

Propriétés universelles des algorithmes de recherche combinatoire probabilistes, 2ème partie

Rémi Monasson

LPTENS Paris

et

Christophe Deroulers

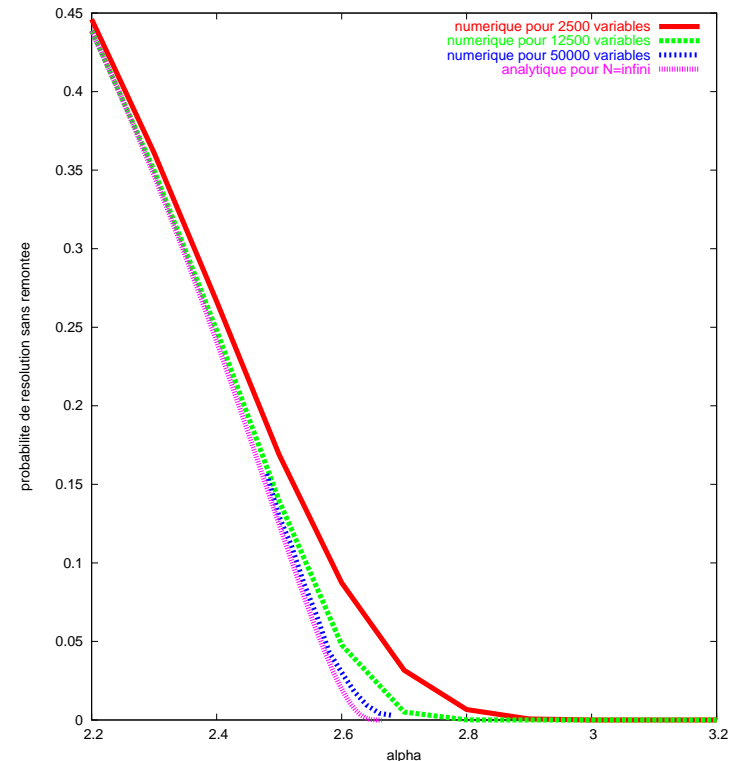
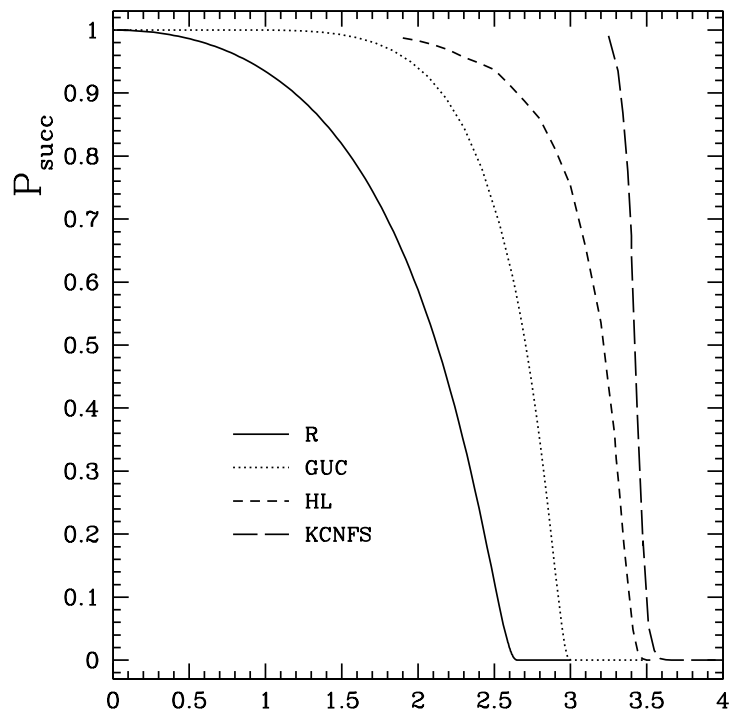
LPTENS Paris

INRIA Rocquencourt, 8 novembre 2004

But de cette partie

Étude des algorithmes gloutons :

- ▶ Calcul de $P(\alpha, N = +\infty)$ pour l'algorithme random ;
- ▶ Calcul de $P(\alpha, N \gg 1)$ pour la classe «P.U.» au voisinage de $\alpha_{\text{poly/exp}}$.



Réduction du problème

Notations

N variables, αN clauses.

$0 \leq T \leq N$: nombre de variables éliminées depuis le début de la résolution.

Un état des αN clauses lors de la résolution, après élimination de T variables sur N :

	VRAI		i		g OU \bar{h}
d OU \bar{g} OU j		g OU h OU \bar{i}			VRAI
	VRAI		VRAI		VRAI

Il y a C_1 1-clauses, C_2 2-clauses, C_3 3-clauses.

Première réduction

Avec cette distribution des instances,

Quand on a moyenné sur les entrées de l'algorithme, la distribution de formules à 1-, 2- et 3-clauses à un instant T donné est uniforme, conditionnée à la donnée de $\vec{C}(T) := (C_3(T), C_2(T), C_1(T))$.

Franco

Matrice d'évolution

On peut écrire de façon exacte les taux de transition de la «chaîne de Markov» $K(\vec{C} \leftarrow \vec{C}', T)$. Exemple de random :

$$\binom{C'_3}{C'_3 - C_3} \left(1 - \frac{3}{N - T}\right)^{C_3} \left(\frac{3}{2(N - T)}\right)^{C'_3 - C_3} \sum_{w_2=0}^{C'_3 - C_3} \binom{C'_3 - C_3}{w_2} \times$$

$$\sum_{z_2=0}^{C'_2} \binom{C'_2}{z_2} \left(1 - \frac{2}{N - T}\right)^{C'_2 - z_2} \left(\frac{1}{N - T}\right)^{z_2} \sum_{w_1=0}^{z_2} \binom{z_2}{w_1} \delta_{C_2 = C'_2 + w_2 - z_2} \times$$

$$\left\{ (1 - \delta_{C'_1=0}) \sum_{e_1=0}^{C'_1-1} \binom{C'_1-1}{e_1} \left(\frac{1}{2(N - T)}\right)^{e_1} \left(1 - \frac{1}{N - T}\right)^{C'_1-1-e_1} \times \right. \\ \left. \delta_{C_1 = C'_1 + w_1 - e_1 - 1} + \delta_{C'_1=0} \delta_{C_1 = w_1} \right\}$$

Deuxième réduction

Il y a concentration de la mesure pour C_2 et C_3 .

Donc on calcule juste les moyennes $\langle C_2(T) \rangle$ et $\langle C_3(T) \rangle$ plutôt que toutes les distributions.

Chao & Franco

Pour random :

$$\begin{aligned}\langle C_3(T+1) \rangle - \langle C_3(T) \rangle &= -\frac{3\langle C_3 \rangle}{N-T} \\ \langle C_2(T+1) \rangle - \langle C_2(T) \rangle &= -\frac{2\langle C_2 \rangle}{N-T} + \frac{1}{2} \frac{3\langle C_3 \rangle}{N-T}\end{aligned}$$

Deuxième réduction

Et pour tous les algorithmes avec Propagation Unitaire,

$$\langle C_3(T + 1) \rangle - \langle C_3(T) \rangle = f_1(\langle C_3 \rangle, \langle C_2 \rangle, N - T)$$

$$\langle C_2(T + 1) \rangle - \langle C_2(T) \rangle = f_2(\langle C_3 \rangle, \langle C_2 \rangle, N - T)$$

car

- ▶ On le vérifie lors d'un pas de Propagation Unitaire ($C_1 > 0$),
- ▶ C'est forcément vrai si $C_1 = 0$,
- ▶ Donc c'est vrai pour la somme des deux contributions.

↪ On peut calculer $\langle C_2(T) \rangle$ et $\langle C_3(T) \rangle$.

Temps discret \rightarrow temps continu

Chao & Franco

$C_2(T), C_3(T)$ varient sur des temps T d'ordre 1, mais $\frac{C_j}{N}$ varient sur des temps T d'ordre N :

$\exists c_2(t)$ et $c_3(t)$ fonctions régulières telles que

$$C_j(T) = Nc_j\left(\frac{T}{N}\right) + O(\sqrt{N})$$

Alors (pour random)

$$\begin{aligned}\frac{dc_3}{dt} &= -\frac{3c_3}{1-t} \\ \frac{dc_2}{dt} &= -\frac{2c_2}{1-t} + \frac{1}{2} \frac{3c_3}{1-t}\end{aligned}$$

et $c_3(t) = \alpha(0)(1-t)^3$, $c_2(t) = \frac{3}{2}\alpha(0)t(1-t)^2$ pour $0 \leq t \leq 1 \dots$

Temps discret \rightarrow temps continu

Chao & Franco

$C_2(T), C_3(T)$ varient sur des temps T d'ordre 1, mais $\frac{C_j}{N}$ varient sur des temps T d'ordre N :

$\exists c_2(t)$ et $c_3(t)$ fonctions régulières telles que

$$C_j(T) = Nc_j\left(\frac{T}{N}\right) + O(\sqrt{N})$$

Alors (pour random)

$$\begin{aligned}\frac{dc_3}{dt} &= -\frac{3c_3}{1-t} \\ \frac{dc_2}{dt} &= -\frac{2c_2}{1-t} + \frac{1}{2} \frac{3c_3}{1-t}\end{aligned}$$

et $c_3(t) = \alpha(0)(1-t)^3$, $c_2(t) = \frac{3}{2}\alpha(0)t(1-t)^2$ pour $0 \leq t \leq 1$...

...tant qu'il n'y a pas de **contradiction** entre deux clauses unitaires.

Contradictions ?

Proba de ne pas remarquer de contradiction de T à $T + 1$:

$$\left(1 - \frac{1}{2(N - T)}\right)^{\max(C_1 - 1, 0)}$$

Contradictions ?

Proba de ne pas remarquer de contradiction de T à $T + 1$:

$$\left(1 - \frac{1}{2(N - T)}\right)^{\max(C_1 - 1, 0)}$$

d'où la proba de ne pas rencontrer de contradiction de 0 à $T = Nt$:

$$\exp\left(-\int_0^t dt \frac{\langle \max(C_1 - 1, 0) \rangle}{2(1 - t)}\right)$$

finie quand $N \rightarrow +\infty$ si C_1 est borné par $O(1)$.

→ il faut étudier la distribution de $C_1(T)$.

Étude de la distribution de C_1

Matrice de transition pour C_1

$$\begin{aligned}
 H_N[C'_1 \leftarrow C_1; T, C_2] = & \sum_{s_2, r_2=0}^{C_2} \binom{C_2}{s_2, r_2} \left(\frac{1}{N-T} \right)^{s_2+r_2} \left(1 - \frac{2}{N-T} \right)^{C_2-s_2-r_2} \times \\
 & \left[\delta_{C_1=0} \delta_{C'_1=r_2} + (1 - \delta_{C_1=0}) \times \right. \\
 & \left. \sum_{s_1=0}^{C_1-1} \binom{C_1-1}{s_1} \left(\frac{1}{2(N-T)} \right)^{s_1} \left(1 - \frac{1}{N-T} \right)^{C_1-1-s_1} \delta_{C'_1=C_1-1-s_1+r_2} \right]
 \end{aligned}$$

Matrice de transition pour C_1

$$H_N[C'_1 \leftarrow C_1; T, C_2] = \sum_{s_2, r_2=0}^{C_2} \binom{C_2}{s_2, r_2} \left(\frac{1}{N-T}\right)^{s_2+r_2} \left(1 - \frac{2}{N-T}\right)^{C_2-s_2-r_2} \times$$

$$\left[\delta_{C_1=0} \delta_{C'_1=r_2} + (1 - \delta_{C_1=0}) \times \right.$$

$$\left. \sum_{s_1=0}^{C_1-1} \binom{C_1-1}{s_1} \left(\frac{1}{2(N-T)}\right)^{s_1} \left(1 - \frac{1}{N-T}\right)^{C_1-1-s_1} \delta_{C'_1=C_1-1-s_1+r_2} \right]$$

- ▶ H ne dépend (explicitement) que de C_1 , C_2 et T .
- ▶ Équation exacte. Contient toute l'information (sauf C_2).
- ▶ Expression valable uniquement pour la distribution «random K-SAT»
- ▶ mais pour tous les algorithmes qui appliquent la **Propagation Unitaire**.

C_1 reste fini si...

$P_N(C_1, T)$: proba que, au temps T , on n'ait pas remarqué de contradiction et qu'il y ait C_1 1-clauses.

C_1 reste fini si...

$P_N(C_1, T)$: proba que, au temps T , on n'ait pas remarqué de contradiction et qu'il y ait C_1 1-clauses.

Hypothèse : $P_N(C_1, T) = p_0(C_1, t) + \frac{1}{N}p_1(C_1, t) + \mathcal{O}(\frac{1}{N^2})$.

...On reporte dans la matrice de transition de C_1 ...

C_1 reste fini si...

$P_N(C_1, T)$: proba que, au temps T , on n'ait pas remarqué de contradiction et qu'il y ait C_1 1-clauses.

Hypothèse : $P_N(C_1, T) = p_0(C_1, t) + \frac{1}{N}p_1(C_1, t) + \mathcal{O}(\frac{1}{N^2})$.

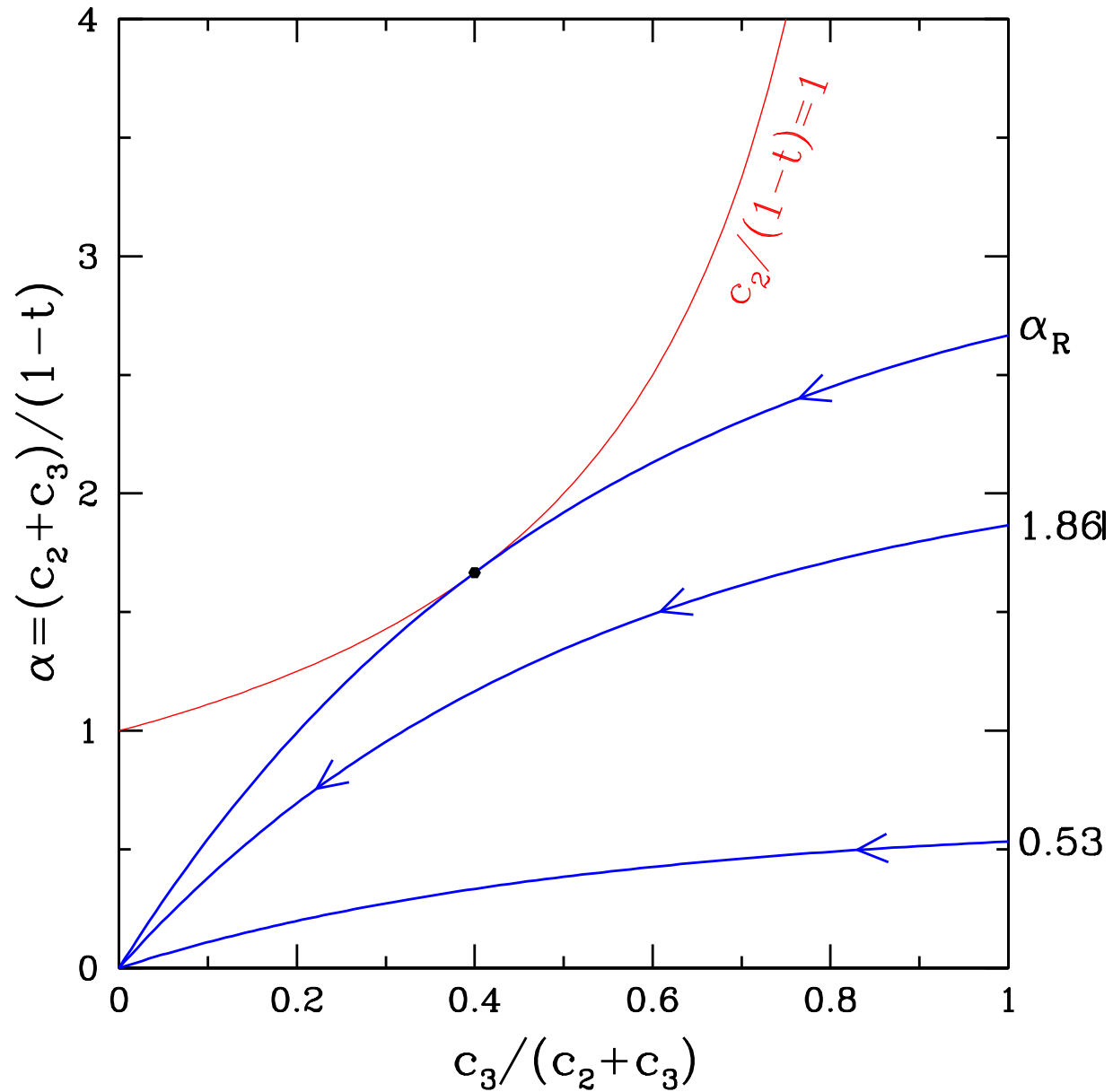
...On reporte dans la matrice de transition de C_1 ...

→ pour C_1 grand, $p_0(C_1, t) \asymp e^{-\rho C_1}$ où

$$\rho \frac{1-t}{c_2} = 1 - e^{-\rho}$$

Tant que $\frac{c_2}{1-t} < 1$, $\rho > 0$ et C_1 est borné p.s. quand $N \rightarrow +\infty$.

Trajectoires de résolution



Probabilité de succès finie

→ Tant que $\alpha < \alpha_{\text{random}} = \frac{8}{3}$, $C_1(T)$ est borné p.s. et on connaît sa distribution en fonction de $c_2(t)$. (Si $\alpha \geq \alpha_{\text{random}}$, C_1 n'est plus toujours borné.)

Probabilité de succès finie

→ Tant que $\alpha < \alpha_{\text{random}} = \frac{8}{3}$, $C_1(T)$ est borné p.s. et on connaît sa distribution en fonction de $c_2(t)$. (Si $\alpha \geq \alpha_{\text{random}}$, C_1 n'est plus toujours borné.)

$$P_{\text{succ}} = \exp \left(- \int_0^1 dt \frac{\langle \max(C_1 - 1, 0) \rangle}{2(1-t)} \right)$$

d'où pour random, $\alpha < \frac{8}{3}$ et $N = +\infty$:

$$\ln(P_{\text{succ}}) = \frac{3\alpha}{16} - \frac{1}{2\sqrt{\frac{8}{3\alpha}-1}} \arctan \left(\frac{1}{\sqrt{\frac{8}{3\alpha}-1}} \right).$$

On peut faire le calcul pour d'autres algorithmes.
Universalité ?

Le régime critique

Préliminaire

Fonction génératrice de C_1 :

$$p_N(T, x) := \sum_{C_1=0}^{+\infty} x^{C_1} P_N(C_1, T)$$

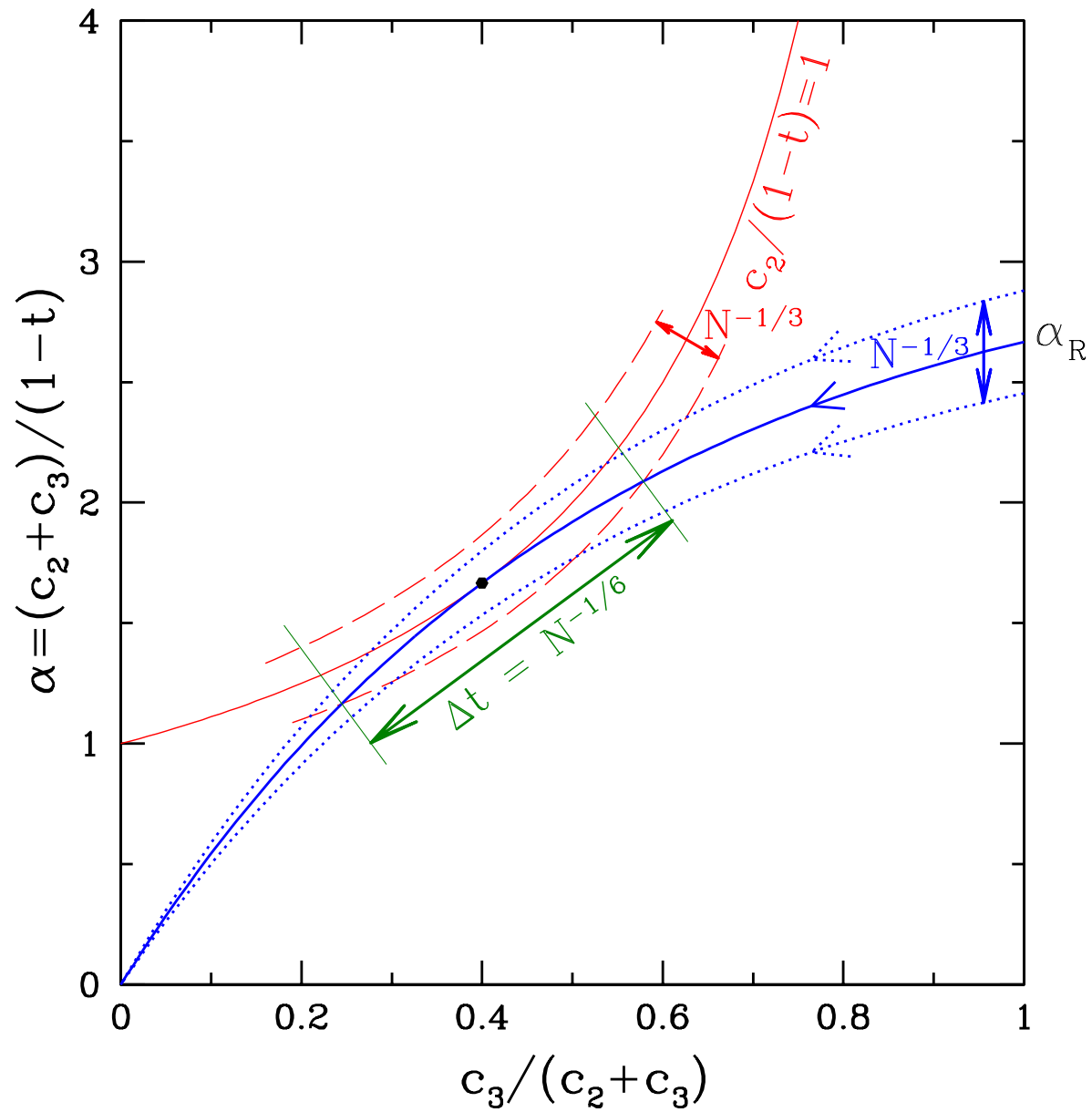
Matrice $H(C'_1 \leftarrow C_1)$ devient :

$$p_N(T+1, x) = \left(1 + \frac{x-1}{N-T}\right)^{C_2(T)} \times \left(\frac{1}{A} p_N(T, A) + \left(1 - \frac{1}{A}\right) p_N(T, 0)\right)$$

avec $A := \frac{1}{2(N-T)} + \left(1 - \frac{1}{N-T}\right)x$.

Exact, contient toute l'information sauf C_2 ; valable pour cette distribution des instances seulement mais pour tous les algorithmes avec Propagation Unitaire.

Zoom in



Zoom in

Zoom in autour du point de tangence $\frac{c_2}{1-t} = 1$. Dans le cas de random :

$$\blacktriangleright \alpha = \frac{8}{3}(1 + \epsilon_0 N^{-\theta})$$

$$\blacktriangleright t = \frac{1}{2}(1 + t_0 N^{-\tau})$$

$$\blacktriangleright C_2(T) = 4T(1 - \frac{T}{N})^2 + \mathcal{O}(\sqrt{N})$$

θ, τ à déterminer pour capturer le 1er ordre non nul.

Hypothèses

Hypothèses sur la distribution de $C_1(T)$ (on ne sait pas tout démontrer ☹) :

- ▶ $P_N(C_1, T) = e^{-N^\lambda \phi(t_0)} F(C_1, t_0)$ avec F normalisée ($F =$ distribution de C_1 conditionnée à l'absence de contradiction remarquée).

Hypothèses

Hypothèses sur la distribution de $C_1(T)$ (on ne sait pas tout démontrer ☹) :

- ▶ $P_N(C_1, T) = e^{-N^\lambda \phi(t_0)} F(C_1, t_0)$ avec F normalisée ($F =$ distribution de C_1 conditionnée à l'absence de contradiction remarquée).
- ▶ La valeur typique de C_1 se comporte comme N^γ ,
 $0 < \gamma < 1$: $F(C_1, t_0) = \frac{1}{N^\gamma} f\left(\frac{C_1}{N^\gamma}\right)$ avec $f(c)$ fonction régulière.

Hypothèses

Hypothèses sur la distribution de $C_1(T)$ (on ne sait pas tout démontrer ☹) :

- ▶ $P_N(C_1, T) = e^{-N^\lambda \phi(t_0)} F(C_1, t_0)$ avec F normalisée ($F =$ distribution de C_1 conditionnée à l'absence de contradiction remarquée).
- ▶ La valeur typique de C_1 se comporte comme N^γ ,
 $0 < \gamma < 1$: $F(C_1, t_0) = \frac{1}{N^\gamma} f\left(\frac{C_1}{N^\gamma}\right)$ avec $f(c)$ fonction régulière.
- ▶ Grande déviation : $F(C_1 = 0, t_0) = \frac{f_0(t_0)}{N^{\gamma_0}}$.

Hypothèses

Hypothèses sur la distribution de $C_1(T)$ (on ne sait pas tout démontrer ☹) :

- ▶ $P_N(C_1, T) = e^{-N^\lambda \phi(t_0)} F(C_1, t_0)$ avec F normalisée ($F =$ distribution de C_1 conditionnée à l'absence de contradiction remarquée).
- ▶ La valeur typique de C_1 se comporte comme N^γ ,
 $0 < \gamma < 1$: $F(C_1, t_0) = \frac{1}{N^\gamma} f\left(\frac{C_1}{N^\gamma}\right)$ avec $f(c)$ fonction régulière.
- ▶ Grande déviation : $F(C_1 = 0, t_0) = \frac{f_0(t_0)}{N^{\gamma_0}}$.

On s'intéresse aux valeurs de C_1 typiques, donc en N^γ , donc dans la fonction génératrice il faut prendre $x = 1 - \frac{x_0}{N^\gamma}$ (zoom in) :

$$p_N(T, x) = e^{-N^\lambda \phi(t_0)} \pi(x_0, t_0)$$

Calculs...

On remplace ces notations et hypothèses dans

$$p_N(T+1, x) = \left(1 + \frac{x-1}{N-T}\right)^{C_2(T)} \times \left(\frac{1}{A}p_N(T, A) + \left(1 - \frac{1}{A}\right)p_N(T, 0)\right)$$

avec $A := \frac{1}{2(N-T)} + \left(1 - \frac{1}{N-T}\right)x$.

θ et τ sont libres. Si on demande que les termes non nuls au premier ordre soient tous du même ordre :

$$\theta = \gamma = \gamma_0 = \frac{1}{3}, \quad \lambda = \tau = \frac{1}{6}.$$

Et

$$-\partial_{t_0} \phi(t_0) \pi(x_0, t_0) = \partial_{x_0} \pi(x_0, t_0) + \pi(x_0, t_0) \left(\frac{x_0^2}{2} + (t_0^2 - \epsilon_0)x_0 \right) - f_0(t_0)x_0$$

En $x_0 = 0 \rightsquigarrow \partial_{t_0} \phi(t_0) = -\partial_{x_0} \pi(0, t_0) = \bar{c}(t_0)$.

EDP pour π

Donc :

$$\partial_{x_0} \pi(x_0, t_0) + \pi(x_0, t_0) \left(\frac{x_0^2}{2} + (t_0^2 - \epsilon_0)x_0 + \bar{c}(t_0) \right) - f_0(t_0)x_0 = 0$$

avec $\bar{c}(t_0)$ moyenne de c .

Remarque : $\pi(x_0, t_0)$, fonction génératrice, est aussi la transformée de Laplace de $f(c = \frac{C_1}{N}, t_0)$.

f régulière \Rightarrow pour $x_0 \rightarrow +\infty$,

$$\pi(x_0, t_0) = 0 \times x_0 + 0 + \mathcal{O}\left(\frac{1}{x_0}\right)$$

EDP pour π

Donc :

$$\partial_{x_0} \pi(x_0, t_0) + \pi(x_0, t_0) \left(\frac{x_0^2}{2} + (t_0^2 - \epsilon_0)x_0 + \bar{c}(t_0) \right) - f_0(t_0)x_0 = 0$$

avec $\bar{c}(t_0)$ moyenne de c .

Remarque : $\pi(x_0, t_0)$, fonction génératrice, est aussi la transformée de Laplace de $f(c = \frac{C_1}{N}, t_0)$.

f régulière \Rightarrow pour $x_0 \rightarrow +\infty$,

$$\pi(x_0, t_0) = 0 \times x_0 + 0 + \mathcal{O}\left(\frac{1}{x_0}\right)$$

d'où les conditions aux limites pour f :

$$f_0(t_0) = \frac{f(0, t_0)}{2}$$

$$\partial_c f(c = 0, t_0) + (t_0^2 - \epsilon_0)f(c = 0, t_0) = 0$$

EDP pour f

$$\partial_{x_0} \pi(x_0, t_0) + \pi(x_0, t_0) \left(\frac{x_0^2}{2} + (t_0^2 - \epsilon_0)x_0 + \bar{c}(t_0) \right) - f_0(t_0)x_0 = 0$$

→ (transformée de Laplace inverse)

$$\frac{1}{2} \frac{\partial^2 f}{\partial c^2} + (t_0^2 - \epsilon_0) \frac{\partial f}{\partial c} + (c - \bar{c})f = 0$$

EDP pour f

$$\partial_{x_0} \pi(x_0, t_0) + \pi(x_0, t_0) \left(\frac{x_0^2}{2} + (t_0^2 - \epsilon_0)x_0 + \bar{c}(t_0) \right) - f_0(t_0)x_0 = 0$$

→ (transformée de Laplace inverse)

$$\boxed{\frac{1}{2} \frac{\partial^2 f}{\partial c^2} + (t_0^2 - \epsilon_0) \frac{\partial f}{\partial c} + (c - \bar{c})f = 0}$$

On résout pour f :

$$f(c, t_0) \propto e^{-(t_0^2 - \epsilon_0)c} \text{Ai}[\sqrt[3]{2}c + z(t_0^2 - \epsilon_0)]$$

où $z(x)$ est la fonction inverse de $x(z) = \sqrt[3]{2} \frac{\text{Ai}'(z)}{\text{Ai}(z)}$.

Résultat final

Rappel : $P_N(C_1, T) = e^{-N^\lambda \phi(t_0)} F(C_1, t_0)$ avec F normalisée ($F =$ distribution de C_1 conditionnée à l'absence de contradiction remarquée).

Donc $\ln P_{\text{succ}} = -N^\lambda \phi(t_0 = +\infty)$ et $\phi(t_0 = -\infty) = 0$.

$\partial_{t_0} \phi(t_0) = \bar{c}(t_0) \rightsquigarrow$

$$-\ln P_{\text{succ}}((1 + \epsilon_0)\alpha_{\text{random}}, N) = N^\lambda \Phi(\epsilon_0 N^\theta)$$

avec

$$\Phi(\epsilon_0) = \frac{1}{4} \int_{-\epsilon_0}^{+\infty} \frac{dx}{\sqrt{\epsilon_0 + x}} [x^2 - 2^{2/3} z(x)]$$

Portée de ce résultat et conclusion

Le calcul ne dépend de l'algorithme qu'à travers C_2 . En fait, si on fait le zoom autour du point de tangence t_A (qui n'intervient pas directement) :

$$\frac{c_2}{1-t} = 1 + b(t - t_A)^2 \text{ pour } t \rightarrow t_A$$

les calculs sont les mêmes, sauf que b a une valeur différente. On trouve

$$\Phi_A(\epsilon_0) = r_A^\Phi \Phi(r_A^\epsilon \epsilon_0)$$

où les r_A sont fonctions de b .

Portée de ce résultat et conclusion

Le calcul ne dépend de l'algorithme qu'à travers C_2 . En fait, si on fait le zoom autour du point de tangence t_A (qui n'intervient pas directement) :

$$\frac{c_2}{1-t} = 1 + b(t - t_A)^2 \text{ pour } t \rightarrow t_A$$

les calculs sont les mêmes, sauf que b a une valeur différente. On trouve

$$\Phi_A(\epsilon_0) = r_A^\Phi \Phi(r_A^\epsilon \epsilon_0)$$

où les r_A sont fonctions de b .

Cas non générique (possible si $K > 3\text{-SAT}$) :

$$\frac{c_2}{1-t} = 1 + b(t - t_A)^{4,6,\dots}$$

Portée de ce résultat et conclusion

Le calcul dépend de la distribution des instances : la fonction Φ peut changer. Mais les exposants sont les mêmes à cause de la robustesse de la distribution des tailles des composantes connexes d'un graphe aléatoire (cf. première partie).