

Algorithmes distribués probabilistes

A. Zemmari

LaBRI - Université Bordeaux 1

15 septembre 2010

Définitions : Système distribué [Tel]

- A distributed system is an interconnected collection of autonomous computers, processes, or processors. The computers, processes, or processors are referred to as the nodes of the distributed system.
- to be qualified as autonomous, the nodes must at least be equipped with their own private control,
- to be qualified as interconnected, the nodes must be able to exchange information.

A distributed algorithm for a collection $P = \{p_1, p_2, \dots, p_n\}$ of processes is a collection of local algorithms, one for each process in P .

\mathcal{S} : système distribué

\mathcal{A} : Algorithme distribué

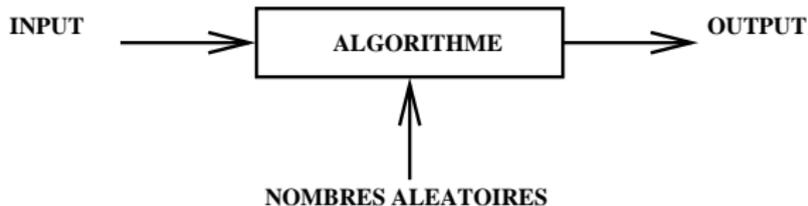
Modèle de \mathcal{S} : $G = (V, E)$ ($|V| = n$, $|E| = m$) graphe connexe, simple et symétrique.

- Processus : tout est local : état, mémoire, horloge, identité, tampons, variables, calculs, etc.
- Communication par messages
 - point-à-point/multipoint (voisins/difusions)
 - synchrone/asynchrone
- \mathcal{S} - \mathcal{A} avec/sans information globale, statique/dynamique, avec identités/anonyme
- Terminaison de \mathcal{A} ? explicite/implicite
- Mesures de complexité ? en messages/bits, en mémoire, en temps/phases-rondes.

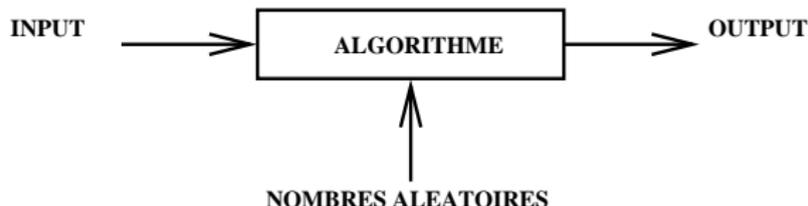
Algorithme probabiliste [Motwani & Raghavan]



Algorithme probabiliste [Motwani & Raghavan]



- In addition to input, algorithm takes a source of random numbers and makes random choices during execution,
- Behavior can vary even on a fixed input.



- Design algorithm + analysis to show that this behavior is likely to be good, on every input.
(The likelihood is over the random numbers only.)

Catégorie des algorithmes probabilistes (Motwani & Ragahavan)

- A Monte Carlo algorithm runs for a fixed number of steps and produces an answer that is correct with probability $\geq \frac{1}{3}$.
- A Las Vegas algorithm always produces the correct answer; its running time is a random variable whose expectation is bounded (say by a polynomial).

Briser la symétrie, Résultat d'impossibilité

Théorème [Angluin-80]

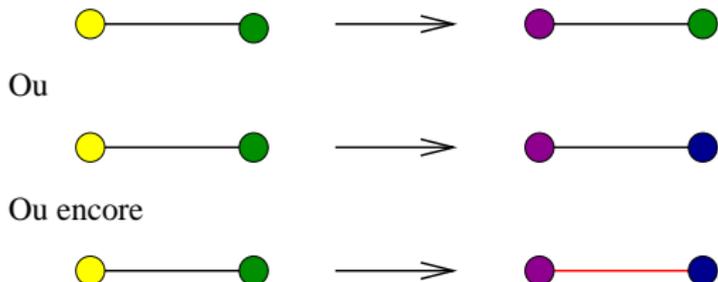
- Il n'existe aucun algorithme d'élection déterministe dans un réseau anonyme, de taille n connue ou non des processus.

Il en résulte, sous l'hypothèse d'anonymat, qu'il n'existe pas d'algorithme déterministe pour :

- colorier un graphe,
- réaliser des *handshake*,
- construire un MIS,
- construire un couplage maximal,
- ...

Problème 1. Réaliser des *handshake*

- Un des modèles de communication de base dans le cas des algorithmes distribués codés par les calculs locaux : les règles de réécriture sont de la forme :



- Exemples :

- Calcul d'arbre couvrant;
- Élection dans un arbre,
- Communication dans un réseau de robots (DHSZ'2004).
- ...

Solution : un algorithme probabiliste

- Chaque sommet v répète tout le temps :
 - Le sommet v choisit un de ses voisins $c(v)$ au hasard;
 - v envoie 1 à $c(v)$;
 - v envoie 0 à tous ses autres voisins
 - v reçoit les messages de tous ses voisins.
(* il y a handshake entre v et $c(v)$ si v reçoit 1 de $c(v)$ *)

- Définition : Soit $G = (V, E)$ un graphe. Un appel sur G est une fonction c de V dans V qui envoie un sommet v sur un de ses voisins.
- \Rightarrow soit c un appel, il y a un handshake sur G ssi il existe deux sommets v et w tels que $c(v) = w$ et $c(w) = v$.
- Définition : un appel c est un succès s'il y a au moins un handshake dans le graphe. Sinon c est dit échec.

- Définition : Soit $G = (V, E)$ un graphe. Un appel sur G est une fonction c de V dans V qui envoie un sommet v sur un de ses voisins.
- \Rightarrow soit c un appel, il y a un handshake sur G ssi il existe deux sommets v et w tels que $c(v) = w$ et $c(w) = v$.
- Définition : un appel c est un succès s'il y a au moins un handshake dans le graphe. Sinon c est dit échec.

- Définition : Soit $G = (V, E)$ un graphe. Un appel sur G est une fonction c de V dans V qui envoie un sommet v sur un de ses voisins.
- \Rightarrow soit c un appel, il y a un handshake sur G ssi il existe deux sommets v et w tels que $c(v) = w$ et $c(w) = v$.
- Définition : un appel c est un succès s'il y a au moins un handshake dans le graphe. Sinon c est dit échec.

- Quelle est la probabilité de succès ?
- Combien d'itérations faut-il faire pour obtenir un succès ?
- Quel est le nombre moyen de handshake dans le graphe ?
- (Question plus dure) Quelle est la probabilité d'obtenir exactement k handshake dans le graphe ?

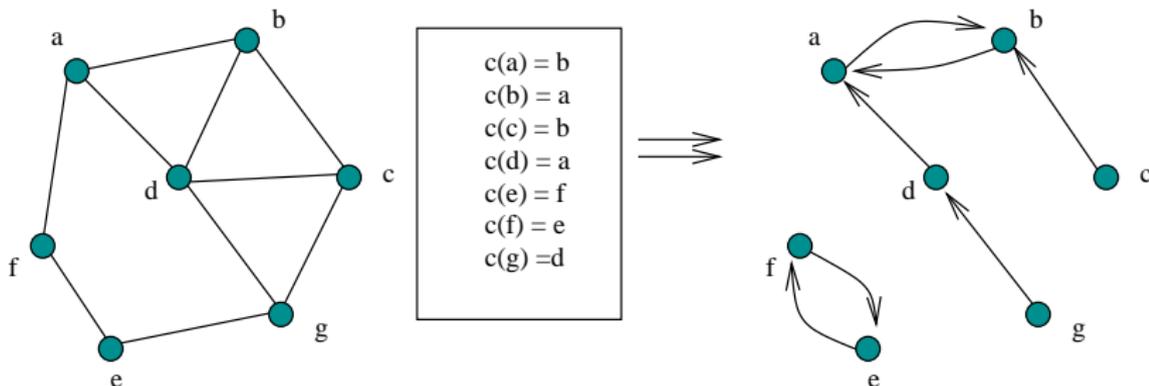
- Quelle est la probabilité de succès ?
- Combien d'itérations faut-il faire pour obtenir un succès ?
- Quel est le nombre moyen de handshake dans le graphe ?
- (Question plus dure) Quelle est la probabilité d'obtenir exactement k handshake dans le graphe ?

- Quelle est la probabilité de succès ?
- Combien d'itérations faut-il faire pour obtenir un succès ?
- Quel est le nombre moyen de handshake dans le graphe ?
- (Question plus dure) Quelle est la probabilité d'obtenir exactement k handshake dans le graphe ?

- Quelle est la probabilité de succès ?
- Combien d'itérations faut-il faire pour obtenir un succès ?
- Quel est le nombre moyen de handshake dans le graphe ?
- (Question plus dure) Quelle est la probabilité d'obtenir exactement k handshake dans le graphe ?

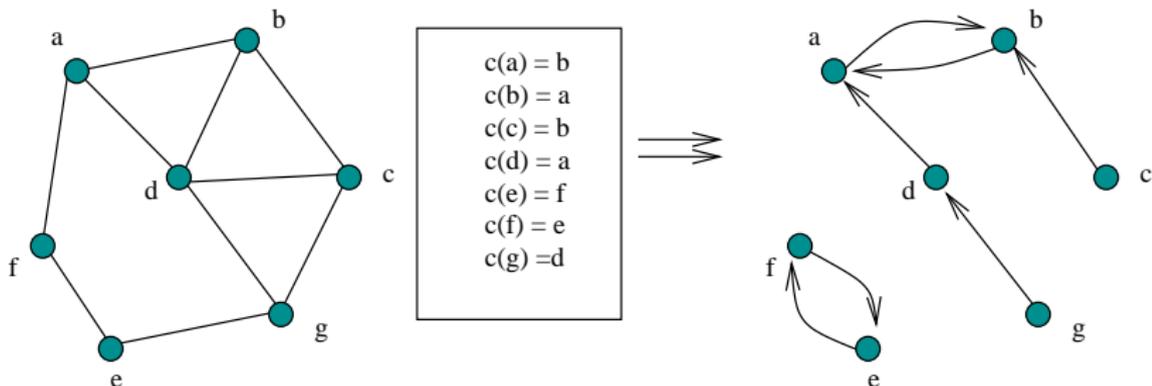
Graphe d'appel

- Soit $G = (V, E)$ un graphe. A chaque appel c sur G correspond un graphe orienté $G_c = (V, A)$, où un arc $a = (v, w) \in A$ si et seulement si $c(v) = w$.
- Exemple



Graphe d'appel

- Soit $G = (V, E)$ un graphe. A chaque appel c sur G correspond un graphe orienté $G_c = (V, A)$, où un arc $a = (v, w) \in A$ si et seulement si $c(v) = w$.
- Exemple



- Fait : Soit c un appel sur un graphe G . c est un échec si et seulement si G_c n'a pas de cycle de longueur 2.
- Corollaire : Si $G = (V, E)$ est un arbre, alors tout appel sur G est un succès.
- Preuve : Par récurrence sur la taille de G .

- Fait : Soit c un appel sur un graphe G . c est un échec si et seulement si G_c n'a pas de cycle de longueur 2.
- Corollaire : Si $G = (V, E)$ est un arbre, alors tout appel sur G est un succès.
- Preuve : Par récurrence sur la taille de G .

- Tous les sommets voisins d'un sommet ont la même probabilité $1/d(v)$ d'être choisi par v ;
- On définit la mesure de probabilité qui affecte à tout appel c sur G une probabilité :

$$\alpha(G) = \prod_{v \in V} \frac{1}{d(v)}.$$

- Lemme : Si on note par $s(G)$ la probabilité de succès et $f(G) = 1 - s(G)$. On a alors :

$$f(G) = \alpha(G)N(G) \text{ et } s(G) = 1 - \alpha(G)N(G),$$

où $N(G)$ est le nombre d'appels c sur G pour lesquels G_c n'a pas de cycle de longueur 2.

- Tous les sommets voisins d'un sommet ont la même probabilité $1/d(v)$ d'être choisi par v ;
- On définit la mesure de probabilité qui affecte à tout appel c sur G une probabilité :

$$\alpha(G) = \prod_{v \in V} \frac{1}{d(v)}.$$

- Lemme : Si on note par $s(G)$ la probabilité de succès et $f(G) = 1 - s(G)$. On a alors :

$$f(G) = \alpha(G)N(G) \text{ et } s(G) = 1 - \alpha(G)N(G),$$

où $N(G)$ est le nombre d'appels c sur G pour lesquels G_c n'a pas de cycle de longueur 2.

- Tous les sommets voisins d'un sommet ont la même probabilité $1/d(v)$ d'être choisi par v ;
- On définit la mesure de probabilité qui affecte à tout appel c sur G une probabilité :

$$\alpha(G) = \prod_{v \in V} \frac{1}{d(v)}.$$

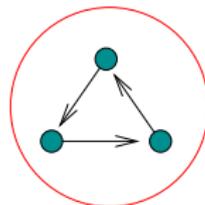
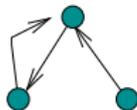
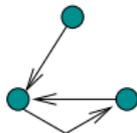
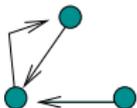
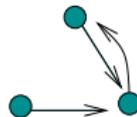
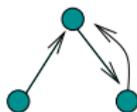
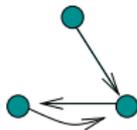
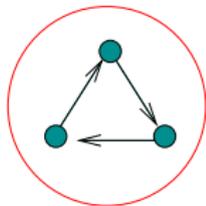
- Lemme : Si on note par $s(G)$ la probabilité de succès et $f(G) = 1 - s(G)$. On a alors :

$$f(G) = \alpha(G)N(G) \text{ et } s(G) = 1 - \alpha(G)N(G),$$

où $N(G)$ est le nombre d'appels c sur G pour lesquels G_c n'a pas de cycle de longueur 2.

Exemple

- Anneau de taille $n = 3$:



- Soit G un anneau de taille n . Le nombre $N(G)$ d'appels sans cycle de longueur 2 est égal à 2. Donc

$$f(G) = \frac{1}{2^{n-1}} \text{ et } s(G) = 1 - \frac{1}{2^{n-1}}.$$

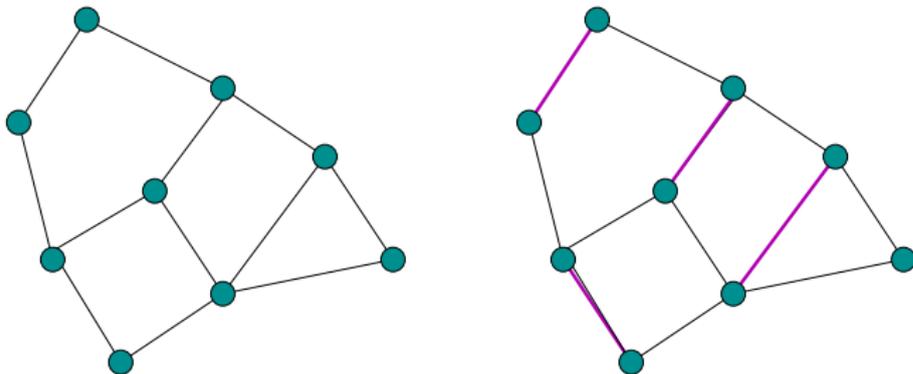
- L'espérance du nombre d'appels pour obtenir un succès est

$$\left(1 - \frac{1}{2^{n-1}}\right) + 2\frac{1}{2^{n-1}} \left(1 - \frac{1}{2^{n-1}}\right) + \dots = \frac{2^{n-1}}{2^{n-1} - 1}$$

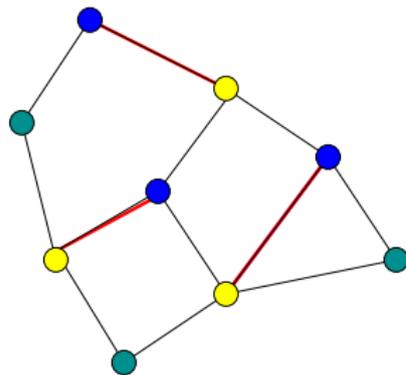
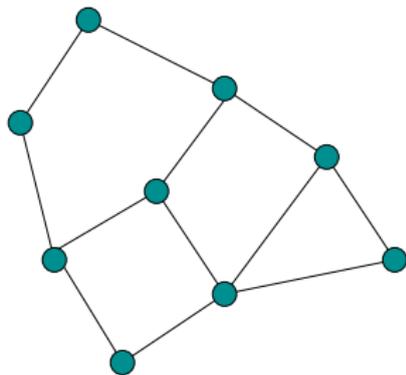
Couplage de graphe

- Définition : Soit $G = (V, E)$ un graphe. Un couplage sur G est un sous-ensemble M de E tel que :

$$\forall e, e' \in M, e \cap e' = \emptyset$$



handshake \leftrightarrow couplages



- Notation : soit $e \in E$, on note par e^1 (rep. e^0) l'événement : il y a un handshake sur e (resp. il n'y a pas de handshake sur e). Si $e = \{v, w\}$, alors $Pr(e^1) = 1/d(v)d(w)$.
- Pour tout couplage $M = \{e_1, e_2, \dots, e_k\}$, la probabilité pour M d'être un ensemble de handshake est

$$Pr(M) = Pr(e_1^1 \wedge e_2^1 \wedge \dots \wedge e_k^1) = \prod_{\{v,w\} \in M} (1/d(v)d(w)),$$

- Soit k un entier, un k -couplage sur G est un couplage de taille k . Soit \mathcal{M}_k l'ensemble des k -couplages sur G . On note :

$$q_k = \sum_{M \in \mathcal{M}_k} Pr(M), k = 0, 1, \dots, n/2$$

- Proposition : Soit q_k la séquence définie ci-dessus pour $k = 0, 1, \dots, n/2$. Pour tout entier l , la probabilité d'obtenir exactement l handshake sur le graphe G est :

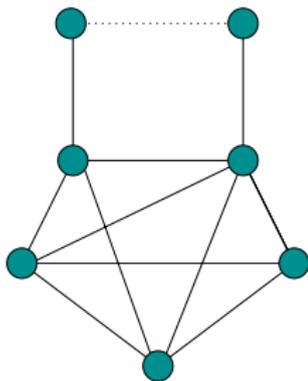
$$P_l = \sum_{l \leq i \leq n/2} (-1)^{i+l} q_i$$

- la probabilité de succès est alors :

$$s(G) = P_1 = \sum_{0 \leq i \leq n/2-1} (-1)^i q_{i+1}$$

Remarque

- Question : quel est l'impact de l'ajout d'une arête dans le graphe sur la probabilité de succès ?
- Si on rajoute une arête à un arbre, la probabilité de succès diminue.
- Si on ajoute une arête au graphe suivant, cette probabilité augmente (contre exemple fournit par V. Dikert et H. Austinat)



Espérance du temps entre deux handshake

- Pour un sommet v :

$$\frac{d(v)}{\sum_{w \text{ voisin de } v} \frac{1}{d(w)}}$$

- Pour une arête $e = \{v, w\}$:

$$d(v)d(w)$$

- Si le graphe est de degré borné par d , alors
 - Pour un sommet, l'espérance est bornée par d
 - Pour une arête, elle est bornée par d^2 .

Espérance du nombre de handshake dans le graphe

- Soit X la v.a. définie par $X =$ nbre de handshake dans G et soit $E(X)$ son espérance mathématique.
- Pour toute arête e , notons par χ_e la v.a. définie par $\chi_e = 1$ si il y a un handshake sur e et 0 sinon. On a alors :

$$X = \sum_{e \in E} \chi_e$$

- Or, la linéarité de l'espérance nous assure que :

$$E(X) = \sum_{e \in E} E(\chi_e)$$

- χ_e est une v.a. qui suit une loi de Bernoulli de paramètre $1/d(v)d(w)$ si $e = \{v, w\}$. Donc $E(\chi_e) = 1/d(v)d(w)$.
- D'où

$$E(X) = \sum_{\{u,v\} \in E} \frac{1}{d(u)d(v)}$$

- Si G est un graphe complet de taille n , alors

$$E(X) = \binom{n}{2} \frac{1}{(n-1)^2} = \frac{n}{2(n-1)}$$

- Si G est un anneau, alors :

$$E(X) = \frac{n}{4}$$

- Si G est un graphe de degré majoré par d , alors :

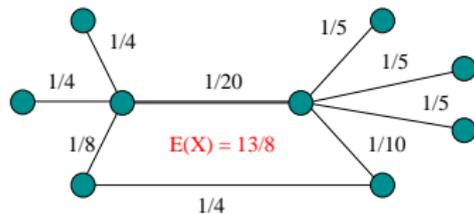
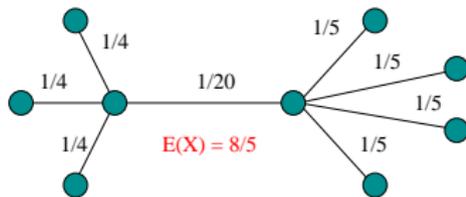
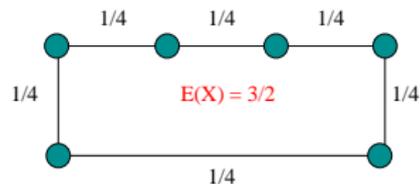
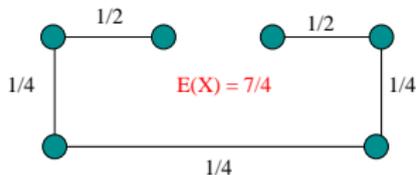
$$E(X) \geq \frac{|E|}{d^2}$$

- Si G est un arbre, alors

$$E(X) \geq \frac{n-1}{d^2}$$

Grphe minimisant $E(X)$

- Impact de l'ajout d'une arête n'est pas monotone :



Graphe minimisant $E(X)$ (2)

- Proposition : Pour n fixe, le graphe complet K_n est le graphe qui minimise l'espérance du nombre de handshake dans tous les graphes de taille n . Le nombre minimum de handshake est alors $E(X) = n/2(n - 1)$.
- Preuve. (voir tableau)

Étude de cas particuliers

1. graphes de degrés bornés

- $G = (V, E)$ un graphe de degré maximum d .
 - Proposition :

$$s(G) \geq 1 - \left(1 - \frac{1}{d^2}\right)^{|E|}$$

- Preuve : (voir le tableau) .
- Corollaire :

$$f(G) \leq e^{-\frac{|E|}{d^2}}$$

2. Graphes complets

- Soit K_n le graphe complet de taille n .

- Proposition :



$$s(K_n) = \sum (-1)^{k+1} \frac{n!}{k! 2^k (n-2k)! (n-1)^{2k}}$$

- $s(K_n)$ est asymptotiquement égal à $1 - e^{-1/2}$.
- L'espérance du nombre d'appels nécessaires pour obtenir un handshake est asymptotiquement égale à

$$\frac{\sqrt{e}}{\sqrt{e} - 1}$$

Borne générale pour $s(G)$

- Théorème : Soit $G = (V, E)$ un graphe quelconque. La probabilité de succès $s(G)$ est minorée par $1 - e^{-E(X(G))}$, où $X(G)$ désigne le nombre de handshake dans G et $E(X(G))$ son espérance.
- Preuve : (voir le tableau).
- Corollaire : La probabilité de succès $s(G)$ est minorée par $1 - e^{-1/2}$
- Question : est-ce que le graphe complet minimise la probabilité de succès ?
- Réponse : OUI (d'après Martin Dietzfelbinger (dans ISAAC 2002) puis Martin Dietzfelbinger et Hisao Tamaki (dans Random Struct. Algorithms 2005)).

- R. Motwani et P. Raghavan, Randomized Algorithms. Cambridge.
- G. Tel, Introduction to Distributed Algorithms, Cambridge Press.
- C. Lavault, Analyse d'algorithmes distribués, Hermès