

# Théorème PCP. Applications à l'inapproximabilité. PC 9. Correction

Soit  $q, r : \mathbb{N} \rightarrow \mathbb{N}$ . Un  $(r(n), q(n))$ -vérificateur de preuve probabiliste est un algorithme  $V$  en temps polynomial. Pour une entrée  $x$  et une preuve (= un certificat) que l'on considère être un mot  $P \in \{0, 1\}^*$ , l'algorithme  $V$  reçoit  $x$  et un vecteur  $r \in \{0, 1\}^{\mathcal{O}(r(|x|))}$  de bits aléatoires.  $V$  calcule alors un total de  $\mathcal{O}(q(|x|))$  positions et reçoit les valeurs des bits de  $P$  à ces positions précises. Avec cette information,  $V$  décide ou non d'accepter le triplet  $\langle x, r, P \rangle$ .

Un langage  $L$  appartient à  $\text{PCP}(r(n), q(n))$  s'il existe un  $(\mathcal{O}(r(n)), \mathcal{O}(q(n)))$  vérificateur tel que :

- $x \in L$  implique que  $\Pr(V \text{ accepte } x) = 1$ .
- $x \notin L$  implique que  $\Pr(V \text{ accepte } x) \leq 1/2$ .

On peut toujours supposer la preuve de longueur au plus  $q(n)2^{\mathcal{O}(r(n))}$ , où  $n = |x|$ .

## 1 De PCP au théorème PCP

1. Montrer que :
  - (a)  $\text{P} = \text{PCP}(0, 0)$ .
  - (b)  $\text{NP} = \text{PCP}(0, \text{poly})$ .
  - (c)  $\text{coRP} = \text{PCP}(\text{poly}, 0)$ .
2. Montrer que si  $L \in \text{PCP}(r(n), q(n))$  alors  $L$  est reconnaissable en temps non-déterministe  $2^{\mathcal{O}(r(n)+\log(n))}$ .

## 2 Problème MAX3SAT : Approximabilité

Montrer qu'il existe un algorithme en temps polynomial qui, étant donnée une formule  $\phi$  en 3CNF avec  $n$  variables et  $m$  clauses, telle que chaque clause possède trois variables distinctes, produit une affectation qui satisfait  $7m/8$  clauses.

Correction: *L'idée de départ est la suivante : on considère une assignation aléatoire uniforme des variables  $x_i$  d'une formule  $\phi$ . Notons  $C_j$  la variable aléatoire*

qui vaut 1 si la clause  $c_j$  est satisfaite, et qui vaut 0 sinon. Alors,

$$\Pr(C_j = 1) = 7/8$$

et donc  $E[C_j] = 7/8$ . Donc si  $C = \sum_j C_j$ , alors  $E[\text{nb de clauses satisfaites}] = E[C] = \sum_j E[C_j] = 7m/8$ .

Une assignation moyenne satisfait  $7/8$  des clauses. Il suffit d'en trouver une qui satisfait plus que  $7/8$  des clauses.

Voici un algorithme déterministe pour cela.

- Soient  $x_1, x_2, \dots, x_n$  les variables booléennes de  $\phi$ .
- Pour  $i = 1, 2, \dots, n$ , on considère les deux valeurs possibles 0 et 1 pour  $x_i$ .
  - Pour chaque clause  $c_j$ , calculer  $E[C_j | x_i = b]$  :
    - supposons que  $c_j$  contienne  $k$  littéraux.
      - si  $c_j$  ne contient ni  $x_i$ , ni son complément, alors  $E[C_j | x_i = 1] = (2^k - 1)/2^k$ .
      - si  $c_j$  contient  $x_i$ , alors  $E[C_j | x_i = 1] = 1$ .
      - si  $c_j$  contient le complément de  $x_i$ , alors  $E[C_j | x_i = 1] = (2^{k-1} - 1)/2^{k-1}$ .
  - Comparer  $\sum_j E[C_j | x_i = 0]$  à  $\sum_j E[C_j | x_i = 1]$  : choisir  $x_i = b$  tel que  $\sum_j E[C_j | x_i = b]$  soit maximal.  
(cette valeur est plus grande que  $7/8m$ , puisque

$$E[C] = 1/2 * \sum_j E[C_j | x_i = b] + 1/2 * \sum_j E[C_j | x_i = 0] \geq 7/8m,$$

et donc la plus grande des deux valeurs est plus grande que  $7/8m$ .)

- Propager cette valeur : continuer en remplaçant  $\phi$  par  $\phi$  où l'on a remplacé  $x_i$  par  $b$ .

(par construction, on a encore et toujours pour la nouvelle formule  $E[C] \geq 7/8m$ ).

Par cette construction on finit par tomber sur des clauses satisfaites, et on produit bien au final une affectation qui satisfait plus que  $7/8m$  des clauses.

### 3 Problème MAX3SAT : Difficulté d'approximation

Rappel : le théorème PCP affirme le résultat surprenant  $\text{NP} = \text{PCP}(\log n, 1)$ .

On va l'utiliser pour montrer la difficulté d'approximer MAX3SAT.

1. Montrer que s'il existe  $\delta > 0$ , et un problème NP complet  $K$  et une réduction  $f$  de  $K$  vers 3SAT calculable en temps polynomial telle que
  - $x \in K$  implique que  $f(x)$  est une formule satisfiable
  - $x \notin K$  implique que au plus  $(1 - \delta)m$  clauses de  $f(x)$  sont satisfiables (on parle de réduction écartante) alors MAX3SAT n'est pas approximable en temps polynomial avec un rapport meilleur que  $1 - \delta$ , sauf si  $P = \text{NP}$ .

Correction: Soit  $K$  un problème NP-complet. Supposons qu'il existe une réduction  $f$  de  $K$  vers 3SAT calculable en temps polynomial telle que

- $x \in K$  implique que  $f(x)$  est une formule satisfiable
- $x \notin K$  implique que au plus  $(1 - \delta).m$  clauses de  $f(x)$  sont satisfiables.

Supposons qu'il existe un algorithme  $A$  en temps polynomial qui approxime MAX3SAT avec rapport  $1 - \delta$ .

Alors on peut résoudre  $K$  en temps polynomial : en effet, sur une entrée  $x$ ,

- (a) on calcule  $f(x)$ .
- (b) on lance l'algorithme  $A$  sur la formule  $f(x)$ .
- (c) si le résultat de cet algorithme  $A$  est plus grand que  $(1 - \delta).m$  où  $m$  est le nombre de clauses de  $f(x)$ , on accepte. Sinon on rejette.

En effet, d'une part si  $x \in K$ , alors  $f(x)$  est satisfiable, et donc  $A$  retourne une valeur entre  $m$  et  $(1 - \delta)m$ , soit toujours plus grande que  $(1 - \delta).m$ . D'autre part si  $x \notin K$ , alors  $f(x)$  possède au plus  $(1 - \delta).m$  clauses satisfiables, et donc  $A(x)$  va produire une réponse qui sera entre  $(1 - \delta)^2.m$  et  $(1 - \delta).m$ , et donc plus petite que  $(1 - \delta).m$ .

2. Utiliser le théorème PCP pour montrer que pour un certain  $\delta > 0$ , il existe une réduction  $f$  de 3SAT vers 3SAT telle que
  - (a) si  $\phi$  est une formule satisfiable, alors  $f(\phi)$  est satisfiable
  - (b) si  $\phi$  est insatisfiable, alors une fraction au plus  $1 - \delta$  des clauses de  $f(\phi)$  sont satisfiables simultanément.

Correction: Puisque 3SAT est dans NP, par le théorème PCP (= le résultat  $NP = PCP(\log n, 1)$ ), on sait que pour des constantes  $c$  et  $k$ , il existe un vérificateur probabiliste pour 3SAT qui utilise au plus  $c \log n$  bits probabilistes et consulte au plus  $k$  bits de la preuve.

Soit  $N = 2^{c \log n} \leq n^c$ . Il existe au plus  $N$  choix de bits probabilistes et chacun mène la consultation de  $k$  bits de la preuve. On peut donc supposer que la preuve est de longueur  $kN$  et que  $\{0, 1\}^{kN}$  est l'ensemble des preuves possibles.

Soit  $C$  l'ensemble des 3-clauses de  $\phi$  et  $n$  le nombre de variables booléennes. Pour chaque  $0 \leq r \leq N - 1$ , on construit la fonction  $f_r : \{0, 1\}^{kN} \rightarrow \{0, 1\}$ , telle que

- (a)  $f_r(p) = 1$  si la preuve  $p$  est acceptée par le vérificateur lorsque l'on utilise la suite de choix aléatoires donnée par  $r$ .
- (b)  $f_r(p) = 0$  si la preuve  $p$  est rejetée par le vérificateur lorsque l'on utilise la suite de choix aléatoires donnée par  $r$ .

Observons que pour un  $r$  fixé, la fonction  $f_r$  ne dépend que de  $k$  bits de son entrée. Nous notons ces positions  $j_r(1) < j_r(2) < \dots < j_r(k)$ .

Par définition de PCP, on sait que

- (a) si  $\phi$  est satisfiable, alors il existe une preuve  $\pi \in \{0,1\}^{kN}$  telle que  $f_r(\pi) = 1$  pour tous les  $r$ .
- (b) si  $\phi$  est insatisfiable, alors on a que  $f_r(\pi) = 1$  pour au plus la moitié des  $r$ .

Puisque chaque fonction  $f_r$  ne dépend que de  $k$  bits d'entrée, on peut donc écrire une formule en forme normale conjonctive de taille au plus  $k2^k$  décrivant  $f_r$ . Les variables booléennes utilisées ici sont  $\pi_{j_r(1)}, \dots, \pi_{j_r(k)}$ .

Par la technique habituelle pour transformer une formule SAT en formule 3SAT on peut transformer cette formule en formule 3SAT, quitte à introduire de nouvelles variables. On peut donc construire pour chaque  $f_r$ , un ensemble de  $k^*2^k$  clauses de trois littéraux ( $k^* = \max(1, k - 2)$ ) telles que

- (a) si  $f_r(\pi) = 1$ , alors toutes ces clauses sont satisfiables simultanément
- (b) si  $f_r(\pi) = 0$ , alors au moins une clause est insatisfaite.

En construisant ces clauses pour chaque  $f_r$ , on obtient un ensemble d'au plus  $k^*2^k N$ , avec  $k^* = \max(1, k - 2)$  clauses de longueur 3 telles que :

- (a) si  $\phi$  est satisfiable, alors toutes ces clauses sont satisfiables
- (b) si  $\phi$  est insatisfiable, alors au moins  $N/2$  de ces nouvelles clauses sont insatisfiables pour toute assignation (il y en a au moins une pour au moins la moitié des  $f_r$ ).

On prend  $\delta = k^*2^{k+1}$ . Si  $\phi$  est insatisfiable, alors la proportion de nos nouvelles clauses qui sont simultanément satisfiables est  $1 - (N/2)/(Nk^*2^k) = 1 - \delta$ .

3. Montrer que MAX3SAT n'est pas  $\alpha$ -approximable pour un certain  $\alpha$ , sauf si  $P = NP$ .

Correction: La question précédente donne une réduction écartante de 3SAT vers 3SAT. Le résultat découle directement de la première question ( $\alpha = 1 - \delta$ ).

## 4 Calcul du permanent

Soit  $A = (a_{i,j})$  une matrice carrée à coefficients entiers. Le *permanent* est la quantité

$$\sum_{\sigma} \prod_{i=1}^n a_{i,\sigma(i)}.$$

1. Rappeler la définition du déterminant. Quelle est la complexité du calcul du déterminant.
2. Rappeler la formule de développement selon la première colonne du déterminant. Donner une formule similaire pour le permanent.
3. En utilisant ce développement, proposer un protocole de preuve interactif pour le permanent.