Lecture notes for Apr 12, 2023
2-Disjoint Path Problem, FPT and kernelization

Chun-Hung Liu

April 12, 2023

Recall the following notions: Let G be a simple graph. Let x_1, x_2, y_1, y_2 be distinct vertices. Then we say that (x_1, x_2, y_1, y_2) is *feasible* in G if there exist two disjoint paths P_1 and P_2 in G such that for every $i \in [2]$, the ends of P_i are x_i and y_i; otherwise, we say that (x_1, x_2, y_1, y_2) is *infeasible* in G. For a separation (A, B) of G of order at most three, the *3-reduction of* (A, B) is the graph obtained from $G[A]$ by adding edges such that A becomes a clique. Let x_1, x_2, y_1, y_2 be distinct vertices of G. A *3-reduction of* (x_1, x_2, y_1, y_2) *(for* G*)* is the 3-reduction of a separation (A, B) of G of order at most three with $\{x_1, x_2, y_1, y_2\} \subseteq A$ and $B - A \neq \emptyset$ such that either $|A \cap B| \leq 1$, or there exists no separation (A', B') of order at most two such that $A' \supseteq A$ and $B' - A' \neq \emptyset$. A *full 3-reduction of* (x_1, x_2, y_1, y_2) *(for* G*)* is a 3-irreducible graph G' with respect to (x_1, x_2, y_1, y_2) such that there exists a sequence of graphs $G_1, G_2, ..., G_t$ such that $G_1 = G$, $G_t = G'$, and G_{i+1} is a 3-reduction of (x_1, x_2, y_1, y_2) for G_i for every $1 \leq i \leq t-1$.

We prove the following theorem that characterizes the feasibility.

Theorem 1 (Seymour; Thomassen) *Let G be a graph. Let x_1, x_2, y_1, y_2 be distinct vertices of G. Then (x_1, x_2, y_1, y_2) is infeasible in G if and only if for every full 3-reduction G' of (x_1, x_2, y_1, y_2) for G, the graph $G' + x_1x_2y_1y_2x_1$ can be drawn in the plane such that the 4-cycle $x_1x_2y_1y_2x_1$ bounds the outer face.*

Proof. ($\Leftarrow$) Proved in the previous lecture.

(⇒) We prove it by induction on $|V(G)|$. The case $|V(G)| = 4$ is obvious. So we may assume $|V(G)| \geq 5$ and the theorem holds if $|V(G)|$ is smaller.

Suppose to the contrary that $G + x_1x_2y_1y_2x_1$ cannot be drawn in the plane such that the 4-cycle $x_1x_2y_1y_2x_1$ bounds the outer face.

We prove that $|V(G')| = |V(G)|$, and hence $G = G'$, in the previous lecture.

Claim 1: There exists no separation (A, B) of G of order at most three such that $\{x_1, x_2, y_1, y_2\} \subseteq A$ and $B - A \neq \emptyset$.

Proof of Claim 1: Proved in the previous lecture. □

Claim 2: If (A, B) is a separation of G of order at most four such that $\{x_1, x_2, y_1, y_2\} \subseteq A$ and $B - A \neq \emptyset \neq A - B$, then $|A \cap B| = 4$, $|B - A| = 1$, and the unique vertex in $B - A$ is adjacent to all four vertices in $A \cap B$.

Proof of Claim 2: Let (A, B) be a separation of G of order at most four such that $\{x_1, x_2, y_1, y_2\} \subseteq A$ and $B - A \neq \emptyset \neq A - B$. We choose such (A, B) such that B is maximal. By Claim 1, every component of $G[B] - A$ is adjacent to all vertices in $A \cap B$, so for every pair of vertices $u, v \in A \cap B$, there exists a path in $G[B]$ between u and v internally disjoint from $A \cap B$.

By Claim 1, (A, B) has order four. By Claim 1 and Menger's theorem, there exist four disjoint paths $P_{x_1}, P_{x_2}, P_{y_1}, P_{y_2}$ in $G[A]$ from $\{x_1, x_2, y_1, y_2\}$ to $A \cap B$. Denote $A \cap B$ by $\{x'_1, x'_2, y'_1, y'_2\}$, and we may assume that for every $z \in \{x_1, x_2, y_1, y_2\}$, P_z is between z and z'. Since (x_1, x_2, y_1, y_2) is infeasible in G, (x'_1, x'_2, y'_1, y'_2) is infeasible in $G[B]$. Since $A - B \neq \emptyset$, $|V(G[B])| < |V(G)|$. So by the induction hypothesis, for every full 3-reduction H_B of (x'_1, x'_2, y'_1, y'_2) for $G[B]$, the graph $H_B + x'_1x'_2y'_1y'_2x'_1$ can be drawn in the plane such that the 4-cycle $x'_1x'_2y'_1y'_2x'_1$ bounds the outer face.

Note that if there exists a separation (A', B') of $G[B]$ of order at most three such that $\{x'_1, x'_2, y'_1, y'_2\} \subseteq A'$ and $B' - A' \neq \emptyset$, then the separation $(A' \cup A, B')$ is a separation of G of order at most three such that $\{x_1, x_2, y_1, y_2\} \subseteq A \cup A'$ and $B' - (A \cup A') \neq \emptyset$, contradicting Claim 1. Hence the only full 3-reduction of (x'_1, x'_2, y'_1, y'_2) for $G[B]$ is $G[B]$ itself. That is, $G[B] + x'_1x'_2y'_1y'_2x'_1$ can be drawn in the plane such that the 4-cycle $x'_1x'_2y'_1y'_2x'_1$ bounds the outer face.

Let $G_1 = G[A] + x'_1x'_2y'_1y'_2x'_1$. We first assume that (x_1, x_2, y_1, y_2) is infeasible in G_1. Since $B - A \neq \emptyset$, $|V(G_1)| < |V(G)|$. So by the induction hypothesis, $G_1 + x_1x_2y_1y_2x_1$ can be drawn in the plane such that $x_1x_2y_1y_2x_1$ bounds the outer face. We further choose the drawing such that $x'_1x'_2y'_1y'_2x'_1$ bounds a face if possible. If $x'_1x'_2y'_1y'_2x'_1$ bounds a face in this drawing, then we can combine the drawing of $G_1 + x_1x_2y_1y_2x_1$ and

$G[B]+x_1'x_2'y_1'y_2'x_1'$ to obtain a plane drawing of $G+x_1x_2y_1y_2x_1$ such that $x_1x_2y_1y_2x_1$ bounds the outer face, and we are done. So we may assume that $x_1'x_2'y_1'y_2'x_1'$ does not bound a face and hence is a separating cycle. This implies that if $\{x_1,x_2,y_1,y_2\}=\{x_1',x_2',y_1',y_2'\}$, then (x_1,x_2,y_1,y_2) is feasible in G, a contradiction. So $\{x_1,x_2,y_1,y_2\}\neq\{x_1',x_2',y_1',y_2'\}$. Hence there exists a separation (A',B') of G of order four such that $A'\cap B'=\{x_1',x_2',y_1',y_2'\}$, $\{x_1,x_2,y_1,y_2\}\subset A'$ and $B'\supset B$, contradicting the choice of (A,B).

So we may assume that (x_1,x_2,y_1,y_2) is feasible in G_1. So there exist two disjoint paths P_1 and P_2 in G_1, where P_1 is between x_1 and y_1, and P_2 is between x_2 and y_2. Since (x_1,x_2,y_1,y_2) is infeasible in G_1, and for every pair of vertices $u,v\in A\cap B$, there exists a path in $G[B]$ between u and v internally disjoint from $A\cap B$, we know that both P_1 and P_2 contains edges in the 4-cycle $x_1'x_2'y_1'y_2'x_1'$. Since P_1 and P_2 are disjoint, each P_i contains exactly one edge e_i and two vertices in the 4-cycle $x_1'x_2'y_1'y_2'x_1'$. Then $(P_1-e_1)\cup(P_2-e_2)$ is a union of four disjoint paths in G from $\{x_1,x_2,y_1,y_2\}$ to $\{x_1',x_2',y_1',y_2'\}$.

For $i\in[2]$, let u_i and v_i be the ends of e_i. Since e_1 and e_2 are edges in the 4-cycle C_B bounding the outer face of a plane drawing of $G[B]+x_1'x_2'y_1'y_2'x_1'$, we may assume by symmetry that u_1,v_1,u_2,v_2 appear in C_B in the order listed. By the planarity, there exist no two disjoint paths in $G[B]$, where one is between u_1 and u_2, and the other is between v_1 and v_2. Since (x_1,x_2,y_1,y_2) is infeasible in G, there exist no two disjoint paths in $G[B]$, where one is between u_1 and v_1 and the other is between u_2 and v_2. Hence there exist no two disjoint paths in $G[B]$ from $\{u_1,v_2\}$ to $\{u_2,v_1\}$. So by Menger's theorem, there exists a separation (A'',B'') of $G[B]$ of order at most one such that $\{u_1,v_2\}\subseteq A''-B''$ and $\{u_2,v_1\}\subseteq B''-A''$. Then $(A\cup A'',B'')$ and $(A\cup B'',A'')$ are separations of G of order at most $2+|A''\cap B''|\leq 3$ such that $\{x_1,x_2,y_1,y_2\}\subseteq A\subseteq(A\cup A'')\cap(A\cup B'')$. By Claim 1, $B''-(A\cup A'')=\emptyset=A''-(A\cup B'')$. Since $B-A\neq\emptyset$, $|A''\cap B''|=1$. So there exists the unique vertex c in $B-A$. By Claim 1, c has degree at least four, so c is adjacent to all vertices in $A\cap B$. $\Box$

Claim 3: For every edge e of G, if at least one end of e is not in $\{x_1,x_2,y_1,y_2\}$, then $(G/e)+x_1x_2y_1y_2x_1$ is planar.

Proof of Claim 3: Let e be an edge of G with at least one end not in $\{x_1,x_2,y_1,y_2\}$. Let $G_1=G/e$. Since at least one end of e is not in $\{x_1,x_2,y_1,y_2\}$, we can assume $\{x_1,x_2,y_1,y_2\}\subseteq V(G_1)$. Since (x_1,x_2,y_1,y_2) is infeasible in G, (x_1,x_2,y_1,y_2) is infeasible in G_1. Since $|V(G_1)|=|V(G)|-1$, by the induction hypothesis, for every full 3-reduction G_2 of (x_1,x_2,y_1,y_2) for G_1, $G_2+x_1x_2y_1y_2x_1$ can be drawn in the plane. Hence we may assume

$G_2 \neq G_1$, for otherwise we are done.

Let (A, B) be a separation of G_1 of order at most three such that $\{x_1, x_2, y_1, y_2\} \subseteq A$ and $B - A \neq \emptyset$. Since $G_1 = G/e$, by recontracting e, we know that there exists a separation (A', B') of G of order at most four such that $\{x_1, x_2, y_1, y_2\} \subseteq A'$ and $B' - A \neq \emptyset$, where the order of (A', B') equals four if and only if the vertex of $G_1 = G/e$ obtained by contracting e is in $A \cap B$ and both ends of e are in $A' \cap B'$. By Claim 1, (A', B') has order four. So both ends of e are in $A' \cap B'$. Hence at least one vertex in $\{x_1, x_2, y_1, y_2\}$ is in $A' - B'$. In particular, $A' - B' \neq \emptyset$. By Claim 2, $B' - A' = B - A$ contains exactly one vertex c, and c is adjacent in G to all four vertices in $A' \cap B'$.

This implies that $G_1 + x_1x_2y_1y_2x_1$ can be obtained from the plane drawing of $G_2 + x_1x_2y_1y_2x_1$ by repeatedly picking a face on 3 vertices and adding a vertex adjacent to those 3 vertices. So $G_1 + x_1x_2y_1y_2x_1$ is planar. $\square$

Claim 4: $G + x_1x_2y_1y_2x_1$ is planar.

Proof of Claim 4: Suppose to the contrary that $G + x_1x_2y_1y_2x_1$ is not planar. By Kuratowski's theorem, there exists a subgraph H of $G + x_1x_2y_1y_2x_1$ isomorphic to a subdivision of K_5 or $K_{3,3}$. If there exists a vertex v of H with degree 2 in H such that $v \notin \{x_1, x_2, y_1, y_2\}$, then v is incident with an edge e with at least one end not in $\{x_1, x_2, y_1, y_2\}$, so Claim 3 implies that $(G/e) + x_1x_2y_1y_2x_1$ is planar, but H/e is a subdivision of K_5 or $K_{3,3}$ contained in $(G/e) + x_1x_2y_1y_2x_1$, a contradiction. So every vertex of H with degree 2 in H is in $\{x_1, x_2, y_1, y_2\}$. Moreover, both edges of H incident with a degree-2 vertex in H have all ends in $\{x_1, x_2, y_1, y_2\}$. So H has at most two degree-2 vertices.

Since $|V(G)| \geq 5$, if there exists $v \in \{x_1, x_2, y_1, y_2\}$ such that v has no neighbor in $V(G) - \{x_1, x_2, y_1, y_2\}$, then there exists a separation (A, B) of order at most three such that $\{x_1, x_2, y_1, y_2\} \subseteq A$ and $B - A \neq \emptyset$, contradicting Claim 1. So every vertex v in $\{x_1, x_2, y_1, y_2\}$ is incident with an edge e_v not incident with $\{x_1, x_2, y_1, y_2\} - \{v\}$. If some vertex $v \in \{x_1, x_2, y_1, y_2\}$ is not contained in H, then H is a subgraph of $(G/e_v) + x_1x_2y_1y_2x_1$ isomorphic to a subdivision of K_5 or $K_{3,3}$, contradicting Claim 3. So every vertex in $\{x_1, x_2, y_1, y_2\}$ is contained in H. Similarly, e_v has both ends in $V(H)$ for every $v \in \{x_1, x_2, y_1, y_2\}$. Then it is easy to show that (x_1, x_2, y_1, y_2) is feasible in G by considering e_v for $v \in \{x_1, x_2, y_1, y_2\}$, a contradiction. $\square$

Take a plane embedding of $G + x_1x_2y_1y_2x_1$. Suppose to the contrary that the 4-cycle $C = x_1x_2y_1y_2x_1$ does not bound a face. Let D be the disk bounded by C. Let A be the set consisting of vertices in C and the vertices drawn outside the disk D. Let B be the set consisting of vertices in C and the vertices drawn inside the disk D. Then (A, B) is a separation of G with $A \cap B = \{x_1, x_2, y_1, y_2\}$. Since C does not bound a face, $A - B \neq \emptyset \neq B - A$. By Claim 2, $|A - B| = |B - A| = 1$, the only vertex in $A - B$ adjacent to all vertices in $\{x_1, x_2, y_1, y_2\}$, and the only vertex in $B - A$ adjacent to all vertices in $\{x_1, x_2, y_1, y_2\}$. So (x_1, x_2, y_1, y_2) is feasible in G, a contradiction. ■

1 Fixed-Parameter Tractability

Recall that determine whether a graph G has a vertex-cover with size at most k is NP-hard when k is part of the input. On the other hand, there are at most $O(|V(G)|^k)$ subsets S of $V(G)$ with size at most k, and testing whether each such S is a vertex-cover can be done in time $O(|V(G)|)$. So there exists a $O(|V(G)|^{k+1})$ time algorithm to determine whether G has a vertex-cover with size at most k. That is, if k is a fixed integer instead of part of the input, we can determine whether G has a vertex-cover with size at most k in polynomial time. On the other hand, we mentioned (without a proof) that the Disjoint Path Problem is NP-hard, but the k-Disjoint Path Problem can be solved in $f(k)|V(G)|^3$ time.

Hence by fixing the "parameter" k, we can make an NP-hard problem become in P. But there are two kinds of polynomial time algorithms as mentioned above. One runs in time $n^{g(k)}$ and the other runs in time $f(k)n^c$ for some constant c. We usually prefer the second one, because it usually gives better complexity. For example, when $k = \log \log n$ (which grows to infinity slowly with respect to n), if f is an exponential (or even double exponential) function, then $f(k)n^c = O(n^{c+1})$; while for any increasing non-constant function g, $n^{g(k)}$ is not polynomial. The problems that have the second kind of algorithm is said to be fixed-parameter tractable.

We give a more precise definition. The *parameterization* of a decision problem is a function p that maps every every instance of the problem to an integer. (For example, both the inputs of the Vertex-Cover Problem and Disjoint Path Problem are (G, k), and the function p that maps (G, k) to k is a parameterization.) A decision problem with a parameterization p is *fixed-parameter tractable* (a.k.a *FPT*) if there exist a function f and a constant c such that for every input x, it can be solved in time $f(p(x))n^c$, where n is the size of x.

Hence the k-Disjoint Path Problem is FPT, while our above naive algorithm for k-Vertex-Cover Problem does not build the fixed-parameter tractability. But we will show that k-Vertex-Cover Problem is indeed FPT.

1.1 Kernalization

One trick to obtain an FPT algorithm is to find a "kernel".

Let P be a decision problem and let p be a parameterization. We usually write the instance of this parameterized problem as (I, k), where I is an stance of P and $k = p(I)$. A *kernalization of P with size s*, where s is a function, is a polynomial time algorithm that transforms each instance (I, k) to another instance (I', k') such that

- (I, k) is a positive instance if and only if (I', k') is a positive instance,
- $k' \leq k$, and
- $|I'| \leq s(k)$.

We call (I', k') the *kernel* of (I, k) (for this kernalization).

Note that having a kernalization implies fixed-parameter tractability, since we can do brute-force on the kernel (whose running time only depends on the size of the kernel and hence only depends on k but not n) to decide whether the original input is a positive instance or not.

================================

An algorithm for finding a kernel for Vertex-cover with size k^2
Input: A simple graph G and an integer k.
Output: A simple graph G' and an integer k' with $|V(G')| \leq k^2$ and $k' \leq k$ such that G has a vertex-cover with size at most k if and only if G' has a vertex-cover with size at most k'.
Procedure:

Step 0: Set $G' = G$ and $k' = k$.

Step 1: If G' has an isolated vertex, then redefine $G' = G' - v$.

Step 2: If there exists a vertex v in G' with degree at least $k + 1$, redefine $G' = G' - v$ and $k' = k' - 1$.

Step 3: If there exists a vertex v in G' with degree equal to 1, then let u be the unique neighbor of v, redefine $G' = G' - u$ and $k' = k - 1$.

Step 4: Repeat Steps 1-3 until no modification can be made. If $|E(G')| \leq k^2$, then output G' and k'; otherwise, redefine $G' = K_2$ and $k' = 0$, and output G' and k'.

==================================

Proposition 2 *The above algorithm runs correctly in time $O(|V(G)|^3)$. And k-Vertex Cover can be solved in time $O(|V(G)|^3 + k^{2k}) = O(k^{2k}|V(G)|^3)$.*

Proof. We first prove that (G', k') is a positive instance if and only if it remains positive when a round of Steps 1, 2 or 3 is done. It is clearly true for Step 1.

Assume that v is a vertex of degree at least $k+1$. If v is not used in a vertex-cover, then in order to cover at least $k+1$ edges that are incident with v, we need to put at least $k+1$ vertices in a vertex-cover. So if G' has a vertex-cover of size at most k, then v is in it; if G' does not have a vertex-cover of size at most k, then $G' - v$ does not have a vertex-cover of size $k-1$. Hence Step 2 preserves the positivity and negativity.

Assume v' is a vertex of degree 1. Let u' be the unique neighbor of v'. If a vertex-cover S contains v', then $(S - \{v'\}) \cup \{u'\}$ is a vertex-cover with the same size, and $S - \{v'\}$ is a vertex-cover of $G' - u'$. So if G' has a vertex-cover of size at most k, then $G' - u'$ has a vertex-cover of size at most $k-1$; if G' does not have a vertex-cover of size at most k, then $G' - u'$ does not have a vertex-cover of size $k-1$. Hence Step 3 preserves the positivity and negativity.

So (G', k') is a kernel when no further Steps 1-3 can be applied. Note that G' has minimum degree at least two at this point. So $|E(G')| \geq |V(G')|$. Hence if $|E(G')| \leq k^2$, then $|V(G')| \leq |E(G')| \leq k^2$ and (G', k') is output, so we are done. If $|E(G')| > k^2$, then since Step 2 is not applicable, G' has minimum degree at most k, so k vertices can cover at most $k^2 < |E(G')|$ edges, and hence (G, k) is a negative instance, and the algorithm outputs a negative instance with size $2 \leq k^2$.

Hence the algorithm works correctly. And it clearly runs in time $O(|V(G)|^3)$.

Note that for the final graph G' and integer k', we can test whether G' has a vertex-cover in time $O(|V(G')|^{k'}) = O(k^{2k})$. So the above process decides whether G has a vertex-cover of size at most k or not in time $O(|V(G)|^3 + k^{2k})$.

■

We remark that we actually obtain an algorithm whose running time is of the form $f(k) + n^c$ in Proposition 2, which looks better than the required running time $f(k)n^c$ for FPT. But $f(k)+n^c$ and $f(k)n^c$ are in fact equivalent (with different functions f and constants c) since $f(k) + n^c \leq f(k)n^c$ and $f(k)n^c \leq (f(k))^2 + (n^c)^2 = g(k) + n^{c'}$, where $g(k) = (f(k))^2$ and $c' = 2c$. (We use the easy fact that for any positive numbers a and b, $ab \leq \max\{a, b\} \cdot \max\{a, b\} = \max\{a^2, b^2\} \leq a^2 + b^2$.)

We also remark that the size of the kernel mentioned in Proposition 2 can be reduced to be linear in k by using other tricks. And the running time for k-Vertex-Cover Problem in Proposition 2 can be further reduced, as we will see in next section.

As we mentioned above, having kernelization implies fixed-parameter tractability. But in fact they are equivalent.

Proposition 3 *A parameterized problem has a kernelization if and only if it is fixed-parameter tractable.*

Proof. It suffices to show that if a parameterized problem is FPT, then it has a kernelization. Assume the running time of this problem is $f(k)n^c$ for some function f and constant c. Now we describe a kernalization with size f.

If the input size $n \leq f(k)$, then we just output the input as the kernel. If the input size $n > f(k)$, then we just run the FPT algorithm (which takes time $f(k)n^c \leq n^{c+1}$) to know whether the input is positive or negative, and output a trivially positive instance or trivially negative instance as a kernel. Note that this process for producing the kernel takes time $O(n) + O(n^{c+1}) = O(n^{c+1})$, which is polynomial. ■