Propriétés facilement testables

Rémi de Verclos, Louis Esperet

4 novembre 2015

Property testing

But : tester une propriété de graphe $\mathcal{P}$ en ne regardant qu'un petit nombre d'arêtes de G.
Cela se ramène à l'algorithme suivant.

Algorithme

Entrée : un Graphe G
On tire un ensemble aléatoire X de sommet et on accepte ssi $G[X] \in \mathcal{P}$.

Le nombre de requêtes d'un tel algorithme est à peu près $|X|^2$.

- Si $G \in \mathcal{P}$ et $\mathcal{P}$ est héréditaire, l'algorithme accepte toujours (pas de faux négatif).
- Si G ressemble beaucoup à un graphe de $\mathcal{P}$, l'algorithme a de bonnes chances de se tromper.

Une distance entre graphes

La distance $d(G, H)$ entre deux graphes G et H (de même taille) est le plus petit ϵ tel qu'on peut passer de G à H en changeant ϵn^2 arêtes.
G est dit ϵ-*loin* de $\mathcal{P}$ si

$$\forall H \in \mathcal{P}, d(G, H) > \epsilon.$$

Propriétés testables (avec algorithme one-sided)

Definition

Une propriété (héréditaire) $\mathcal{P}$ est testable si pour tout $\epsilon > 0$ il existe $m(\epsilon)$ tel que pour tout graphe G ϵ-loin de $\mathcal{P}$,

$$\mathbb{P}(G[X] \notin \mathcal{P}) < \frac{1}{2}.$$

où $|X| = m(\epsilon)$

Remarques :

- On ne suppose rien sur le comportement de l'algorithme si G n'est ni dans $\mathcal{P}$ ni ϵ-loin de $\mathcal{P}$.
- $m(\epsilon)$ est **indépendant** de la taille de G.

Theorem (Alon, Shapira 2008)

Une propriété est testable (avec un testeur one-sided) ssi elle est presque héréditaire.

Problème : le nombre $m(\epsilon)$ de sommets à examiner vient d'une version renforcée du lemme de régularité. C'est une tour d'exponentielles de hauteur $\frac{1}{\epsilon}$.

Propriétés facilement testables

On cherche donc à savoir pour quelles classe $m(\epsilon)$ est polynomial en $\frac{1}{\epsilon}$.

Definition

Une propriété (héréditaire) $\mathcal{P}$ est facilement testable si pour tout $\epsilon > 0$ il existe $m(\epsilon) = \left(\frac{1}{\epsilon}\right)^c$ tel que pour tout graphe G ϵ-loin de $\mathcal{P}$,

$$\mathbb{P}(G[X] \notin \mathcal{P}) < \frac{1}{2}$$

où $|X| = m(\epsilon)$

Classes H-free

Si $\mathcal{P}$ est la classe des graphes sans H comme sous-graphe induit.

- (Alon, Shapira 2006) $\mathcal{P}$ est difficilement testable sauf si $H = C_4$ ou P_i pour $i \in \{1, 2, 3, 4\}$ ou leurs complémentaires.
  Ces preuves sont des constructions : les auteurs construisent un graphe G_ϵ, ϵ-loin de $\mathcal{P}$ avec au plus $\epsilon^{\log(1/\epsilon)} n^{|H|}$ copies induites de H.
- $\mathcal{P}$ est facilement testables pour $H = P_i$ pour $i = 1, 2, 3, 4$ et leurs complémentaire.
  Toutes ces classes ont une structure inductive très fortes.
- Le problème est ouvert pour $H = C_4$.

Les preuves de facilité de testabilité utilisent la structure des classe.

Classes difficilement testables : méthode sandwich

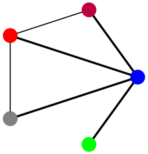
Alon et Fox (2015) ont construit une famille de graphe

- Avec beaucoup de C_5.
- Tel que que $G[X]$ est de comparabilité avec grande probabilité si $|X| = \left(\frac{1}{\epsilon}\right)^c$.

$\rightarrow$ Si *Comparability* $\subset \mathcal{P} \subset C_5$-Free alors $\mathcal{P}$ n'est pas facilement testable (ex : Les graphes parfaits).

Graphes trivialement parfaits

Les graphes *trivialement parfaits* sont les graphes sans P_4 ni C_4 induit.



Ce sont aussi les graphes d'intersection d'intervalles I_v tels que si $I_v \cap I_u$ alors $I_v \subset I_u$ ou $I_u \subset I_v$.

Theorem

La classe des graphes trivialement parfaits est facilement testable.

Sous-classes de classes facilement testables

Si $\mathcal{P} \subset \mathcal{G}$ et $\mathcal{Q}$ est facilement testable, $\mathcal{P}$ n'est pas forcément facilement testable. En revanche,

Theorem

Les sous-classes héréditaires des graphes trivialement parfaits stables par ajout de jumeaux sont facilement testables.*

*$\mathcal{P}$ est *stable par ajout de jumeaux* si en dupliquant un sommet v d'un graphe de $\mathcal{P}$ par v' de même voisinage et en ajoutant éventuellement une arête entre v et v', on obtient un graphe de $\mathcal{P}$.

Classes semi-algébriques

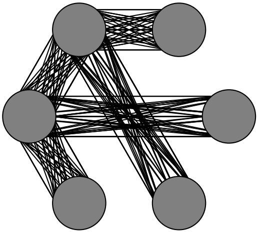
Une classe de graphe $\mathcal{P}$ est *semi-algébrique* si dans chaque $G \in \mathcal{P}$,

- Chaque sommet v correspond à un vecteur $a_u \in \mathbb{R}^k$.
- L'arêtes uv est déterminées par des inégalités polynomiales en les coordonnées de a_u et a_v.

Les graphes d'intervalles et toutes leurs sous-classes sont semi-algébriques.

Les structures semi-algébriques ont un lemme de régularité très fort.

Lemme de régularité polynomial

$G \approx$ 

Theorem (Fox,Pach,Suk)

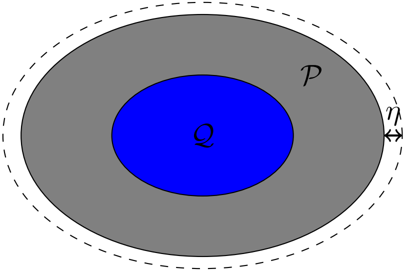
Soit $\mathcal{P}$ une classe de graphe semi-algébrique, il existe une constante c telle que pour tout graphe G, il existe une partition $V_1, \dots, V_k$ des sommets de G telle que

- $||V_i| - |V_j|| \leq 1$ *(La partition est équilibrée)*
- $k \leq \left(\frac{1}{\epsilon}\right)^c$ *(La partition est de taille polynomiale)*
- *G contient soit toutes soit aucune arêtes de $V_i \times V_j$, pour toute paire i, j sauf au plus ϵk^2.*

Conséquence du lemme de régularité polynomial

Corollary

Si $\mathcal{P}$ est semi-algébrique et $\mathcal{Q} \subset \mathcal{P}$ est stable par ajouts de jumeaux alors $\mathcal{Q}$ est facilement testable dans $\mathcal{P}$

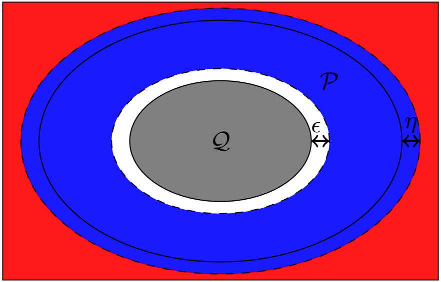


En fait, on peut même tester $\mathcal{Q}$ dans $\{G/d(G,\mathcal{P}) \leq \eta(\epsilon)\}$.

Pseudo-transitivité de la testabilité

Theorem

Si $\mathcal{P}$ est semi-algébrique et facilement testable et $\mathcal{Q} \subset \mathcal{P}$ est stable par ajout de jumeaux alors $\mathcal{Q}$ est facilement testable.



Preuve : On teste si $G[X] \in \mathcal{Q}$, pour $|X|$ de taille adéquate.

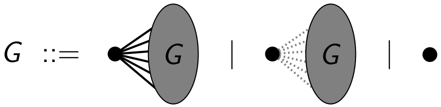
- Si $d(\mathcal{P}, G) > \eta(\epsilon)$, alors G est rejeté avec probabilité $\frac{1}{2}$ car $\mathcal{P}$ est facilement testable.
- Si $d(\mathcal{P}, G) \leq \eta(\epsilon)$ et $d(\mathcal{Q}, G) > \epsilon$, alors G est rejeté avec probabilité $\frac{1}{2}$ car $\mathcal{Q}$ est facilement testable dans $\{G \mid d(\mathcal{P}, G) \leq \epsilon\}$.

Definition

$G = (V, E)$ est un graphe de seuil si on peut associer à chaque sommet v un nombre réel a_v de sorte que

$$uv \in E \Leftrightarrow a_u + a_v \geq 1.$$

Ce sont exactement les graphes construits inductivement à partir du graphe vide par ajout de sommet universel ou isolé.



Theorem

La classe des graphes de seuil est facilement testable.

La preuve est une version simplifiée de celle celles de la facilité de la testabilité des graphes sans P_4 ou des graphes trivialement parfaits.

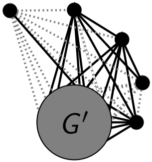
Preuve : les graphes de seuils sont facilement testables

Lemma

Si pour tout $v \in V(G)$, $\epsilon n < d(v) < (1-\epsilon)n$ *alors* $G[X]$ *n'a pas de sommet universel ni isolé avec probabilité* $\frac{2}{3}$, *où* $|X| = c \cdot \frac{1}{\epsilon}\ln(\frac{1}{\epsilon})$.

Rq : Dans ce cas, $G[X]$ n'est pas de seuil.

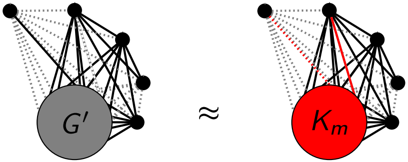
On décompose G : Tant que G a un sommet v vérifiant $d(v) \geq (1-\epsilon)n$ ou $d(v) \leq \epsilon n$, on supprime v.



Lorsqu'on s'arrête on obtient un graphe G', vérifiant les hypothèses du lemme ci-dessus.

Preuve : les graphes de seuils sont facilement testables

- Si $|G'| \leq \epsilon n$, alors G est ϵ-proche d'un graphe seuil.



- Si $|G'| > \epsilon n$, alors en tirant $|Y| = c' \cdot \frac{1}{\epsilon^2} \ln(\frac{1}{\epsilon})$ sommets aléatoires, alors $\mathbb{P}(|Y \cap G'| \geq c \frac{1}{\epsilon} \ln(\frac{1}{\epsilon})) \geq \frac{3}{4}$.

Donc si G est ϵ-loin d'être un graphe de seuil, l'algorithme rejette G avec probabilité $\frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2}$.

Merci pour votre attention

Problèmes ouverts :

- La classe des graphes d'intervalles est-elle facilement testable ?
- L'intersection de deux classes facilement testables est-elle facilement testable ?