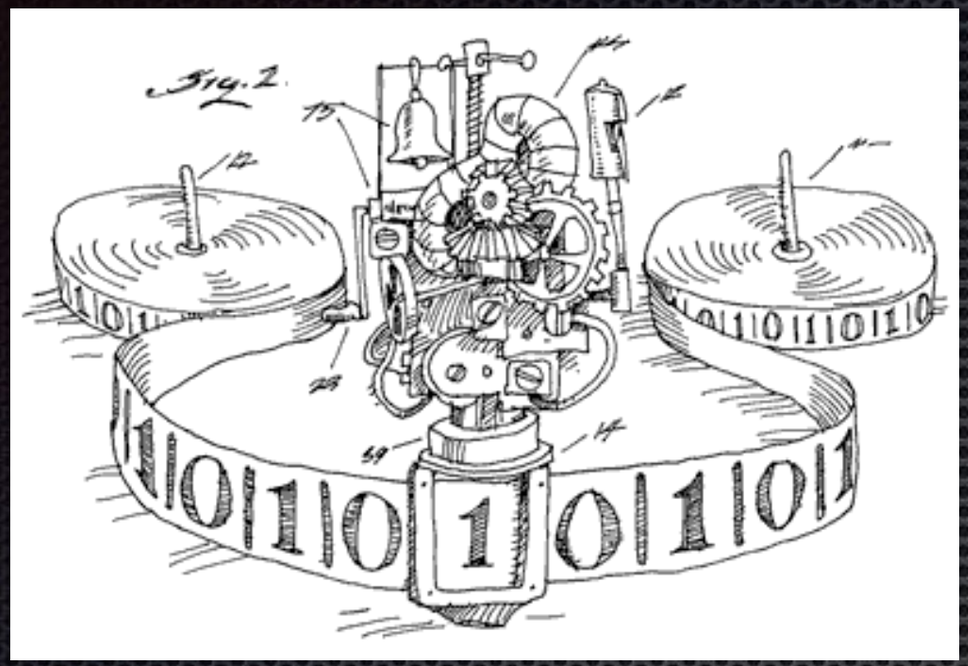
- Objets: opérations formelles
  - Mécanismes digitaux
  - Opérations symboliques
- Modèle: algorithme
- Prédiction: coût
- Vérification: test, preuve
- Amélioration: raffinement de l'algorithme

Alan Turing (1912-1954)

Science informatique

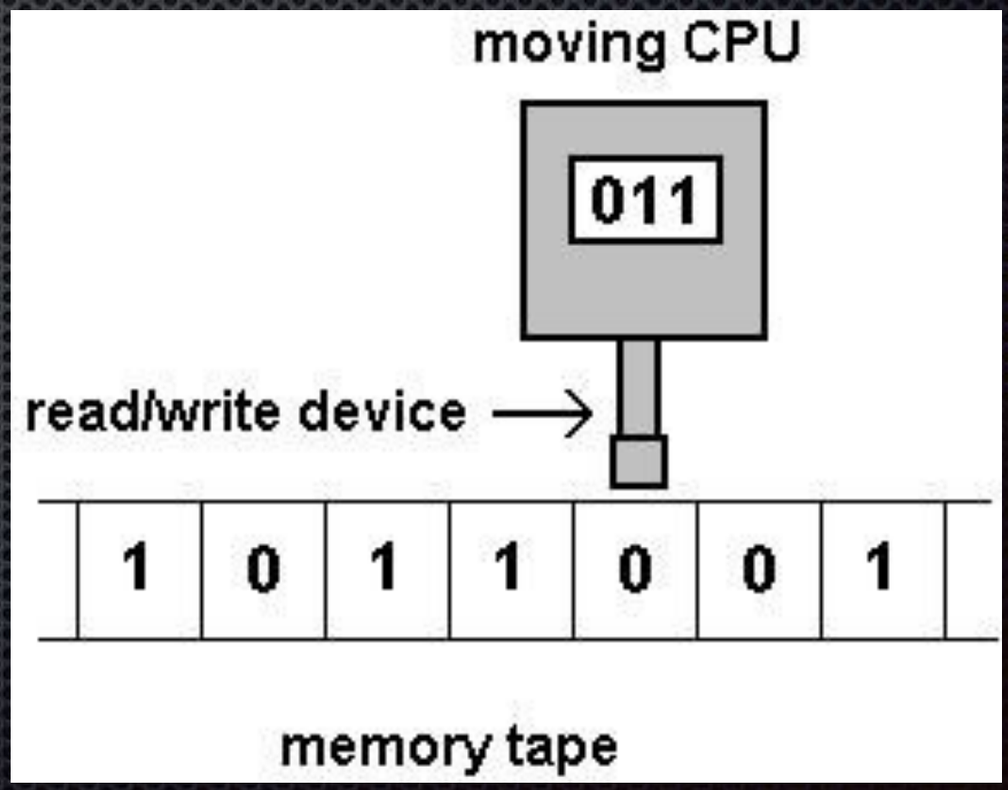
- Seul ce qui est constructible existe
  - Raisonnement sans tiers exclu
- Notion centrale: coût
  - Temps
  - Espace
  - Mise en oeuvre
  - Pire cas, moyenne
- Approches spécifiques
  - Approche stochastique
  - Correction approchée
  - Correction sous oracle

Modèle de référence: Machine de Turing



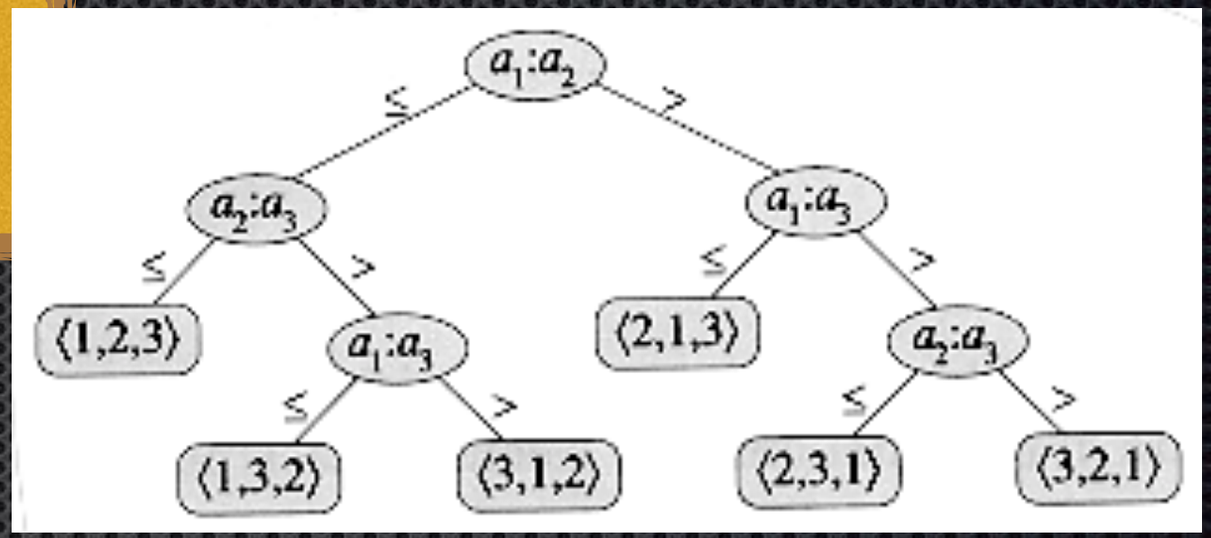
Nombre fini de symboles
Nombre fini d'états
Ruban infini

- Comportement séquentiel
  - Une seule chose à la fois
- Unité de coût
  - Temps: une transition
  - Espace: une cellule



Borne inférieure du tri de n objets

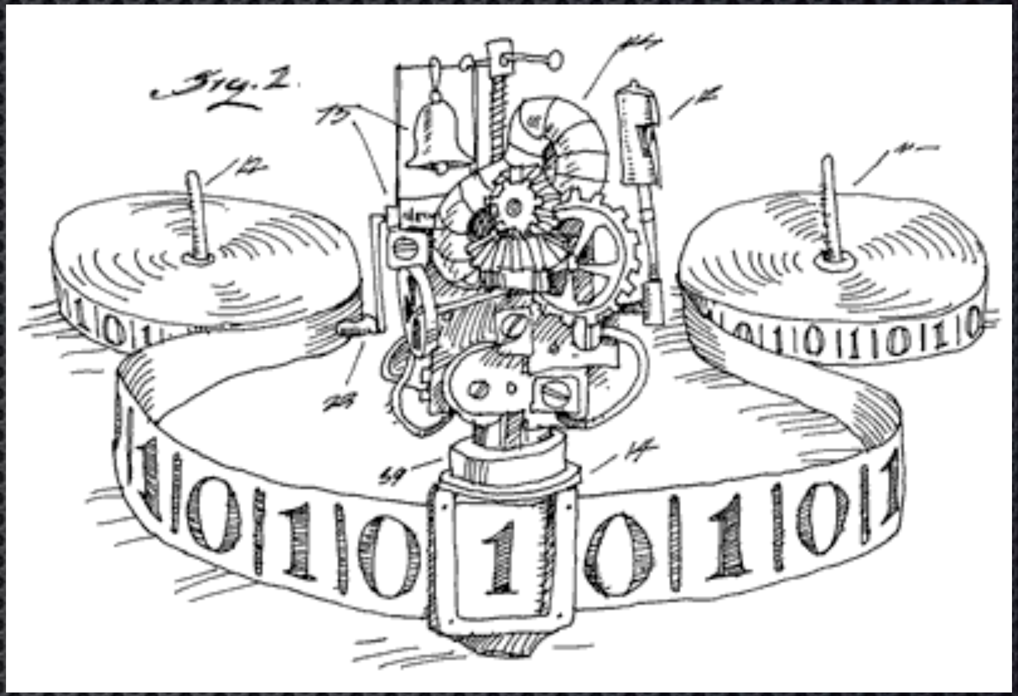
Tout algorithme déterministe fondé sur les comparaisons d'éléments deux à deux doit effectuer Ω(n log n) comparaisons dans le pire cas



- Je considère en entrée une suite de n valeurs distinctes
- Je cherche quelle permutation c'est parmi les n! possibles
- À chaque comparaison, je partage les k permutations candidates en deux parties. Au moins l'une d'elle est de taille supérieure à k/2
- Il y a existe donc une permutation d'entrée qui me demande au moins p comparaisons, avec $n!/2^p \leq 1$
- Donc $\log_2(n!) \leq p$
- Donc $\alpha \times n \times \log_2(n) \leq p$ à pour p grand et une certaine constante α
- $p(n) \in \Omega(n \times \log_2(n))$

Indécidabilité de l'arrêt d'une MT

Il n'existe pas de MT qui puisse décider de manière déterministe de l'arrêt de toutes les MT possibles

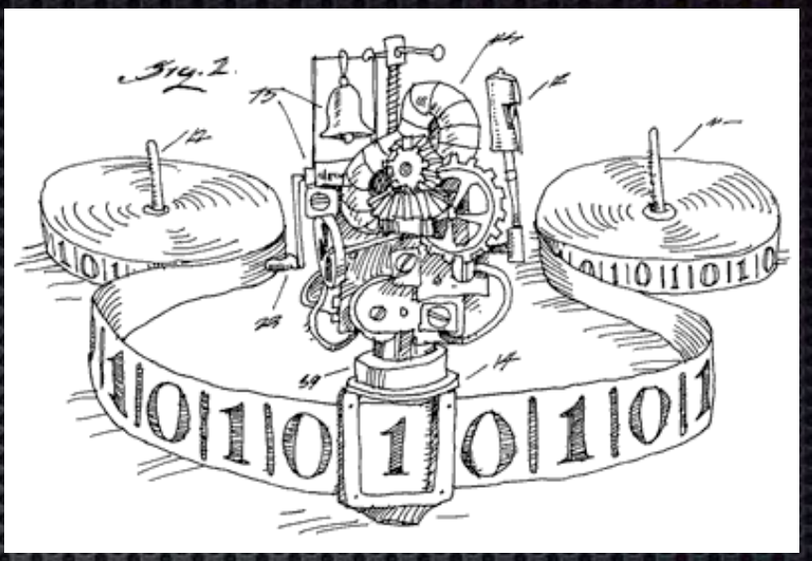


- Il n'existe pas de programme qui puisse répondre à toutes les questions de l'univers
- Toutes les questions non triviales sont indécidables par des machines
- L'homme n'est pas tout-puissant
- Les robots et l'IA le sont encore moins

Thèse de Church-Turing (vers 1950)

Tout procédé (humainement) effectif de calcul peut être simulé par une machine de Turing

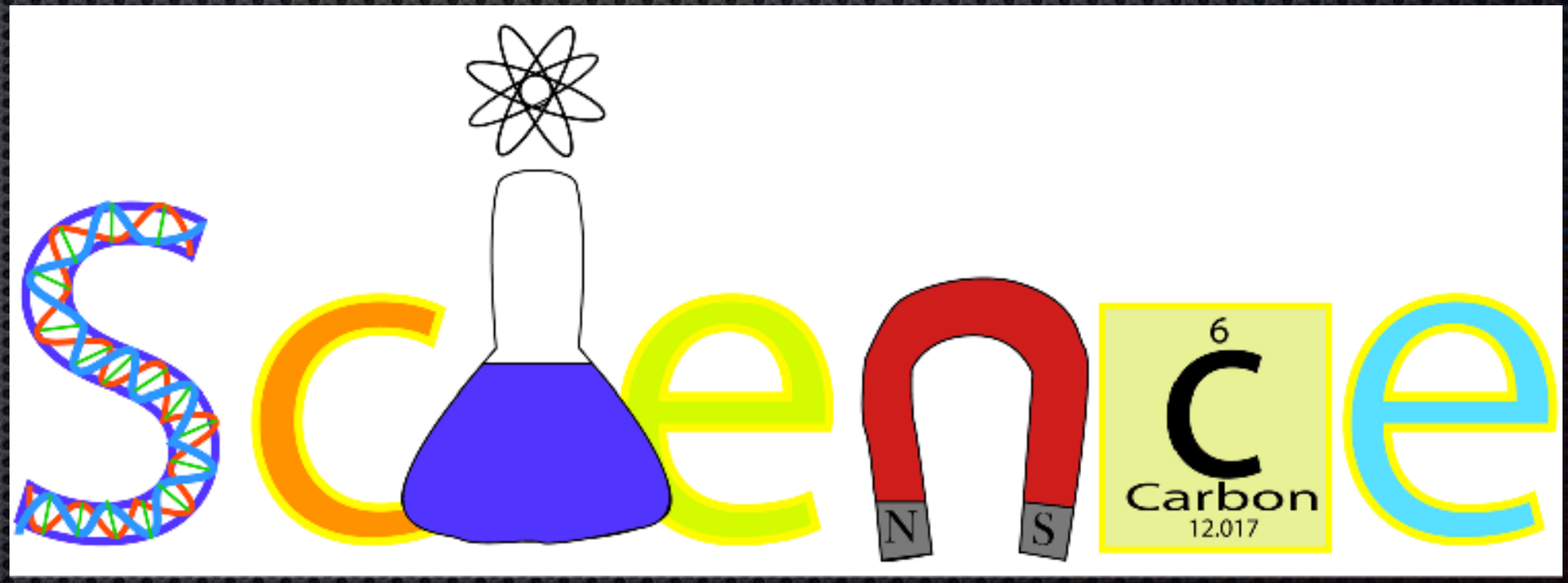
Indécidabilité: une preuve rapide



- Une MT M(s) sur un ruban s est définie par un programme Ⓜ
- Un programme est un texte fini, donc une entrée sur un ruban
- Supposons qu'il existe une MT A(s) qui s'arrête sur toutes les entrées Ⓜ et qui réponde 'oui' si M(Ⓜ) s'arrête et 'non' si M(Ⓜ) ne s'arrête pas
- On construit une machine B(s) telle que B(Ⓜ) ne s'arrête pas si M(Ⓜ) s'arrête et s'arrête si M(Ⓜ) ne s'arrête pas
- Est-ce que B(Ⓑ) s'arrête?

Réification

Explorer les frontières



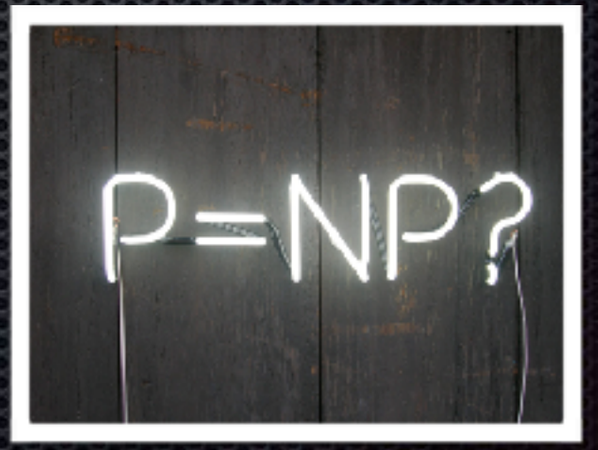
Structures algébriques

- Science mathématique
  - Groupe, anneau, corps, espace vectoriel, espace de Hilbert, etc.
  - Structures riches
  - Question: Propriétés de la structure
    - Groupes quotients
    - Dimensions d'un EV
    - Corps finis
- Science informatique
  - Monoïde (langage), machine de Turing, graphe, logique
  - Structures pauvres
  - Question: Coût de la décision sur les propriétés
    - Reconnaissance d'un langage
    - Arrêt d'une machine de Turing
    - Calcul des composantes connexes d'un graphes

Le problème SAT

3	2		1				3	1
3	2	1	2		3			1
	2	3		1		2	2	1
			1	2				2
		3		3		3		2
1	1	2	3	2	3		3	2
		1	3	1		2	2	
1	1		2		2	1	2	
2		2	2	3			1	3

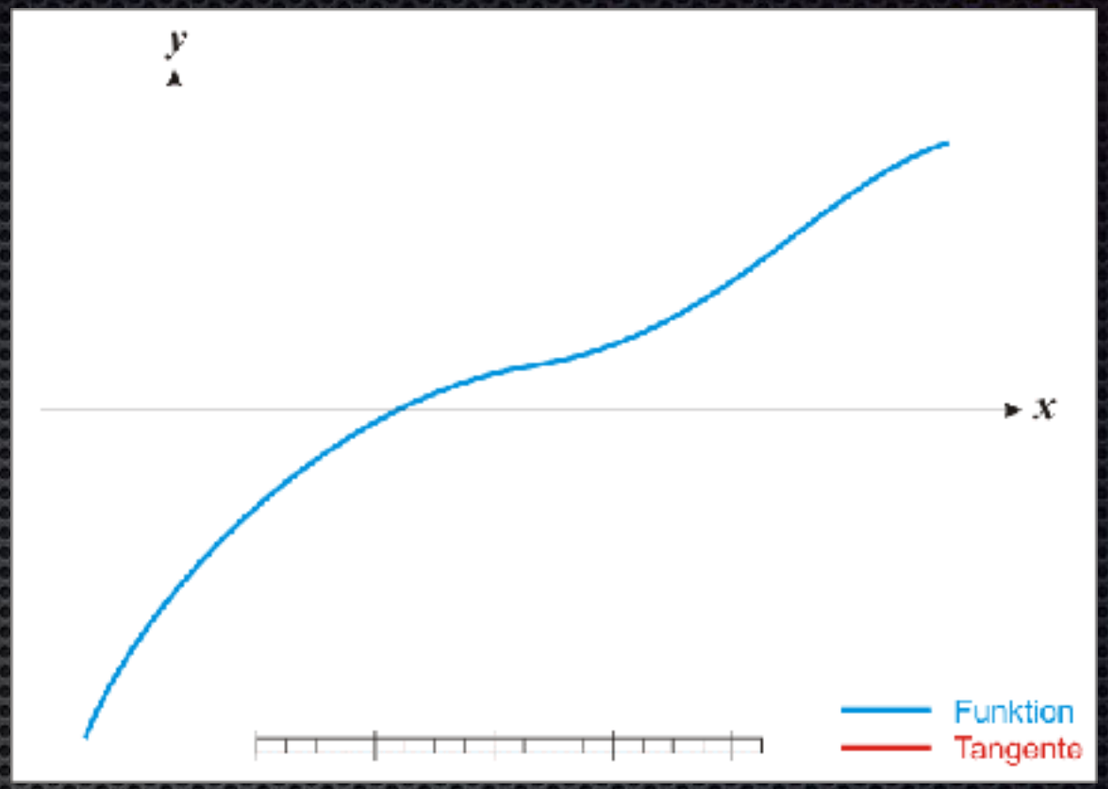
$$(v\,1 \vee \neg\, v\,2) \wedge \neg\, v\,3 \wedge (v\,3 \vee \neg\, v\,1)$$



- SAT1: Trivial
- SAT2: Difficile, mais polynomial
  - Réduction à un problème de coloriage de graphe
- SAT3: NP-complet
  - Au moins aussi difficile que tous les autres problèmes NP-complets
  - S'il existe un algorithme polynomial SAT3 comme pour SAT2, alors beaucoup de choses sur la Terre changeront!
- Sudoku: SAT9

Méthodes numériques: Newton

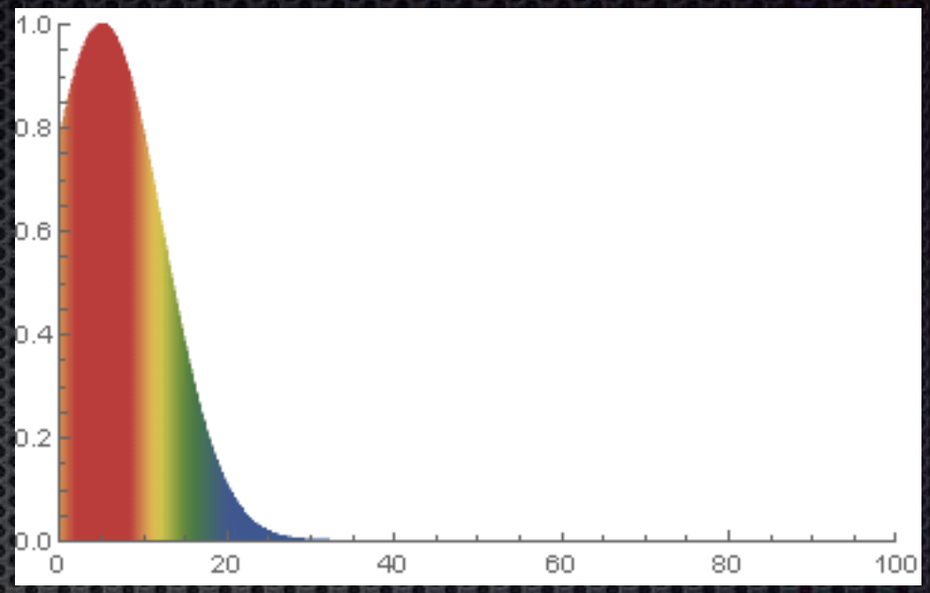
$$x_{n+1} = x_n - f(x_n)/f'(x_n)$$

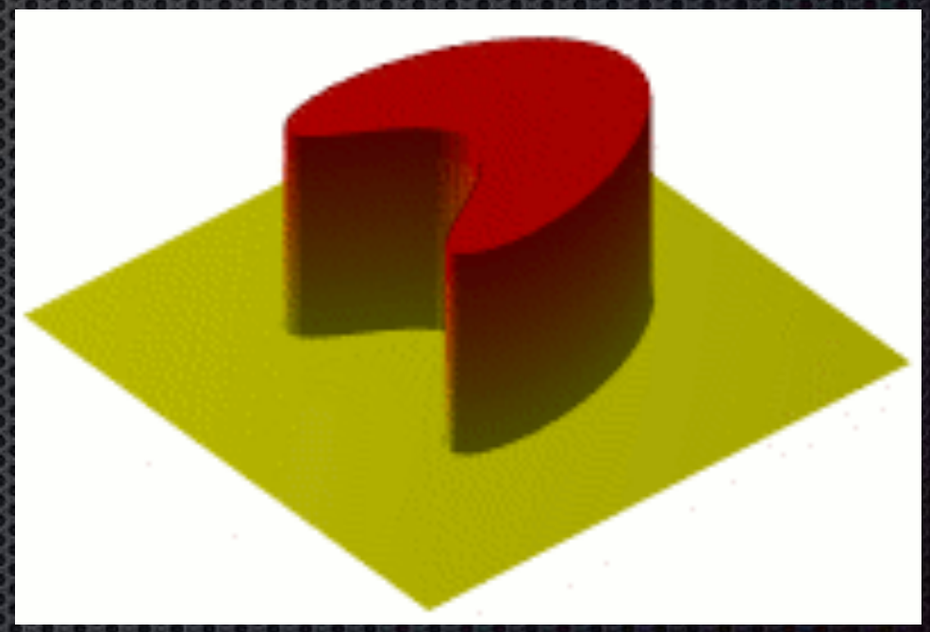


- Approche mathématique
  - La suite des x_n converge vers la racine
- Approche numérique
  - Vitesse de convergence quadratique
  - Stabilité numérique par rapport aux approximations
- Approche informatique
  - Mise en oeuvre sur une architecture
  - Conception d'une architecture adaptée
  - Interface, bibliothèque

Équation différentielles

$$\partial u/\partial t - \alpha \nabla^2 u = 0$$

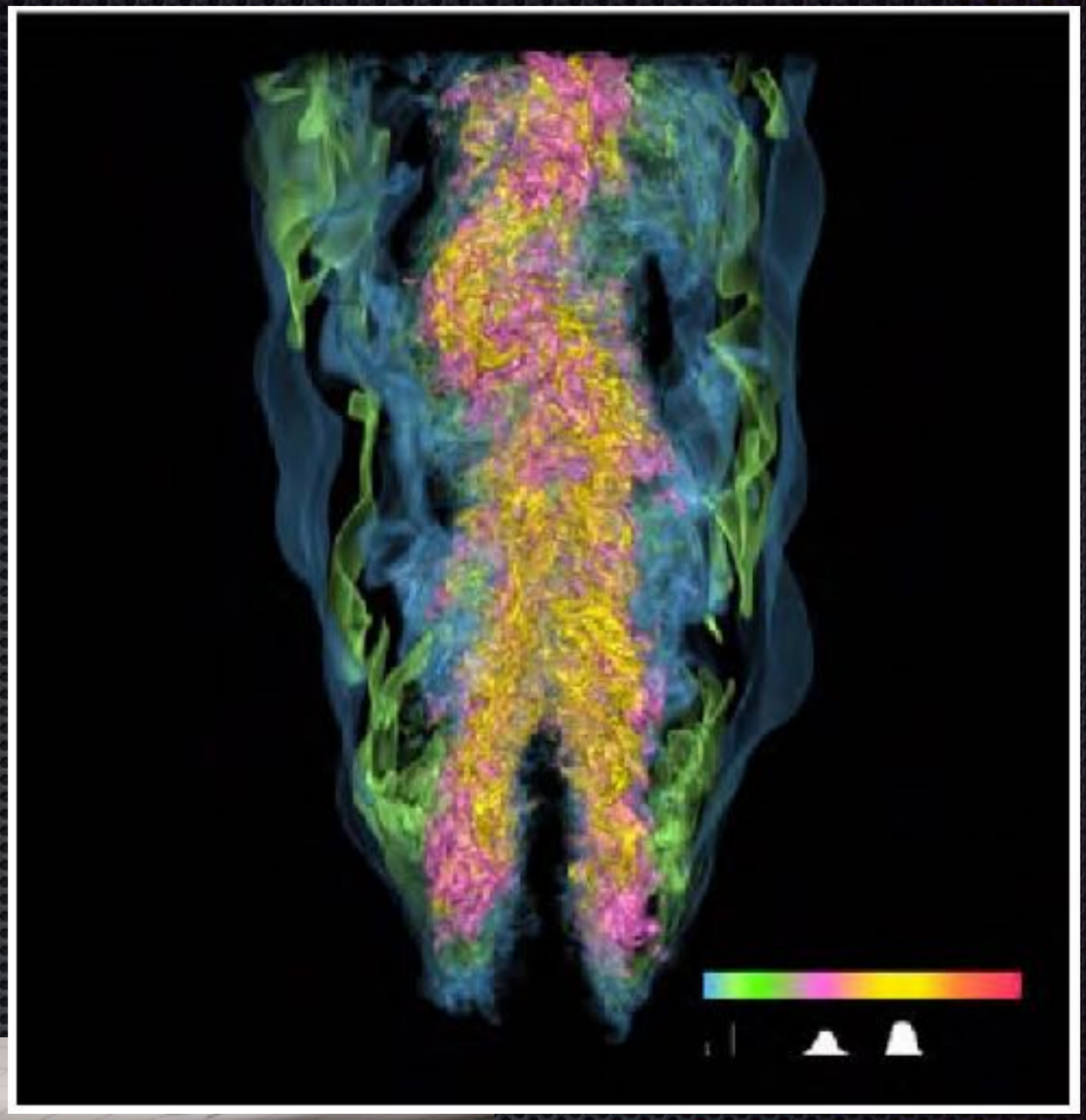




- Discrétiser le modèle: $x_{n+1} = f(x_n)$
  - Temps
  - Espace
  - Étudier la convergence
- Choisir une discrétisation adaptée à l'architecture
  - Processeur
  - Mémoire
  - Distribution
- Mettre en place des outils de visualisation des résultats
- Vérifier la représentativité des calculs par rapport à la théorie mathématique
- Prédire des comportements

$H_2+O_2 = H_2O$

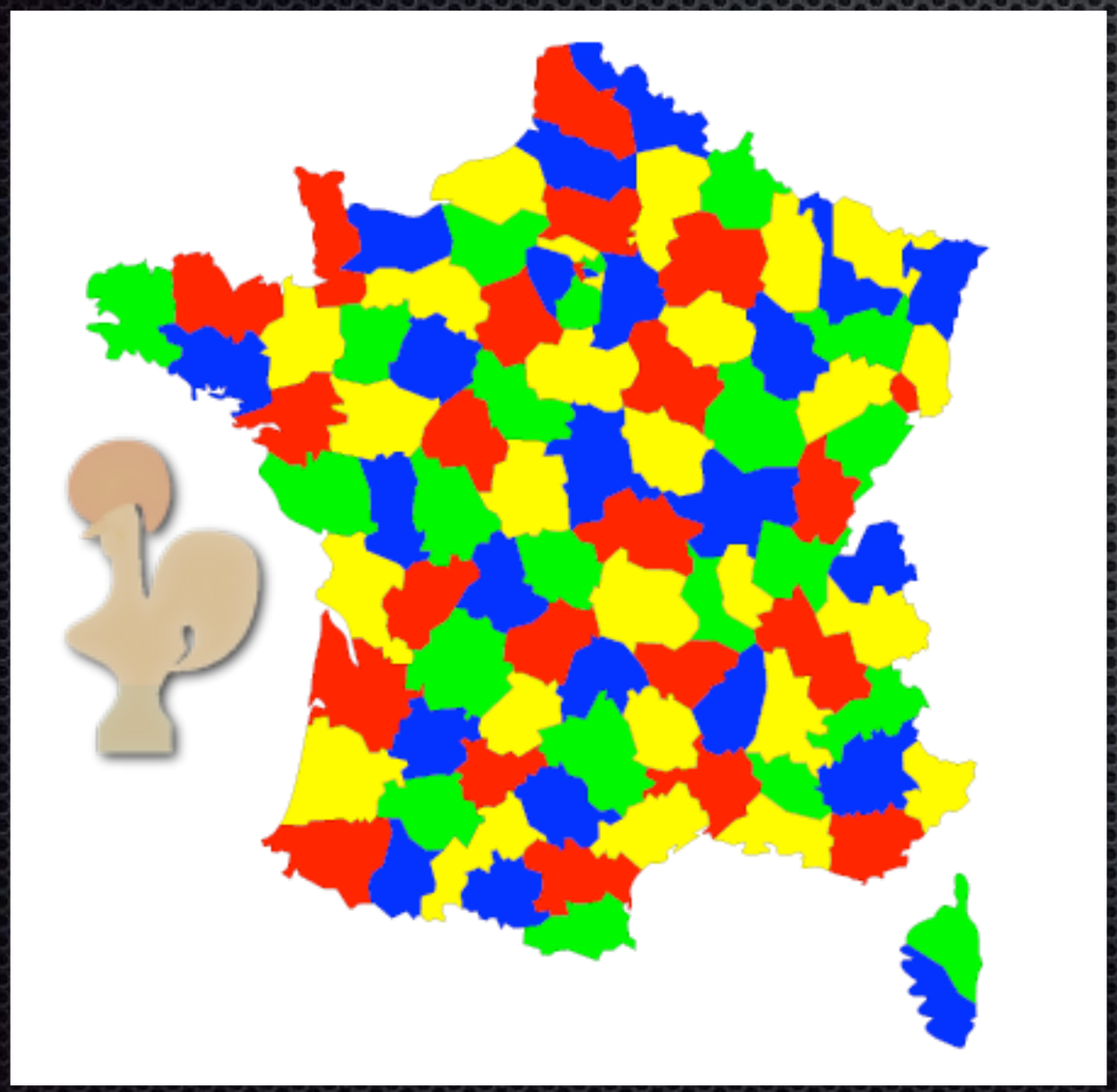
Analyse des turbulences inaccessibles
2,5 millions
d'heures de calcul



#1 en 2012
400 m²
9 MW
100 M$

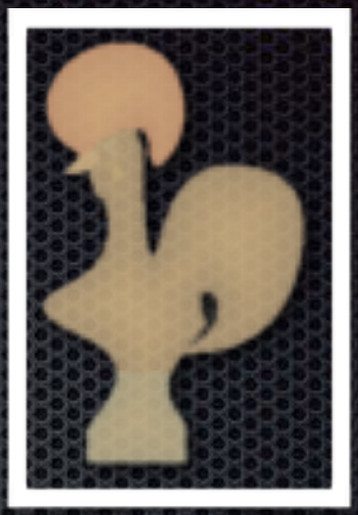
Logique et preuves assistées

Théorème des 4 couleurs



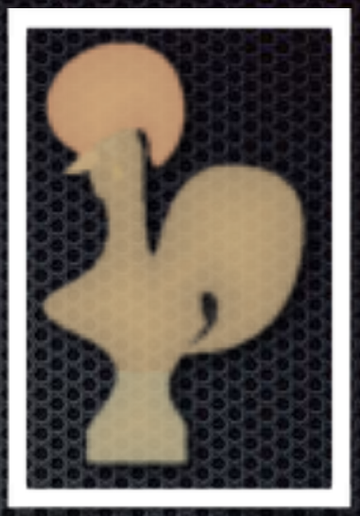
- *1852: Guthrie, énoncé… et nombreuses preuves fausses*
- *1913: Birkhoff, première approche*
- *1969: Heesch, configurations inévitables*
- *1976: Appel et Haken, 1478 configurations, 1200h de calcul*
- *1995: Robertson, Sanders, Seymour et Thomas, 633 configurations*

Coq: vérificateur de preuves



- *Preuve publiée en 2005*
- *2500 lemmes*
- *Environ la moitié triviaux pour un humain*
- *90% font moins de 10 lignes, 40 plus d'une page*
- *Nombres réels intuitionnistes*

- *2012: Preuve du théorème de Feit-Thompson*
- *Tout groupe fini d'ordre impair est résoluble*
- *6 ans de travail*
- *Preuve de 250 pages écrite en 1963*
- *Grand Prix "Sciences de l'informatique" de la Fondation d'entreprise EADS en 2011*



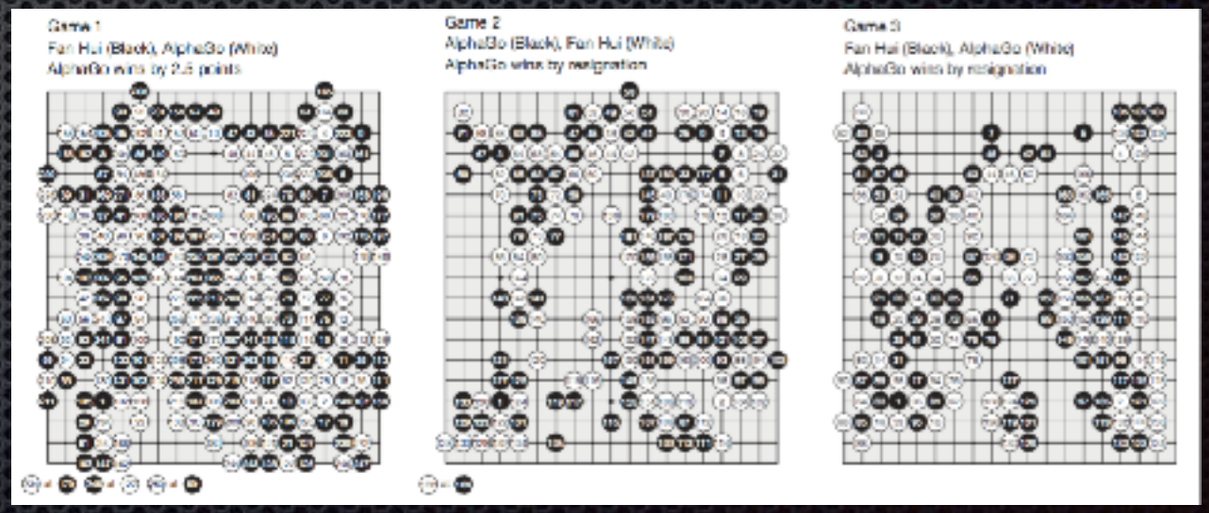
```
Coq < Goal (A -> B -> C) -> (A -> B) -> A -> C.
1 subgoal

  A : Prop
  B : Prop
  C : Prop
  ============================
   (A -> B -> C) -> (A -> B) -> A -> C
```

```
Coq < intros H' HA.
1 subgoal

  A : Prop
  B : Prop
  C : Prop
  H : A -> B -> C
  H' : A -> B
  HA : A
  ============================
   C
```

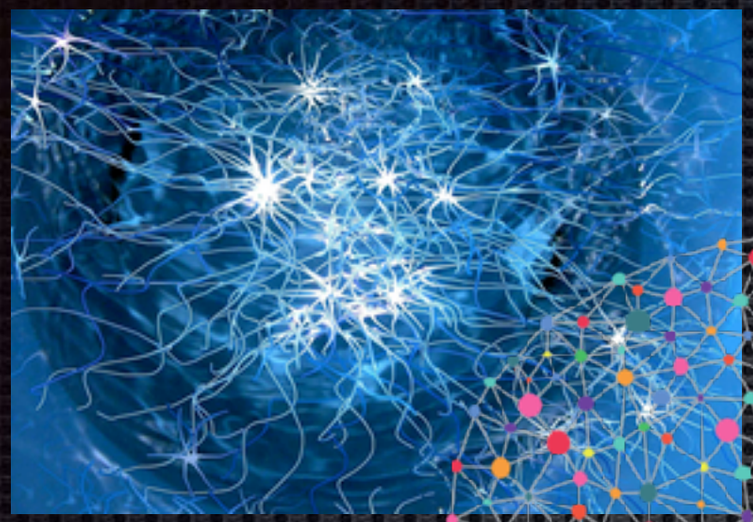
```
Coq < apply H.
2 subgoals

  A : Prop
  B : Prop
  C : Prop
  H : A -> B -> C
  H' : A -> B
  HA : A
  ============================
   A
subgoal 2 is:
 B
```

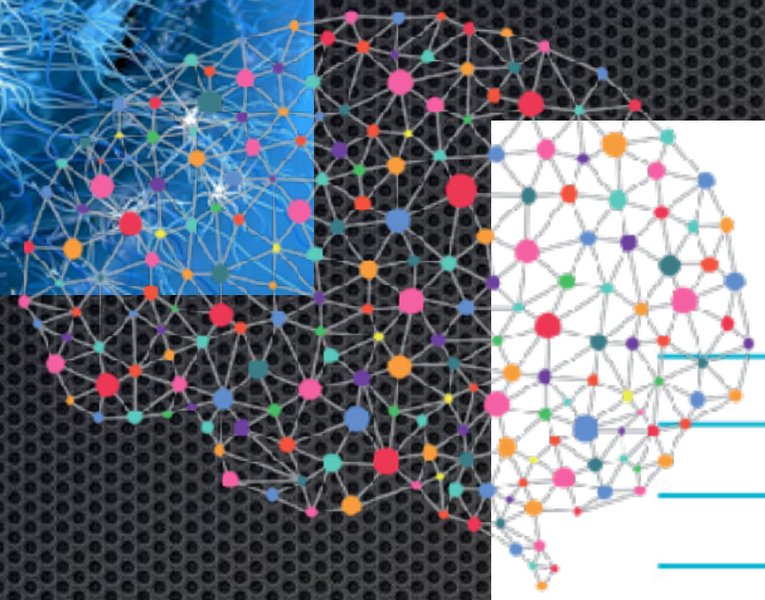
```
Coq < assumption.
Proof completed.
```

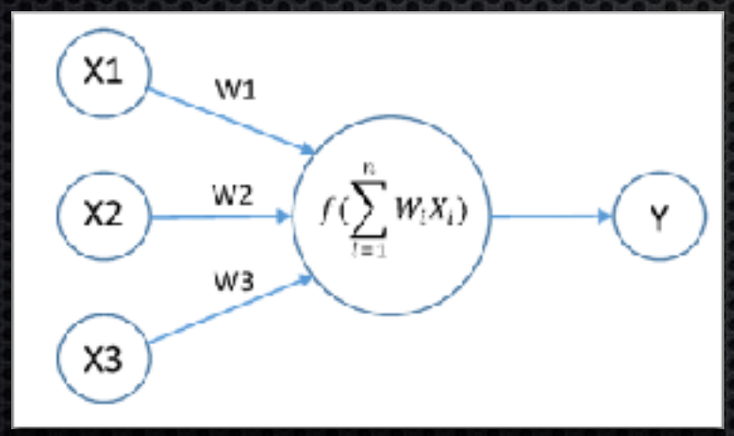
Apprentissage automatique

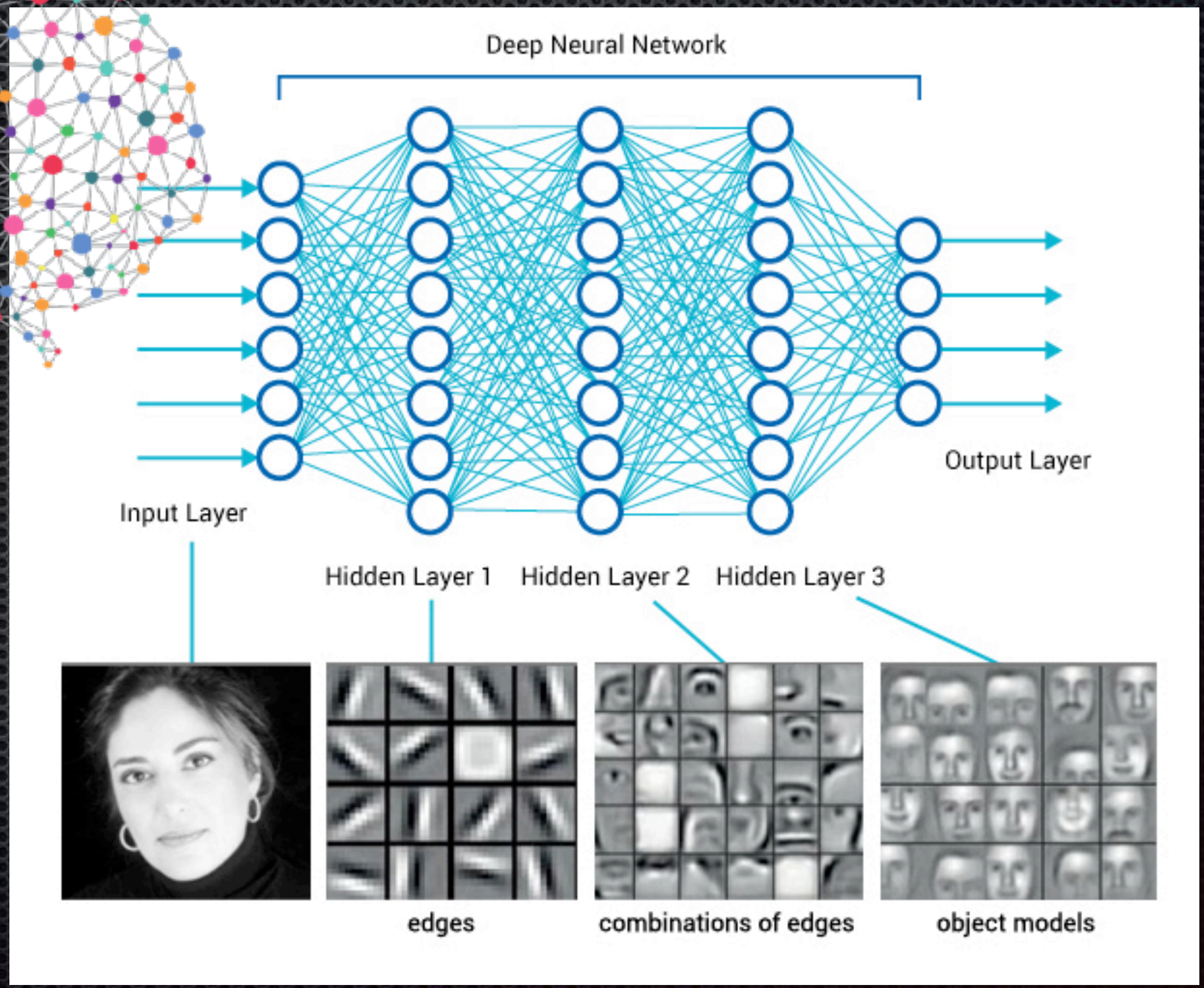


Deep learning

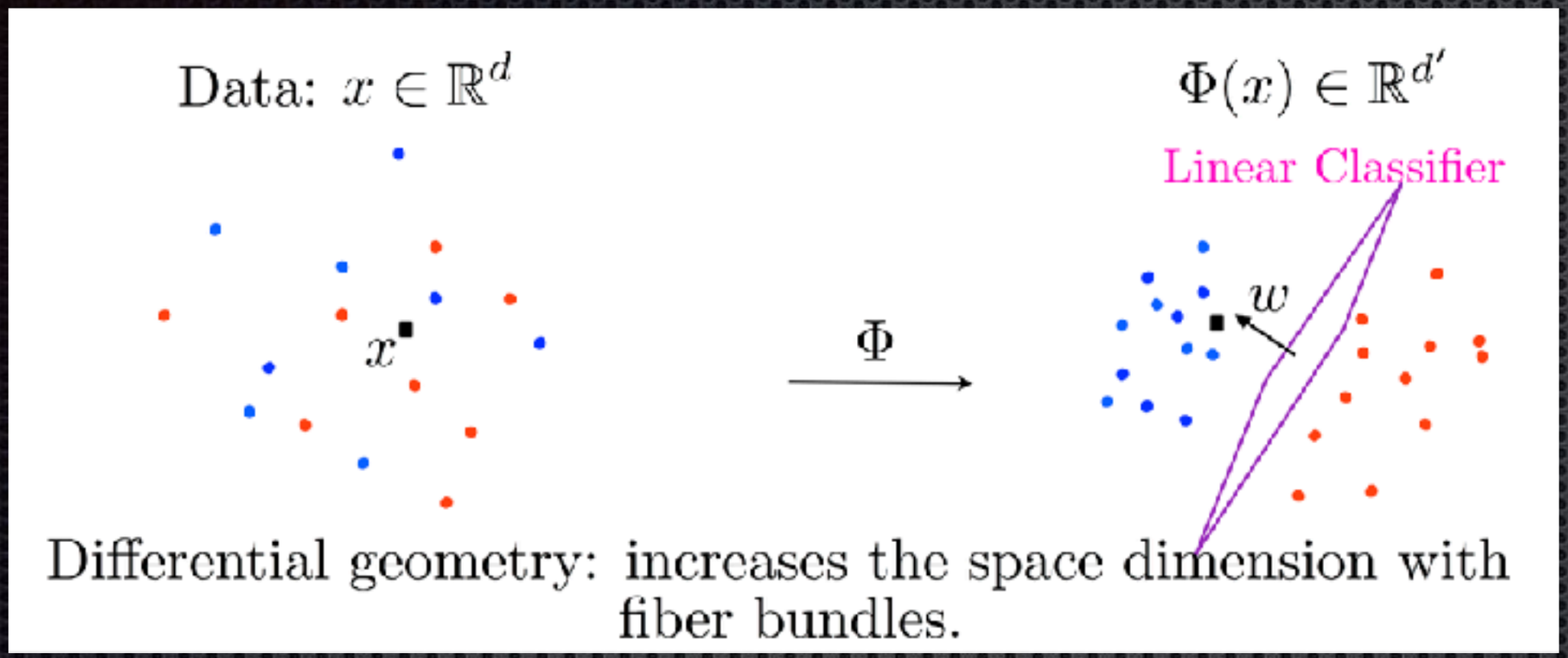






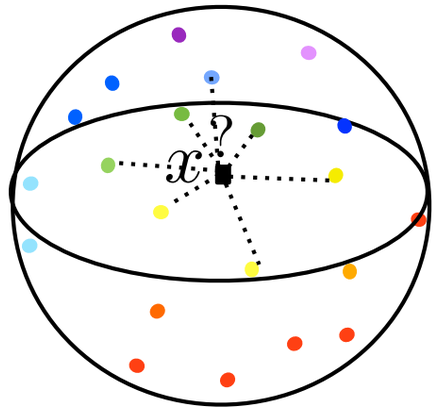


Approche classique: Classifieur et apprentissage



Stéphane Mallat, ENS

La malédiction de la dimension!

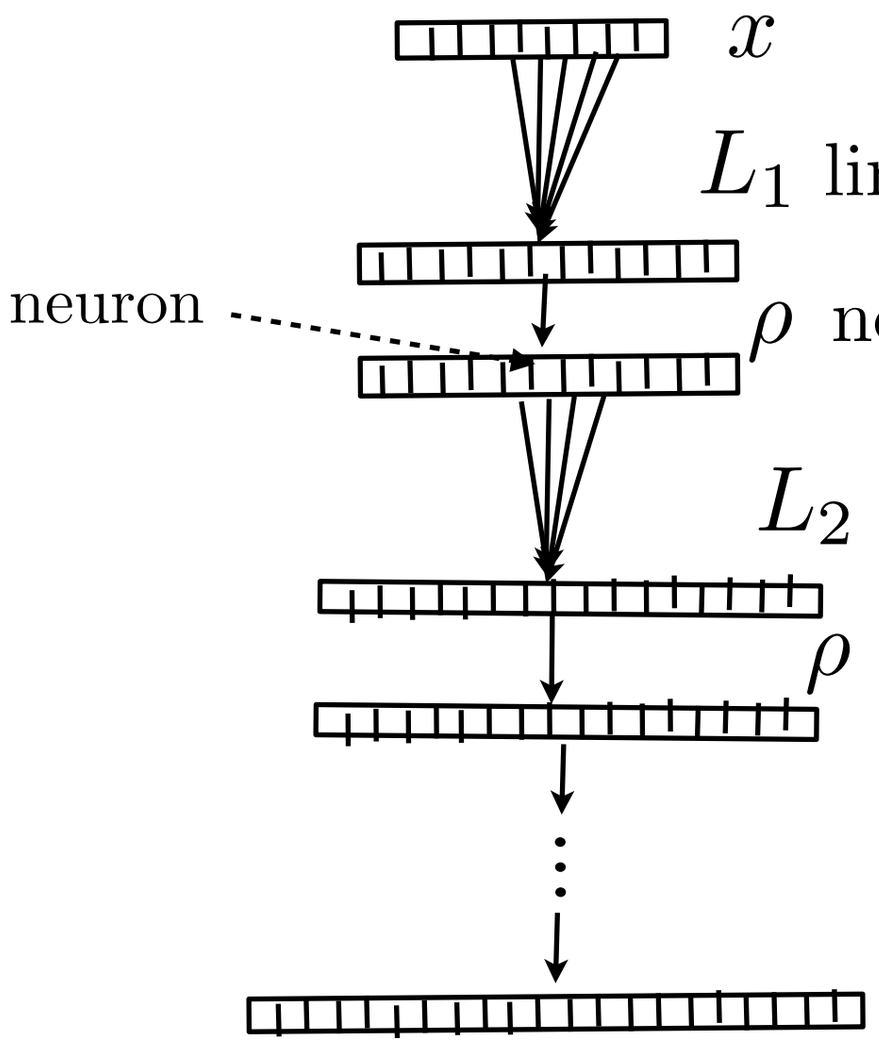


Need ϵ^{-d} points to cover $[0,1]^d$ at a Euclidean distance ϵ

$\Rightarrow \|x - x_i\|$ is always large

Deep Convolution Neworks

- The revival of an old (1950) idea: *Y. LeCun, G. Hinton*



Optimize the L_k with support constraints: over 10^9 parameters

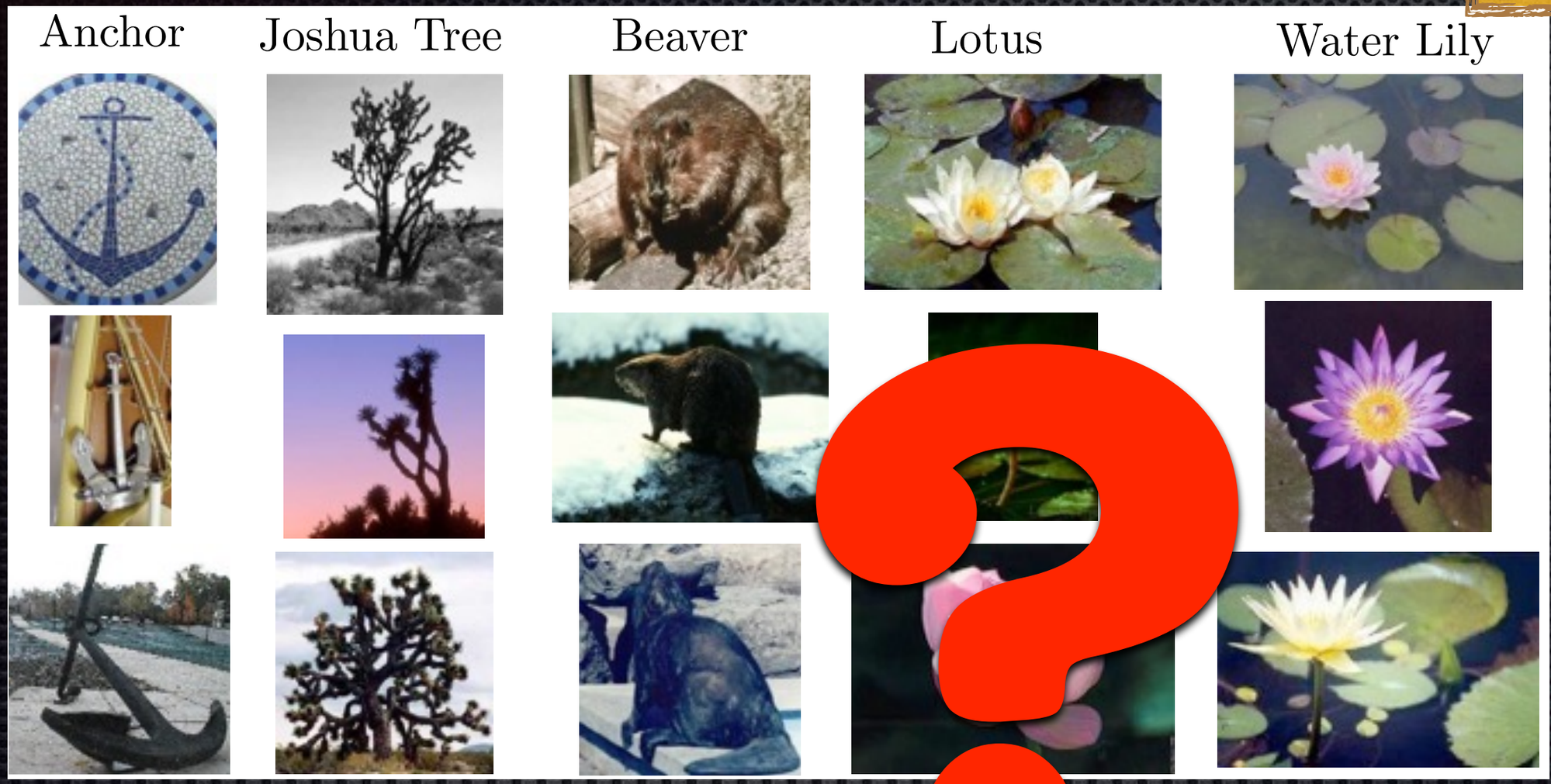
Exceptional results for *images, speech, bio-data* classification.

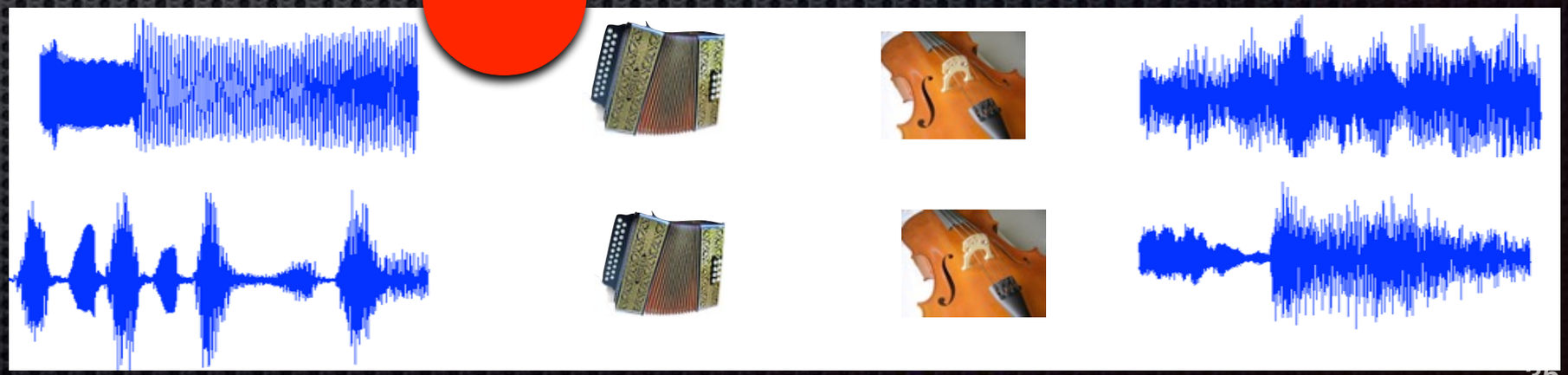
Products by FaceBook, IBM, Google, Microsoft, Yahoo...

Why does it work so well ?

Un mystère mathématique

Turbo-codes
Optimisation par recuit simulé





Un couple d'avenir!

Les mathématiques: Un cadre essentiel pour l'informatique

- Informatique: science du traitement algorithmique de l'information
  - Algèbre: structures discrètes, combinatoire, théorie des graphes
  - Analyse: étude de complexité, modèles continus, stabilité numérique
  - Probabilités: modèles stochastiques, étude en moyenne, algorithmes d'approximation
  - Logique: preuve assistée, vérification de correction
- Des champs parfois peu étudiés traditionnellement
  - Algèbre faible
  - Logique
  - Approche constructive, intuitionniste

L'informatique: Une chance pour les mathématiques

- Cadre conceptuel renouvelé
  - Étude du coût de construction des objet
  - Différence entre syntaxe et sémantique
  - Critère applicatif
- Domaine très ouvert, en forte évolution
  - Objet relativement simples à appréhender
  - Nombreuses possibilités de médiation pédagogique
- Renouvellement des méta-mathématiques
  - Une preuve est un objet mathématique comme les autres!
  - Notion de réification: une machine est un programme

La science informatique
Une science dans grand le réseau des sciences

- Mathématiques
- Sciences de l'ingénieur
  - Robotique
  - Gestion de l'énergie
  - Simulation pour l'ingénierie
- Physique
  - Calcul quantique, théorie de l'information
  - Théorie du signal, codage
  - Simulation à grande échelle
- Biologie
  - Bioinformatique, génomique
  - Réseaux de gènes
  - Simulation de dynamique moléculaire
- Économie, finance, management
- Linguistique, littérature, sociologie, pédagogie, etc.

Merci de votre attention!