Co-nondeterminism in compositions: A kernelization lower bound for a Ramsey-type problem

Stefan Kratsch

September 03, WorKer 2011, Vienna

Universiteit Utrecht

Introduction

Ramsey(k)
Input: A graph G and an integer k.
Parameter: k.
Question: Does G contain an independent set or a clique of size at least k?

Brought to general attention by Rod Downey at WorKer 2010 in Leiden. He asked whether the problem admits a polynomial kernel.

FPT: if $n \geq R(k,k)$ (Ramsey number) then answer YES, else use brute force ($R(k,k) < 4^k$)

Motivation

- spin-off of a classical problem
- a polynomial kernel would speed up computation of Ramsey numbers: essentially replacing brute force on c^k vertices by brute force on $poly(k)$ vertices
- seems to resist standard techniques for upper and lower bounds
- $$$...

Ramsey Numbers

- $\mathbf{R(\ell_1, \ell_2)}$: largest number of vertices among graphs G that contain no ℓ_1-independent set or ℓ_2-clique
- $\mathbf{R(\ell)} := R(\ell, \ell)$
- explicit values are only known for small ℓ (essentially by brute force computation)
- $R(\ell) \sim c^\ell$ (there are exponential upper and lower bounds)

Outline

Outline

A simple composition for Ramsey(k)

- given t instances $(G_1, k), \ldots, (G_t, k)$
- we construct (G', k') with
  - (G', k') YES iff at least one (G_i, k) is YES
  - $k' \in \mathcal{O}(t^{1/2} k)$
- thus Ramsey(k) has no $\mathcal{O}(k^{2-\epsilon})$ kernel unless PH collapses

  [Dell, van Melkebeek 2010 & Hermelin, Wu 2011]

Improvement version

Improvement Ramsey(k)
Input: A graph G and an integer k. Two vertex sets I and K of size $k-1$ each which induce an independent set and a clique in G.
Parameter: k.
Question: Does G contain an independent set or a clique of size at least k?

We will simply continue to call it Ramsey(k). It is straightforward to reduce between the two versions.

The construction

- w.l.o.g. $t = \ell^2$
- group the t instances into ℓ groups of size ℓ each
- let G' contain copies of $G_1, \ldots, G_t$
- add all edges between vertices of G_i and G_j in G' if they are in the same group
- let $k' = \ell(k-1)+1$ thus $k' \in \mathcal{O}(t^{1/2}k)$

note: adjacency between the graphs $G_1, \ldots, G_t$ can be described by a host graph H: a disjoint union of ℓ cliques of size ℓ each

Universiteit Utrecht

Some observations I

- cliques in G' can use vertices from only one group, i.e., from at most ℓ graphs
- independent sets in G' can use vertices from at most one graph per group, i.e., from at most ℓ graphs
- thus a clique of size $\ell(k-1)+1$ must contain at least k vertices from a single G_i
- ditto for independent sets

thus if (G', k') is YES then at least one (G_i, k) is YES

Some observations II

- if some G_i contains a k-clique, then it can be extended by $k-1$ vertices from each other graph in its group in G'
- we get a clique of size $k + (\ell - 1)(k - 1) = \ell(k - 1) + 1$
- similarly for a k-independent set in some G_i
- it is crucial here that we have the improvement version

if some (G_i, k) is YES then (G', k') is YES

We get a composition with dependence of $t^{1/2}$ on t, excluding kernels of size $\mathcal{O}(k^{2-\epsilon})$.

Why did it work...

...and how can we do better?

- in the host graph H (recall: disj. union of ℓ many ℓ-cliques):
  - there are no cliques or independent sets of size $\ell + 1$
  - each vertex is in a clique and an independent set of size ℓ
- $\ell \in \mathcal{O}(t^{1/2})$
- thus arranging and connecting the t instances according to H we get a composition with $\mathcal{O}(t^{1/2})$ dependence on t

To exclude polynomial kernels we need $\ell \in t^{o(1)}$. Unfortunately no deterministic constructions of such graphs are known. (There is work on Ramsey graphs, but they don't include the covering property.)

Outline

Universiteit Utrecht

Co-nondeterministic composition

Let $\mathcal{Q} \subseteq \Sigma^* \times \mathbb{N}$.

coNP-composition for $\mathcal{Q}$: co-nondeterministic algorithm C
input: t instances $(x_1, k), \ldots, (x_t, k) \in \Sigma^* \times \mathbb{N}$
time: polynomial in $\sum_{i=1}^{t} |x_i|$
output: on each computation path an instance (y, k') with $k' \leq t^{o(1)} poly(k)$ such that:

1. if at least one (x_i, k) is YES then each computation path ends with the output of a YES-instance (y, k')
2. if all (x_i, k) are NO then at least one computation path ends with the output of a NO-instance

new: co-nondeterminism, $t^{o(1)}$ dependence on t

Consequence of a coNP-composition

Theorem: If $\mathcal{Q} \subseteq \Sigma^* \times \mathbb{N}$ has a coNP-composition then it admits no polynomial kernelization unless NP $\subseteq$ coNP/poly.

Proof: This follows straightforwardly from the Complementary Witness Lemma [Dell & van Melkebeek 2010].

key: coNP-kernelization & coNP-composition give oracle communication protocol with co-nondeterministic first player

Outline

Universiteit Utrecht

We need better host graphs

- we need a host graph H on t vertices and $\ell \in t^{o(1)}$ such that:
  - H contains no independent set and no clique of size $> \ell$
  - each vertex of H is contained in an independent set and a clique both of size ℓ
- combining t instances according to H will then give a composition
- we will use co-nondeterminism to find such graphs

note: $\alpha(H) = \ell$ cannot be verified, so we will have to cope with graphs H not fulfilling all properties

Making our lives a bit easier

- it suffices if each vertex of H is in a clique or an independent set of size ℓ
- by a simple transformation $G_i \mapsto G_i'$ we get

$$\begin{aligned} & G_i \text{ has a } k\text{-clique or a } k\text{-independent set} \\ \Leftrightarrow\ & G_i' \text{ has a } 2k-1\text{-clique and a } 2k-1\text{-indepenent set} \end{aligned}$$

- it can be seen that embedding graphs G_i' in the relaxed host graph suffices

Ramsey numbers have useful gaps

Lemma: For every integer $t > 3$ there is an integer $\ell \in \{1, \ldots, 8 \log t\}$ such that $R(\ell + 1) > R(\ell) + t$.

Proof (sketch): If no integer $\ell \in \{1, \ldots, 8 \log t\}$ works, then $R(8 \log t)$ would be smaller than known lower bounds.

Thanks to Pascal Schweitzer for the lemma and advice regarding Ramsey numbers.

Finding a host graph

let an integer t be given

- guess smallest $\ell \in \{1, \ldots, 8 \log t\}$ with $R(\ell + 1) > R(\ell) + t$
- guess T such that $T = R(\ell) + t$

there is a graph on T vertices which has no clique or independent set greater than ℓ

- guess a graph H on T vertices

next: covering at least t vertices of H by independent sets and cliques

Partially covering H

assume that we have a graph H with $R(\ell) + t$ vertices

- among any $R(\ell)$ vertices of H there must be an independent set or a clique of size ℓ
- thus there must be a set of (at most t) cliques and independent sets that covers at least t vertices of H
- such a cover can be guessed and verified; on a failure return YES
- let H' be a subgraph of H on at least t vertices, such that all vertices of H' are covered
- use H' as a host graph and return the obtained instance (G', k')

Wrap-Up / Proof sketch

given t instances $(G_1, k), \ldots, (G_t, k)$ of (improvement) Ramsey(k)

- transform to simpler instances $(G_1', 2k-1), \ldots, (G_t', 2k-1)$ for which relaxed host graph suffices
- co-nondeterministically search for a host graph H'
- each computation path returns YES or an instance (G', k')
- in the latter case the used host graph H' is always covered
- there is at least one c-path where H' has no clique or independent set of size $> \ell \in \mathcal{O}(\log t)$

from these facts, we easily get the following:

Theorem: Ramsey(k) has a coNP-composition and hence does not admit a polynomial kernel unless NP $\subseteq$ coNP/poly.

Outline

Universiteit Utrecht

Conclusion

- Ramsey(k) does not admit a polynomial kernel unless NP $\subseteq$ coNP/poly
- Ramsey numbers are the key to both FPT and kernel lower bound for Ramsey(k)
- co-nondeterministic compositions may help for other problems with open existence of polynomial kernels
- is there more to be gained from the $t^{o(1)}$ dependence on t or is $\log t$ all we ever need?

Thank you