Informatique
Amphi 0

Jean-Jacques.Levy@inria.fr

http://jeanjacqueslevy.net

secrétariat de l'enseignement:

Catherine Bensoussan

cb@lix.polytechnique.fr

Aile 00, LIX

tel: 34 67

http://w3.edu.polytechnique.fr/informatique/

Plan

1. Calculabilité
2. Complexité
3. Algorithmes simples
4. Programmation
5. Vaste domaine scientifique
6. Organisation des cours
7. Contenu du Tronc Commun

Un système étiqueté – *Tag system*

Un mot de départ $w = abccba$

Un nombre $n = 3$

Des règles $a \mapsto bab \qquad b \mapsto abba \qquad c \mapsto ca$

$abccba$
$abccbabab$
$\quad cbabab$

$\quad cbabab$
$\quad cbababca$
$\qquad babca$

$\qquad babca$
$\qquad babcaabba$
$\qquad\quad caabba$

$caabba$
$caabbaca$
$\quad bbaca$

$\quad bbaca$
$\quad bbacaabba$
$\qquad caabba$ **(arrêt car déja vu)**

Un autre petit jeu

Le mot de départ $w = cabcba$

Le nombre $n = 3$

Les règles $a \mapsto cac \qquad b \mapsto bab \qquad c \mapsto ca$

$cabcba$

$cbaca$

$caca$

aca

cac

ca

a

c

(arrêt car mot vide)

Un troisième jeu

Le mot de départ	$w = 100110$	
Le nombre	$n = 3$	
Les règles	$0 \mapsto 00$	$1 \mapsto 1101$

100110	1011101110100	
1101101	11011101001101	
11011101	111010011011101	
111011101	0100110111011101	
0111011101	011011101110100	
101110100	01110111010000	
1101001101	1011101000000	
10011011101	11010000001101	
110111011101	100000011011101	
1110111011101	0000110111011101	
01110111011101	011011101110100	(arrêt car déja vu)

Terminaison

- le jeu 1 peut ne pas terminer
- le jeu 2 termine toujours
- la terminaison du jeu 3 est un problème ouvert

Question

Peut-on déterminer par programme la terminaison d'un jeu donné par w, n et $\{a_1 \mapsto A_1, a_2 \mapsto A_2, \ldots a_p \mapsto A_p\} \quad (n \geq 0, p \geq 0)$??

Réponse [Post, 1925]

Non. La terminaison des systèmes étiquetés est indécidable.

Dominos de Wang

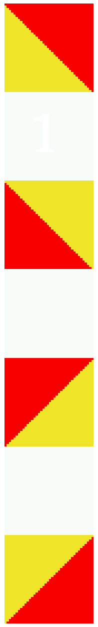
Un ensemble

de $p = 4$ motifs

(carrés à bords colorés)

Paver un carré

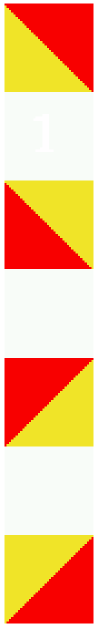
de taille $n = 6$

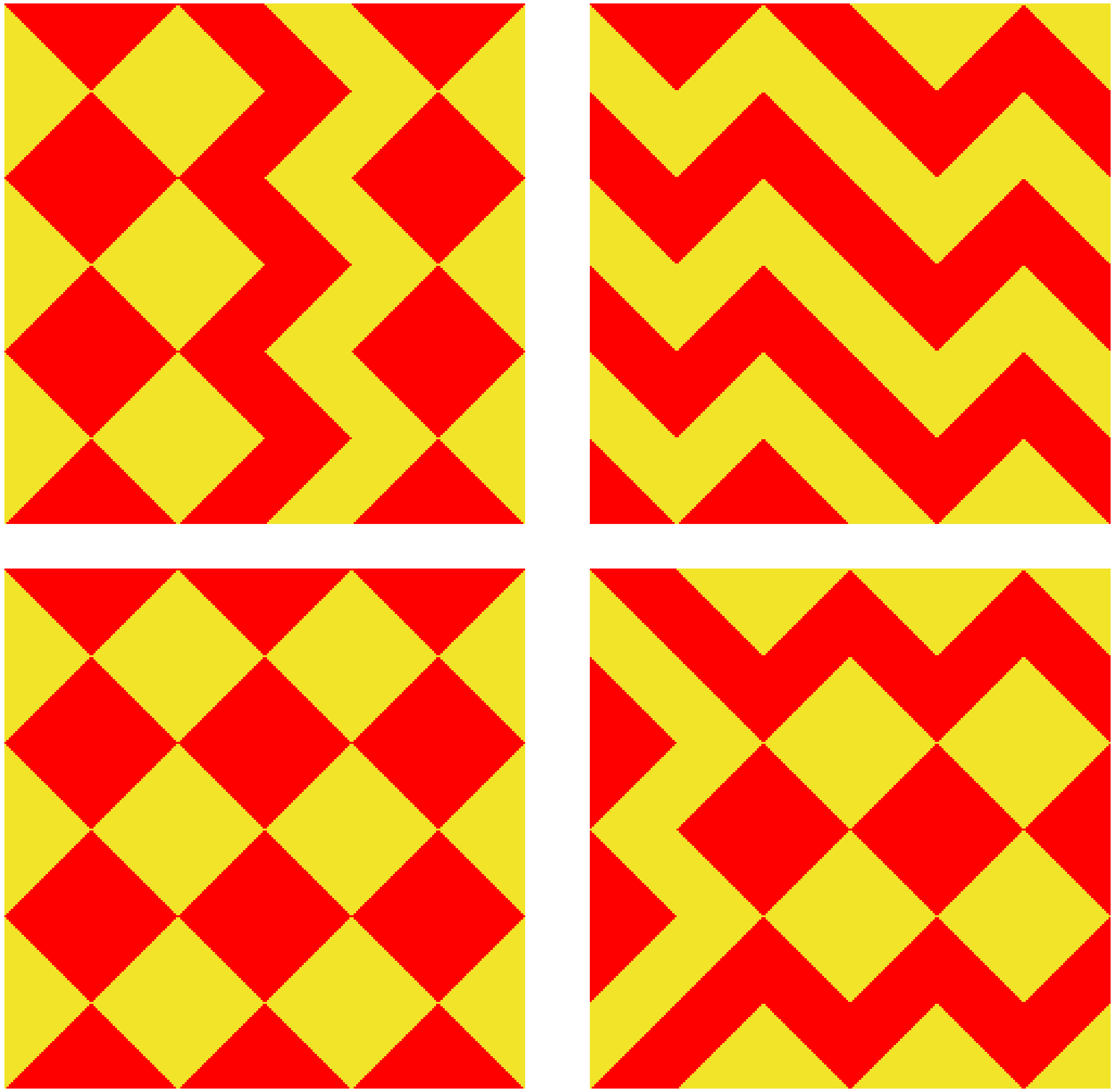


Deux motifs adjacents ssi leurs cotés communs de même couleur.

Rotations et symétries sont interdites.

Dominos de Wang – bis

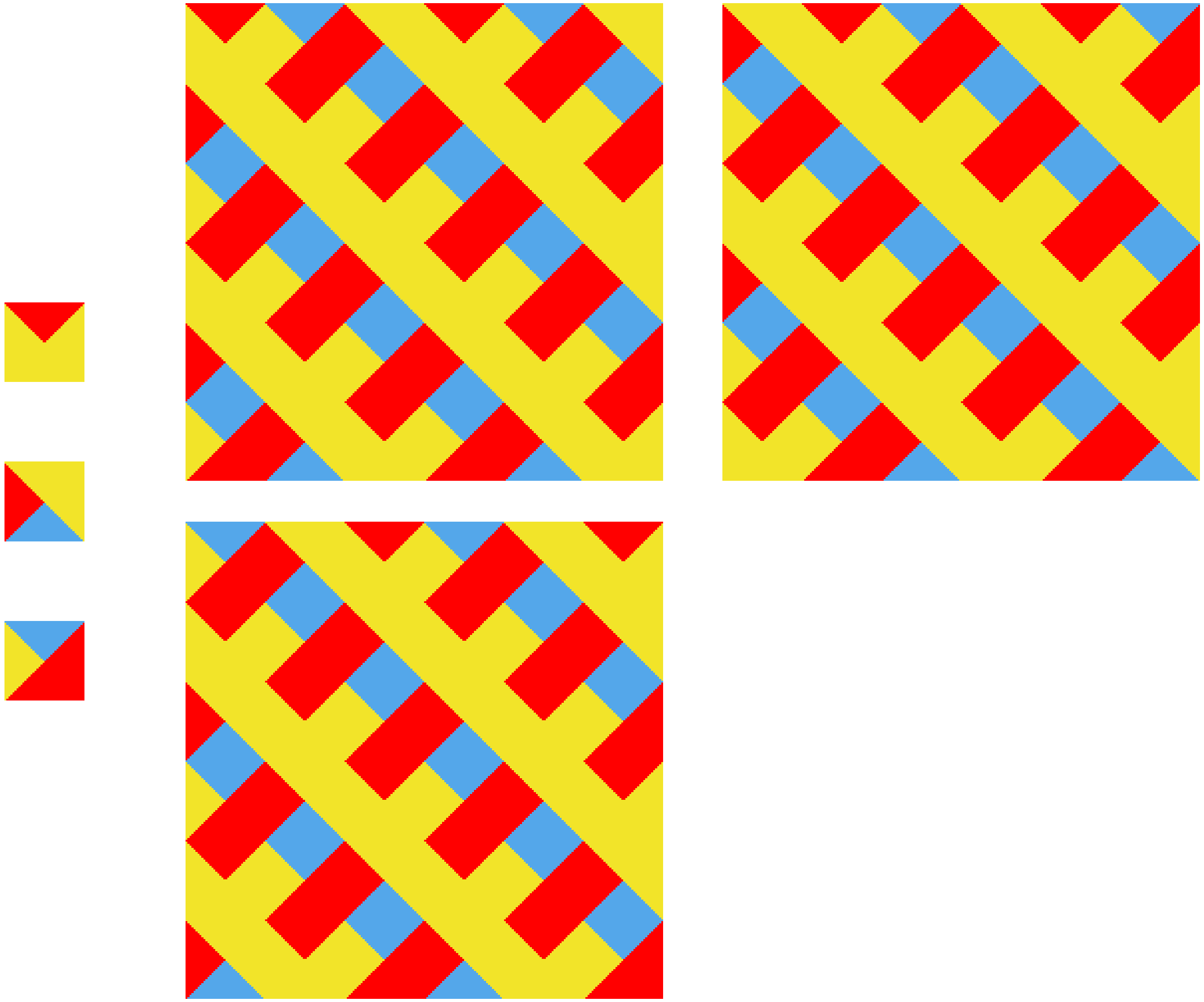




4096

solutions

Au dela du jaune et du rouge



3

solutions

Pavage d'un carré

- **Données:** p motifs et un carré de taille n.
- **Algorithme:** énumérer les p^{n^2} pavages possibles et ne garder que les pavages licites.
- **Complexité:** $O(p^{n^2})$

Dans l'exemple précédent

$n = 6,\ p = 3$

$150094635296999121 \simeq 10^{17}$ tests à effectuer

Soit 57 **jours** en comptant 500×10^6 opérations à la seconde.

Faire mieux?

NP complétude

Personne ne sait faire mieux.

Ce problème est représentatif d'une grande classe de problèmes, dits NP **complets**.

Si on sait faire mieux pour ce problème, on saura le faire pour toute la classe NP.

Conjecture: $P \stackrel{?}{\neq} NP$ **[Cook, 1973]**

Un des 10 problèmes ouverts les plus importants des mathématiques

(cf. Majeure 2 d'Informatique)

Pavage du plan

- **Données:** p motifs
- **Problème:** paver tout le plan

équivalent à

- **Données:** p motifs
- **Problème:** paver les carrés de taille n pour **tout** $n \geq 0$

Réponse

Le pavage du plan par p motifs est **indécidable**

Pavages non périodiques

Conjecture de Wang (1973): tout pavage du plan est périodique.
Et grâce à cette périodicité, on sait décider du pavage du plan.

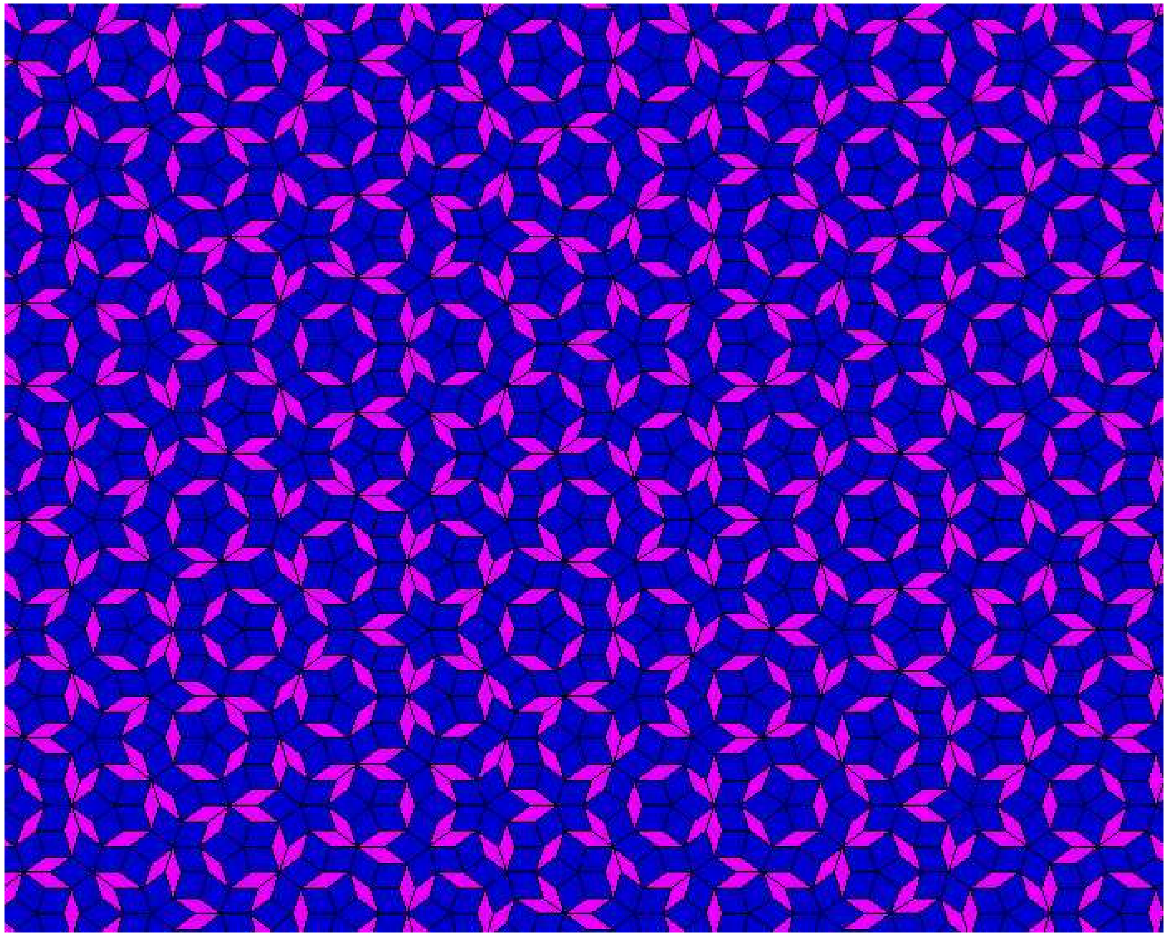
Indécidabilité $\Rightarrow$ existence de pavages non périodiques.

Conclusion: en passant par des théorèmes d'informatique théorique, on peut résoudre des problèmes de math.

(Itou pour la physique)

En 1975, les pavages non périodiques ont fleuri. Pavages quasi périodiques de Penrose.

Exemple de pavage de Penrose



Théorie de la calculabilité

Comment montrer qu'un problème est indécidable?

Par contradiction: si le problème est décidable, un problème indécidable connu deviendrait décidable.

Comment montrer l'indécidabilité de ces problèmes de base?

Par diagonalisation et reflexivité.

Exemple connexe: Soit S l'ensemble de tous les ensembles. Alors $S \in S$. Mais pour l'ensemble $X = \{T \mid T \notin T\}$. Alors $X \in X$ ssi $X \notin X$.

[Gödel, Church, Turing, Kleene, Post; 1920-1940]
[von Neumann; 1940]

Faits

Il existe des

- problèmes indécidables
- problèmes décidables de complexité élevée

En pratique

- problèmes ou sous-cas de faible complexité algorithmique, typiquement en $O(n^\alpha)$ avec $\alpha \leq 3$
- suffisant pour construire les cathédrales informatiques des années 2000

La décidabilité est étudiée en Majeure 1.

La complexité des algorithmes en Tronc Commun et Majeure 2.

Exemples d'algorithmes “simples”

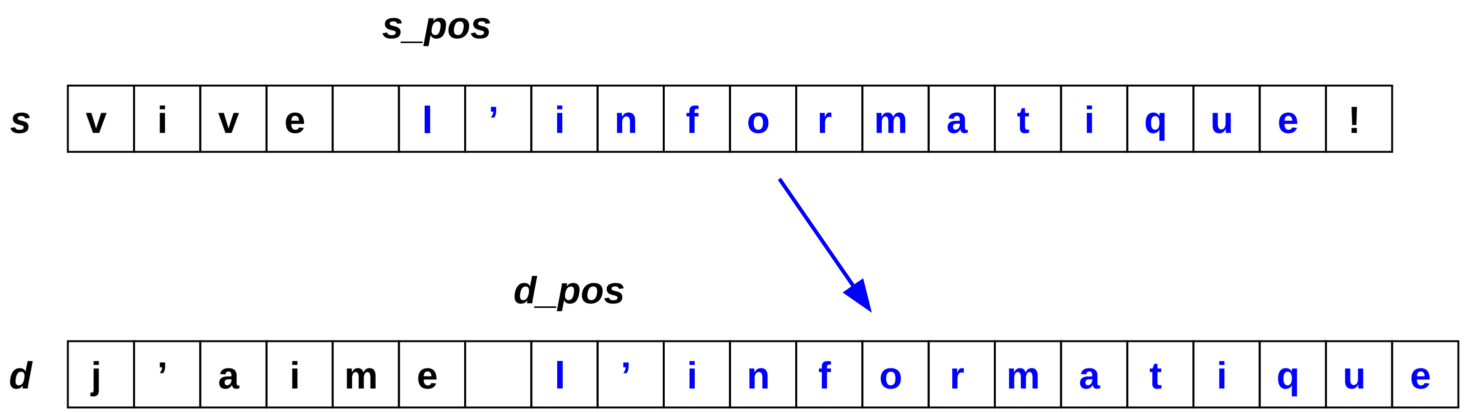
- tri
- recherche en table
- gestion d'une file de priorité
- analyse syntaxique
- évaluation des expressions arithmétiques
- marche du cavalier sur un échiquier (algorithme glouton)
- parcours en profondeur d'abord (connexité dans un graphe)
- recherche du plus court chemin dans un graphe
- etc.

Programmation

algorithmes	langage des mathématiques
programmes	langages de programmation

- exécution par une machine $\Rightarrow$ précision supplémentaire
- complexité dans l'organisation des programmes et des données $\Rightarrow$ modularité
- catalogue de méthodes de programmation bien connues qui s'apprennent
- choix important du langage de programmation

Un programme tout simple




```
static void copie (char[] s, int s_pos, char[] d, int d_pos,
                   int longueur) {
  for (int i = 0; i < longueur; ++i)
    d [d_pos + i] = s [s_pos + i];
}
```

Complexité $O(n)$ si n est la longueur.

Ce programme est-il correct?

Attention aux alias

Bug si s et d sont deux alias ($s = d$)

Le programme correct:

```
static void copie (char[] s, int s_pos, char[] d, int d_pos,
                   int longueur) {
  if ( s != d || d_pos <= s_pos)
    for (int i = 0; i < longueur; ++i)
      d [d_pos + i] = s [s_pos + i];
  else
    for (int i = longueur - 1; i >= 0; --i)
      d [d_pos + i] = s [s_pos + i];
}
```

La recherche d'alias dans un programme est un problème important (Ariane 5, Word 97, Windows 2000, etc).

Analyse statique de programmes.

Sémantique des opérateurs

```
static String paques (int y){                        y = 11571
    int g = (y % 19) + 1;                            g = 1
    int c = y / 100 + 1;                             c = 116
    int x = 3 * c / 4 - 12;                          x = 75
    int z = (8 * c + 5) / 25 - 5;                    z = 32
    int d = 5 * y / 4 - x - 10;                      d = 14378
    int e = (11 * g + 20 + z - x) % 30;              e = -12
    if ( e == 25 && g > 11 || e == 24 )
        ++e;
    int n = 44 - e;                                  n = 56
    if (n < 21)
        n = n + 30;
    int j = n + 7 - ((d + n) % 7);                   j = 63
    if (j > 31)
        return (j - 31) + " avril";                  32 avril
    else
        return j + " mars";
}
```

Que vaut le modulo d'un nombre négatif?

Domaine des valeurs

- Ariane 501: dépassement de capacité dans le système de guidage par inertie en convertissant un flottant 64 bits en entier 16 bits $\Rightarrow$ exception processeur $\Rightarrow$ traitement de l'exception en basculant sur le processeur de secours $\Rightarrow$ arrêt puisque ce processeur exécutait le même programme qui avait lui aussi fait la même exception.

- Coût: 10^9FF

- Leçon: ne pas penser qu'aux erreurs matérielles.
  Le logiciel peut aussi être erronné.

Les principes des langages de programmation sont étudiés en Majeure 1.
L'analyse statique est enseignée en DEA.

Au dela

- certification des programmes
- logique mathématique et preuves mécaniques
- systèmes d'exploitation
- architecture des processeurs et ordinateurs
- compilation
- CAO de circuits
- technologie (semi-conducteurs, physique quantique)
- réseaux
- protocoles et sécurité
- bases de données distribuées (Web)
- traitement et synthèse d'images
- etc.

Quelques domaines d'application

- Transports/Espace
- Médecine
- Biologie
- Télécommunications
- Monétique
- Environnement/Météo
- Cinéma/Multimédia
- Productique/Robotique
- Documentation/Enseignement
- Jeux
- Impact sur l'économie (*start-ups*)
- etc.

Monétique

Cartes à puce programmable en Java (*Javacard*).

Intérêt:

- charger de nouveaux services (banques multiples, carte téléphone, carte médicale, programmes de fidélisation, etc)
- programmation des services dans un langage de haut niveau

Espace limité: 32 KO de ROM, 8 KO d'EEPROM (*bytecode*), 512 octets de RAM (pile).

Les programmes chargés sont-ils sûrs?

$\Rightarrow$ Analyse statique ou preuves formelles.

(La France est bien placée.)

Télécommunications

- puissance de calcul dans un téléphone mobile est de 40 MFlops (opérations flottantes par seconde)
- 600 Gbits/sec passent sur un fil sans répéteur sur 10000km

Une technologie exponentielle

	gain	période
densité des puces	$\times 2$	18 mois
vitesse processeur	$\times 2$	18 mois
capacité mémoire	$\times 4$	3 ans
coût/octet disques	$\times 0.75$	1 an

En 2000,

- Processeur à 750 MHz, 10M transistors, ($\simeq 1300 \times$ **Vax**)
- RAM 128 Mbits, 133 MHz
- Disques 2.5” à 18GO, 3.5” à 20GO (1 kF)
- Réseaux ATM à 640 M bits/sec
- Internet avec 80 M machines

Vers un ordinateur mondial

Le World Wide Web est déjà une bibliothèque mondiale.
Ira-t-on vers un calculateur mondial?

L'exemple de la cryptographie:

- **F9** $= 2^{512} + 1 =$ $13407807929942597099574024998205846127479365820592393377723561443721764030073546976801874298166903427690031858186486050853753882811946569946433649006084097 =$ $2424833 \times 7455602825647884208337395736200454918783366342657 \times 741640062627530801524787141901937474059940781097519023905821316144415759504705008092818711693940737$ **[Manasse, Lenstra] en 1990.**

- **RSA-155** $=$ $10941738641570527421809707322040357612003732945449205990913842131476349984288934784717997257891267332497625752899781833797076537244027146743531593354333897 =$ $102639592829741105772054196573991675900716567808038066803341933521790711307779 \times 106603488380168454820927220360012878679207958575989291$

$5222706082371930628086 43$ (35 ans de machine de 400 MHz avec 300 machines en 7 mois) **[Lenstra, Odlysko, ...Morain]** en 1999.

- ECC2K-108. Soient P et Q deux points sur une courbe elliptique munie d'une opération $\otimes$ formant un groupe de $\sim 2^{108}$ éléments. Trouver n tel que $P^n = Q$.
  Réponse:
  $n = 4745566189622304529974831601 8941 \bmod$
  $32451855365842670148744865646 1467$ (500 ans de machine de 450 MHz avec 9500 machines de 40 pays en 5 mois) **[Harley et al]** en 2000.

- ...à l'écoute des extra terrestres (projet SETI)

L'informatique et l'X

- Futurs ingénieurs doivent connaître l'informatique
- Former de bons informaticiens (vs ENS, ETH, MIT)
- Recherche en informatique en plein essor
- Mathématiques ⇒ rigueur ⇒ atout en informatique (vrai en général)
- Formation initiale très hétérogène des élèves (dramatique)
- L'environnement informatique de l'X est bon. Il faut s'en servir.
- Format du cours demande un peu de travail personnel (beaucoup moins que les 15h/semaine de cours analogues)

Enseignement de l'informatique à l'X

Tronc Commun, obligatoire
1ère année, 1er semestre;
Projet informatique,
1ère année, 2ème semestre

Voie C, optionnelle
1ère année, 2ème semestre

Majeure 1, optionnelle
2ème année, 1er trimestre

Majeure 2, optionnelle
2ème année, 2ème trimestre

DEAs
3ème année

Contenus

Tronc Commun: algorithmes et programmation
Lévy (projet informatique *Gonthier*)

Voie C: optimisation combinatoire
Steyaert

Majeure 1: mathématiques et informatique
Steyaert
automates et calculabilité, architecture [*Vuillemin*],
algèbre $\vee$ bases de donnée $\vee$ lang. de programmation,
1 cours au choix (automates, vision, objets, . . .)

Majeure 2: informatique
Cori et *Toueg*
systèmes et réseaux, complexité,
algorithmes distribués $\vee$ compilation,
1 cours au choix (crypto, synthèse d'images, . . .)

Organisation du Tronc Commun

- 2.5 travaux dirigés d'initiation
- 10 petites classes + 10 travaux dirigés
- contrôle hors-classement, 2h;
  composition de classement, 4h,
  Cœfficient 10. Théorique.
- projets de programmation,
  ~1000 lignes de Java (en binôme)
  20 sujets différents de difficultés variées,
  Cœfficient 8. Pratique, travail personnel.

Groupes de niveaux (Forts, Moyens, Débutants)

L'équipe enseignante du Tronc Commun

- Ecole polytechnique: 1 professeur, 10 maîtres de conférences, 11 chefs de travaux pratiques, 11 vacataires, 1 secrétaire.
- Cours: 4 prof. univ., 5 DR, 1 CR, 1 ingénieur Mines
- TD: 6 CR, 3 MdC, 1 ingénieur Mines, 10 thésards.
- les enseignants sur poste sont tous des chercheurs reconnus.

Centre informatique de l'X

- Climat de confiance. Très difficile à 460 (900) élèves.
- Impression ($\simeq$ 10-20 centimes la page). Ne pas faire des impressions inutiles. Utiliser les photocopieuses.

Le contenu du Tronc Commun

programmes	itération	récursion	
structures de données	tableaux listes files, piles	arbres listes	graphes
algorithmes	tri recherche en table analyse syntaxique exploration		
projet informatique	modularité		

Langage de programmation

Java, [Gosling et Sun microsystems]
langage typé + *garbage collector*

Lisez nos pages Web!

Envoyez nous des e-mails.

Téléphonez nous. Venez nous voir.

Travaillez l'informatique.

On y trouve beaucoup de métiers intéressants.