Approximation of NP-hard problems by moderately exponential time algorithms

Vangelis Th. Paschos

http://www.lamsade.dauphine.fr/~paschos

Joint work with Nicolas Bourgeois and Bruno Escoffier

5 février 2009

Plan

- Polynomial inapproximability
- Exact computation with worst-case time-bounds
- Moderately exponential approximation
- MAX INDEPENDENT SET
- MIN VERTEX COVER
- MAX CLIQUE

Approximation ratio

opt(I): the value of an optimal solution of an instance I of a problem Π

$m_{\mathtt{A}}(I,S)$: the value of the solution S computed by an approximation algorithm A on I

Approximation ratio of A

$$\rho_{\mathtt{A}}(I) = \frac{m_{\mathtt{A}}(I,S)}{\mathrm{opt}(I)}$$

The closer the ratio to 1, the better the performance of A

Polynomial inapproximability

- What is an inapproximability result?
- Examples
- A crucial question

Inapproximability result

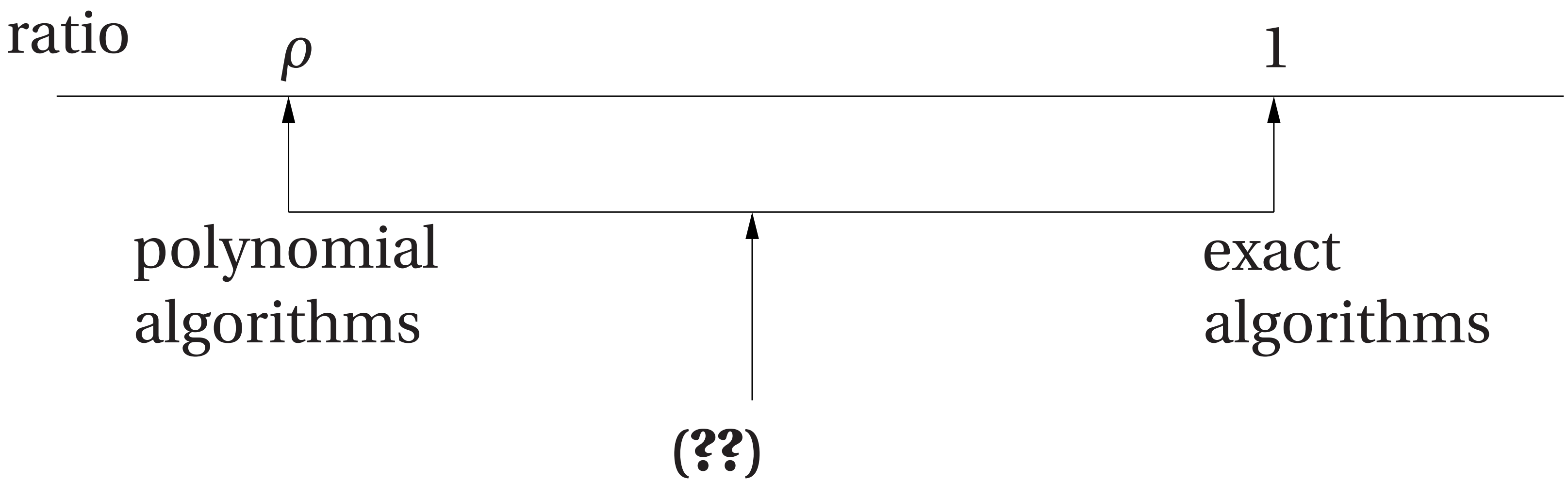
A statement that a problem is inapproximable within ratios better than some approximability level unless something very unlikely happens in complexity theory

- $\boldsymbol{P = NP}$
- *Disproval of the Exponential Time Hypothesis* (problems in **NP** can be solved by slightly superpolynomial algorithms)
- ...

Examples of inapproximability

- MAX INDEPENDENT SET or MAX CLIQUE inapproximable within ratios $\Omega(n^{-1})$
- MIN VERTEX COVER within smaller than 2
- MIN SET COVER within $o(\log n)$
- MIN TSP within better than exponential ratios
- MIN COLORING within $o(n)$
- …

A basic question (goal = max)



What about (**??**)?

Why not taking advantage of the power of modern computers?

For realistic values of n, 1.1^n is not so "worse" than n^5

Exact computation with worst-case guarantees

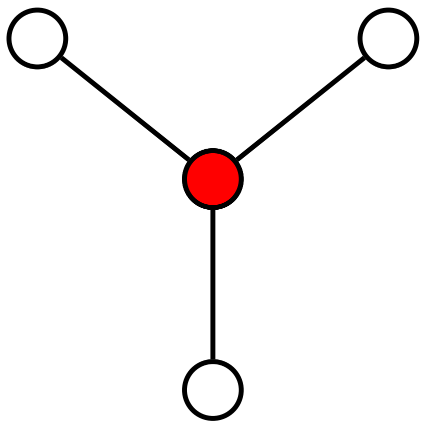
Determine an optimal solution for an ***NP****-hard problem with provably non trivial worst-case time-complexity*

Notation : $O^*(f(n)) = p(n) \times f(n)$ (for a polynomial p)

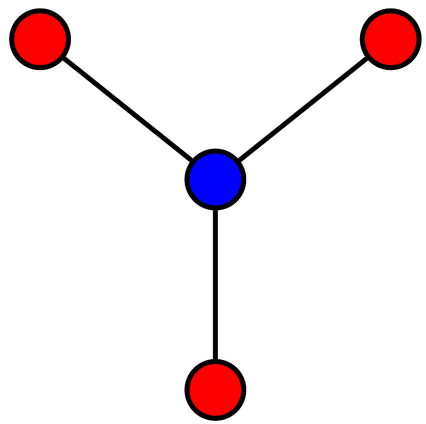
Example: MAX INDEPENDENT SET

- Exhaustively generate any subset of V and get a maximum one among those that are independent sets: $O^*(2^n)$ (trivial exact complexity)
- Find all the maximal independent sets of the input graph: $O^*(1.4422^n)$ (Moon & Moser (1965))

Doing better



(a) 1 vertex fixed



(b) At least 4 vertices fixed

$$T(n) \leqslant T(n-1) + T(n-4) + p(n) \simeq O^*\left(1.385^n\right)$$

Coping with inapproximability: moderately exponential approximation

Approximate optimal solutions of NP-hard problems within ratios “forbidden” to polynomial algorithms and with worst-case complexity provably better than the complexity of an exact computation

Moderately exponential approximation: main stakes

For a problem Π inapproximable within ratios better than r' and solved by an exact algorithm in time $O^*(\gamma^n)$:

- Can we determine an r-approximate solution (r better than r') with complexity essentially better than $O^*(\gamma^n)$?
- More ambitiously, can we do that for any forbidden r?

MAX INDEPENDENT SET

Given a graph, determine a maximum-cardinality set of mutually non-adjacent vertices

- Generate candidate solutions
- Divide and approximate
- Randomization

Something simple: generate candidate solutions

- Generate all the $\sqrt{n}$-subsets of V
- If one of them is independent, then output it
- Else output a vertex at random

Approximation ratio: $n^{-1/2}$
(impossible in polynomial time)

Worst-case complexity: $O^*\left(\binom{n}{\sqrt{n}}\right) \leqslant O^*\left(2^{\sqrt{n}\log n}\right)$
(subexponential)

Divide and approximate - Algorithm `IND_SET`

Set $p/q = r$ (if r fixed, p and q fixed also) and:

- Split G into q induced subgraphs $G_1, \ldots, G_q$ of order n/q
- Build the subgraphs $G'_1, \ldots G'_{C_q^p}$, unions of p subgraphs in $\{G_1, \ldots, G_q\}$ among q
- Optimally solve MAX INDEPENDENT SET in every G'_i
- Output the best of the solutions computed

MAX INDEPENDENT SET Theorem

Assume that an optimal solution for MAX INDEPENDENT SET *can be found in* $O^*(\gamma^n)$
Then, for any fixed p, q, $p < q$, *a* p/q*-approximation can be computed with complexity* $O^*(\gamma^{\frac{p}{q}n})$

It works for every problem defined upon a non-trivial hereditary property

Graph-splitting Corollary

The best independent set computed in the graphs G'_i is an $r = (p/q)$-approximation for MAX INDEPENDENT SET

Sketch of proof (for $p/q = r = 1/2$)

S^*: a maximum independent set in G
S_i: a maximum independent set in G'_i, $i = 1, 2$
Here, G'_i, $i = 1, 2$, the half of G

$$\left|S^* \cap V\left(G'_i\right)\right| \leqslant |S_i| \Rightarrow \left|S^*\right| \leqslant |S_1| + |S_2| \leqslant 2\max\{|S_1|, |S_2|\}$$
$$\Longrightarrow \quad \frac{\max\{|S_1|, |S_2|\}}{|S^*|} \geqslant \frac{1}{2}$$

Complexity: $O^*(\gamma^{n/2})$

Doing better: random splitting

Achieving ratio r with complexity better than $O^*(\gamma^{rn})$?

- Randomly split the graph into subgraphs in such a way that MAX INDEPENDENT SET is to be solved in graphs G'_i of order $r'n$ with $r' < r$
- Compute the probability $\Pr[r]$ that $|S^* \cap V_i| \geqslant r|S^*|$ (r-approximation with probability $\Pr[r]$)
- Repeat splitting $N(r)$ times so that $\Pr[r] \to 1$ (we have an r-approximation with probability ~ 1 in time $N(r)\gamma^{r'n}$)
- Prove that $N(r)\gamma^{r'n} < \gamma^{rn}$

Example: random splitting into 2 subgraphs

- Fix a ratio $r > 1/2$ to attain
- Split the input-graph into two parts of size $n/2$
- Compute the probability $\Pr[r]$ that $|S^* \cap V_1| \geqslant r|S^*|$
- Repeat the 2-splitting $N(r)$ times so that $\Pr[r] \to 1$
- Complexity of the game random splitting - MAX INDEPENDENT SET solving: $N(r)\gamma^{n/2}$
- Complexity of deterministic solution γ^{rn}
- And the winner is ...

The randomization MAX INDEPENDENT SET theorem

For any $r < 1$ and for any β, $r/2 \leqslant \beta \leqslant r$, it is possible to determine with probability $1 - \exp\{-cn\}$ (for some constant c), an r-approximation for MAX INDEPENDENT SET, with running time $O^*(K_n \gamma^{\beta n})$, where:

$$K_n = \frac{n\binom{n}{n/2}}{\binom{\beta n}{rn/2}\binom{n-\beta n}{((1-r)n/2)}}$$

Ratio	Deterministic	Randomized
0.1	1.017^n	1.015^n
0.2	1.034^n	1.031^n
0.3	1.051^n	1.047^n
0.4	1.068^n	1.063^n
0.5	1.086^n	1.080^n
0.6	1.104^n	1.098^n
0.7	1.123^n	1.117^n
0.8	1.142^n	1.136^n
0.9	1.161^n	1.157^n

$\gamma = 1.18$ (Robson (2001))

MIN VERTEX COVER

Given a graph, determine a minimum-cardinality set of vertices that "touch" all the edges

- On the polyhedron of MIN VERTEX COVER
- Aproximation transfer from MAX INDEPENDENT SET to MIN VERTEX COVER
- Using parameterized divide and approximate
- Parameterized approximation

MIN VERTEX COVER

$$\begin{cases} \min & \sum_{v_i \in V} x_i \\ & x_i + x_j \geqslant 1 \quad \forall \left(v_i, v_j\right) \in E \\ & x_i \in \{0,1\} \quad \forall v_i \in V \end{cases}$$

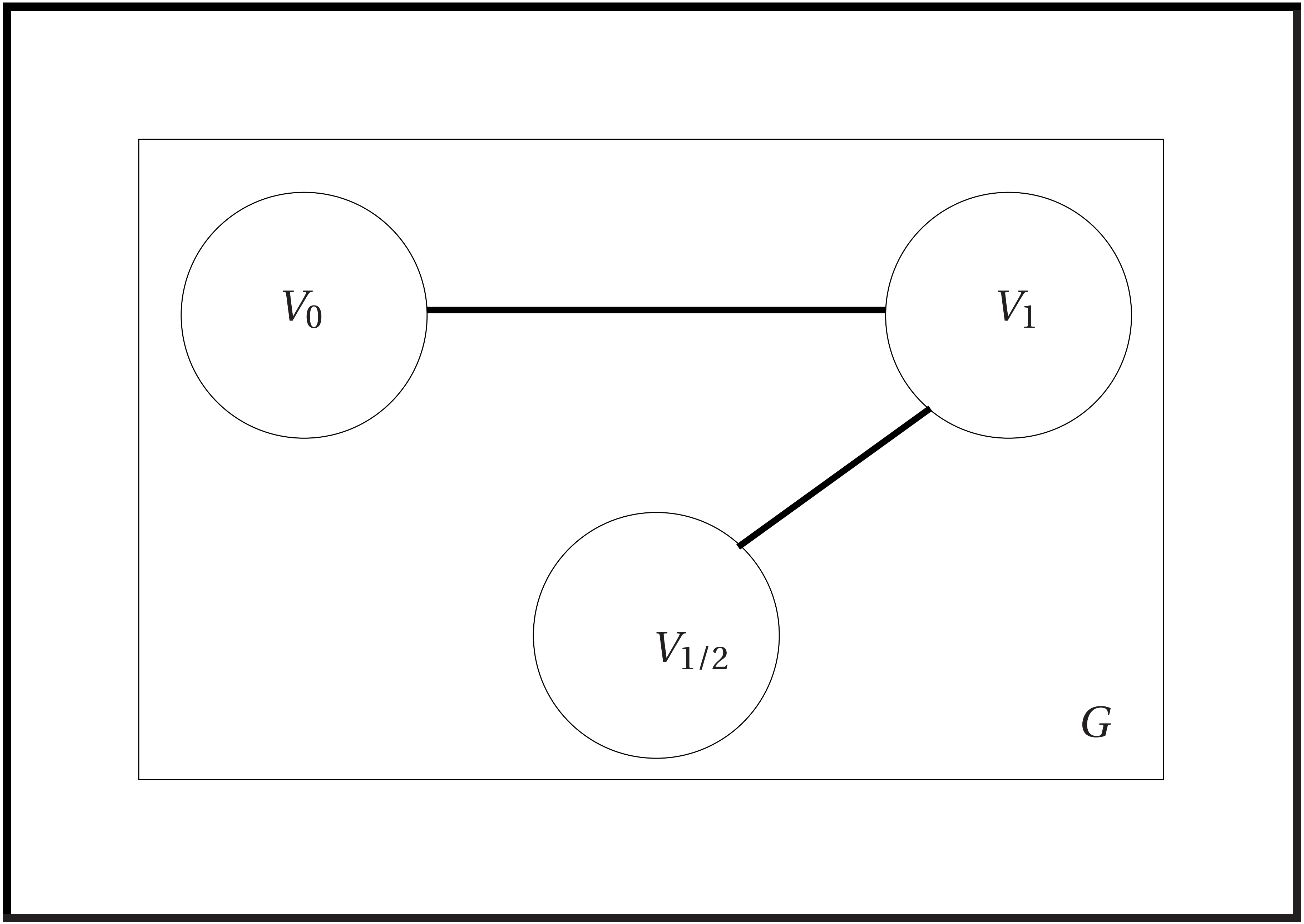
MIN VERTEX COVER-R

$$\begin{cases} \min & \sum_{v_i \in V} x_i \\ & x_i + x_j \geqslant 1 \quad \forall \left(v_i, v_j\right) \in E \\ & x_i \in [0,1] \quad \forall v_i \in V \end{cases}$$

A basic theorem (Nemhauser & Trotter (1975))

The basic optimal solution of MIN VERTEX COVER-R is semi-integral, i.e., it assigns to the variables values from $\{0, 1, 1/2\}$
If V_0, V_1 and $V_{1/2}$ are the subsets of V associated with 0, 1 et 1/2, respectively, then there exists a minimum vertex cover C^* such that:

1. $V_1 \subseteq C^*$
2. $V_0 \subseteq S^* = V \setminus C^*$ (a maximum independent set associated with C^*)


V_0
V_1
$V_{1/2}$
G

Corollaries

1. *The graph $G' = G[V \setminus (V_0 \cup V_1)] = G[V_{1/2}]$ has:*

$$\begin{aligned} |S^*| &\leqslant \frac{|V_{1/2}|}{2} \\ |C^*| = |V \setminus S^*| &\geqslant \frac{|V_{1/2}|}{2} \end{aligned}$$

2. ***Modulo some preprocessing of G, we can reason with respect to G' and assume that $|S^*| \leqslant n/2$ and $|C^*| \geqslant n/2$***

Approximation Transfer Theorem

If S is an r-approximate solution for MAX INDEPENDENT SET, then $C = V \setminus S$ is a $(2-r)$-approximation for MIN VERTEX COVER

C^*: an optimal vertex cover of G
C: an approximate vertex cover

$$\frac{|C|}{|C^*|} = \frac{n - |S|}{n - |S^*|} \leqslant \frac{n - r\,|S^*|}{n - |S^*|} = \frac{1 - r\,\frac{|S^*|}{n}}{1 - \frac{|S^*|}{n}}$$

By first Nemhauser & Trotter corollary:

$$\frac{|S^*|}{n} \leqslant \frac{1}{2} \Longrightarrow \frac{|C|}{|C^*|} \leqslant 2 - r$$

A first MIN VERTEX COVER corollary . . .

A $(2-r)$-approximation for MIN VERTEX COVER can be computed in $O^*(\gamma^{rn})$, for any r

... and a question

Can we achieve ratio $2-r$ in time better than $O^*(\gamma^{rn})$?

A useful parameterized result

> There exists an optimal algorithm `OPT_VC` for MIN VERTEX COVER that, for any $k \leqslant n$, determines if G contains a vertex cover of size k or not and, if yes, it computes it with complexity $O^*(\delta^k)$, $\delta < 2$

$\delta = 1.2852$ (Chen, Kanj & Jia (2001))

First Lemma (small independent sets)

If, for some $\lambda < 1/2$, $|S^*| \leqslant \lambda n$, then a $(2-r)$-approximation of MIN VERTEX COVER *can be found in $O^*\left(\gamma^{(2-r-\frac{1-r}{\lambda})n}\right)$, for any r*

Remark

For $\lambda < 1/2$, $2 - r - \frac{1-r}{\lambda} < r$

From $\frac{|C|}{|C^*|} < \frac{1-r\frac{|S^*|}{n}}{1-\frac{|S^*|}{n}}$, setting $\lambda = |S^*|/n$:

$$\frac{|C|}{|C^*|} \leqslant \frac{1-r\lambda}{1-\lambda} \overset{\lambda<1/2}{<} 2-r$$

So, for $\lambda < 1/2$, ratio $2-r$ is get by approximately solving MAX INDEPENDENT SET with ratio $r' < r$ verifying:

$$2-r = \frac{1-r'\lambda}{1-\lambda} \Longrightarrow r' = 2-r-\frac{1-r}{\lambda}$$

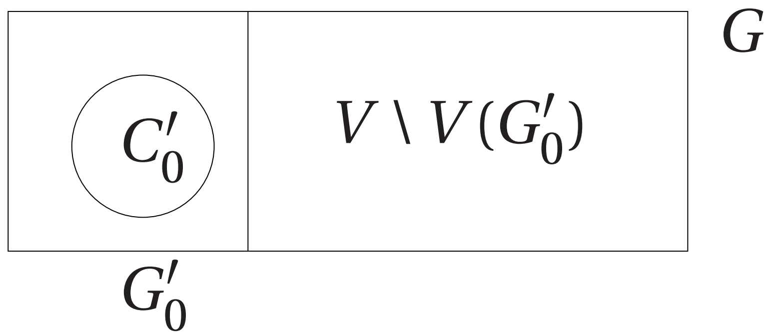
Complexity: $O^*\left(\gamma^{r'n}\right) = O^*\left(\gamma^{\left(2-r-\frac{1-r}{\lambda}\right)n}\right) < O^*\left(\gamma^{rn}\right)$

Second Lemma (large independent sets, i.e., small vertex covers)

If, for some $\lambda < 1/2$, $|S^*| \geqslant \lambda n$, *then a* $(2-r)$*-approximation of* MIN VERTEX COVER *can be found in* $O^*\left(\delta^{r(1-\lambda)n}\right)$

Set $r = p/q$ and run algorithm `VERTEX_COVER`:

1. Split the graph as in the two first steps of `IND_SET`
2. for $i = 1, \dots, C_q^p$, run `OPT_VC`$(G'_i, (1-\lambda)rn)$ and store the best (denoted by C'_0) among the covers that are $\leqslant (1-\lambda)rn$ (if any)
3. if such a cover C'_0 has been computed in Step 2 for a graph G'_0, then output $C = C'_0 \cup (V \setminus V(G'_0))$, else exit



If $|S^*| \geqslant \lambda n$, then there exists a graph G'_0 where a vertex cover C'_0 satisfying $|C'_0| \leqslant (1-\lambda) r n$ has been computed in step 2

$\exists G'_0$ where max independent set S'_0 verifies $\left|S'_0\right| \geqslant r|S^*| \geqslant r\lambda n$ (Graph-splitting Corollary)

Hence $\left|C'_0\right| = \left|V\left(G'_0\right) \setminus S'_0\right| \leqslant rn - r\lambda n = (1-\lambda) r n$

$C = V \setminus S'_0 = C'_0 \cup (V \setminus V(G'_0))$ is a $(2-r)$-approximation for MIN VERTEX COVER (Approximation Transfer Theorem)

Complexity of VERTEX_COVER: $O^*\left(\delta^{r(1-\lambda)n}\right)$

The overall algorithm APPROX_VERTEX_COVER

- Fix an r and determine $\lambda = \lambda(r)$ satisfying $\gamma^{2-r-\frac{1-r}{\lambda}} = \delta^{(1-\lambda)r}$
- Preprocess G
- Set $2 - r - \frac{1-r}{\lambda} = \frac{p}{q}$ and compute $C_0 = V \setminus \mathtt{IND\_SET}(G)$ (First Lemma)
- Set $r = p/q$ and compute $C = \mathtt{VERTEX\_COVER}(G)$ (Second Lemma)
- Output the best among C_0 and C

MIN VERTEX COVER Theorem

For any $r < 1$, MIN VERTEX COVER *can be solved approximately within ratio* $2-r$ *and with time-complexity* $O^*\left(\gamma^{\left(2-r-\frac{1-r}{\lambda}\right)n}\right) \leqslant O^*\left(\gamma^{rn}\right)$

Randomization also works for MIN VERTEX COVER

Parameterized approximation

For any $r = p/q \in \mathbb{Q}$, if there exists a vertex cover of size $\leqslant k$, it is possible to determine a $(2-r)$-approximation of it in time $O^*(\delta^{rk})$

1. Split the graph as in the two first steps of `IND_SET` and set $k = 1$
2. Run `OPT_VC` in G'_i in order to compute a minimum vertex cover of size k; if impossible, repeat step 2 with $k = k + 1$; let C'_{i^*} be a smallest vertex cover so-computed (G'_{i^*} the corresponding graph)
3. Output $C = C'_{i^*} \cup (V \setminus V(G'_{i^*}))$

Complexity: $O^*\left(\delta^{\left|C'_{i^*}\right|}\right)$

$$\begin{aligned} \alpha\left(G'_{i^*}\right) &\geqslant \frac{p}{q}\alpha(G) \text{ (Graph-splitting Corollary)} \\ \left|C'_{i^*}\right| &\leqslant \frac{p}{q}n - \alpha\left(G'_{i^*}\right) \leqslant \frac{p}{q}(n-\alpha(G)) \\ &= \frac{p}{q}\left|C^*\right| = r\left|C^*\right| \end{aligned}$$

$$|C| = \left|C'_{i^*}\right| + \left(1-\frac{p}{q}\right)n \leqslant \frac{p}{q}\left|C^*\right| + n\left(1-\frac{p}{q}\right)$$

Since $|C^*| \geqslant n/2$ (Nemhauser & Trotter Corollary 2):

$$\frac{|C|}{|C^*|} \leqslant 2\left(1-\frac{p}{q}\right) + \frac{p}{q} = 2 - \frac{p}{q} = 2 - r$$

Complexity: $O^*\left(\delta^{r|C^*|}\right)$

Ratio	IND_SET	APPROX_VERTEX_COVER	Parameterized
1.9	1.017^n	1.013^n	1.025^k
1.8	1.034^n	1.026^n	1.051^k
1.7	1.051^n	1.039^n	1.077^k
1.6	1.068^n	1.054^n	1.104^k
1.5	1.086^n	1.069^n	1.131^k
1.4	1.104^n	1.086^n	1.160^k
1.3	1.123^n	1.104^n	1.189^k
1.2	1.142^n	1.124^n	1.218^k
1.1	1.161^n	1.148^n	1.249^k

$\gamma = 1.18$, $\delta = 1.28$

MAX CLIQUE

Given a graph, determine a maximum-size complete induced subgraph

For the efficient approximation of MAX CLIQUE, parameter γ is the same as for MAX INDEPENDENT SET
The exponent for MAX CLIQUE is the maximum degree Δ of the input-graph

- Split G into n induced subgraphs $G_i = G[\{v_i\} \cup \Gamma(v_i)]$
- Solve MAX INDEPENDENT SET in $\bar{G}_i$ (recover cliques in G_i's)
- Output the best of the solutions computed

Hint: any clique is a subset of the neighborhood of a vertex

This is a polynomial reduction MAX CLIQUE $\leq$ MAX INDEPENDENT SET that:

1. transforms approximation ratios functions of n for MAX INDEPENDENT SET into ratios functions of Δ for MAX CLIQUE
2. solves MAX CLIQUE in n graphs of order at most $\Delta + 1$ instead of solving it in one graph of order n

Further research

- Devise proper approximability preserving reductions
- Devise new methods proper to this type of approximation
- Mix polynomial approximation and exact computation
- What about approximation by **sub**exponential algorithms?

What about MIN COLORING, MIN TSP, MIN INDEPENDENT DOMINATING SET, ...?

MERCI