Notes

Date：2017/10/16

Lecture 6：Match Data with Algorithms

Expert Student： Hongyang Liu

1. data of fixed parameters

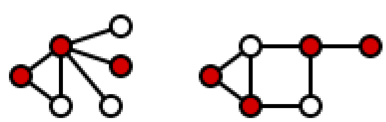
1.1 vertex cover set of fixed constant size k

2. 1.1 Description

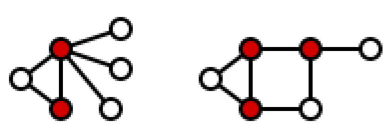
A vertex cover of a graph is a set of vertices such that each edge of the graph is incident to at least one vertex of the set.

2. 1.2 Definition

Formally, a vertex cover $V'$of an undirected graph G = (V, E)} is a subset of V such that $uv \in V' \rightarrow u \in V' \cup v \in V'$, that is to say, it is a set of vertices $V'$where every edge has at least one endpoint in the vertex cover V'. Such a set is said to cover the edges of G. The following figure shows two examples of vertex covers, with some vertex cover V' marked in red.



A minimum vertex cover is a vertex cover of smallest possible size. The vertex cover number τ is the size of a minimum vertex cover, i.e. $\tau = |V'|$ .The following figure shows examples of minimum vertex covers in the previous graphs.



$\binom{n}{k} \approx n^k$, where polynomial of k is constant.

If go over all loops and check, it will be $n^k * \mathrm{m}$

We use FPA to solve the problem, where FPA means fixed parameter algorithm.

Fix k to be the output.

1.2 FPT : reduced time to $O(f(k)) * n^{O(1)}$

Go over all the edges.

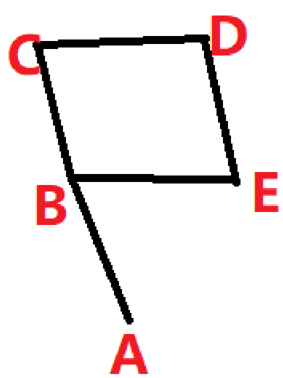
Depth $\leq$ k.

1.3 Kernelization in fixed parameter complexity

Kernelization means we throw away irrelevant data until we can no longer throw. At that time, kernel is reached.

We know the vertex cover of edge is equal to k. In our algorithm, we can eliminate vertex of degree k+1 in our consideration.

Example



Let k = 2

The degree of vertex B is k+1, so we throw it away because it must be in the vertex cover of size 2. If not, all its neighbors must be in the V.C. There are k+1 of them.

Every edge in (f, 1), (f, 2), (f, k+1) must be covered by one node in V.C.

$$f \notin \mathrm{V.C.} \rightarrow \begin{cases} 1 \in V.C. \\ 2 \in V.C. \\ k+1 \ \in V.C. \end{cases}$$

Proof by contradiction.

In our algorithm, we can eliminate vertex of degree k+1. Solve the problem for [G - {f}] with parameter k-1. And for [G - {f}], we remove f and all its incident edge.

Look at all vertex subset of size k in the remaining graph, check if it’s a vertex cover.

2. Data of fixed distribution

2.1 Order statistics

2. 1.1 Notation and examples

The first order statistic (or smallest order statistic) is always the minimum of the sample, that is,

$$\mathrm{X}_{(1)} = \min(X_1, \dots\dots, \ \mathrm{X_n})$$

where, following a common convention, we use upper-case letters to refer to random variables, and lower-case letters (as above) to refer to their actual observed values.

Similarly, for a sample of size n, the nth order statistic (or largest order statistic) is the maximum, that is,

$$\mathrm{X}_{(n)} = \max\left(X_1, \dots\dots, \ \mathrm{X_n}\right).$$

The sample range is the difference between the maximum and minimum. It is clearly a function of the order statistics:

$$\mathrm{Range}\left(\mathrm{X}_1, \dots\dots, \ \mathrm{X_n}\right) = \mathrm{X_{(n)}} - \mathrm{X_{(1)}}$$

2.2 The case for exponential distributions

$$P_r(x > t) = e^{-t}, \mathrm{t} \geq 0$$

$$\mathrm{P(x_1 > t) > P(x_2 > t) > \cdots > P(x_n > t) = e^{-nt}, t \geq 0}$$

$$X_{in} > t \rightarrow \ \forall i, X_{in} > t \rightarrow \ x_1 > t, x_2 > t, \dots\dots, x_n > t$$

2.3 Joint distributions of order statistics

For i < j, the joint probability density function of the two order statistics $U_{(i)} < U_{(j)}$ can be shown to be

$$f_{U_{(i)}U_{(j)}}((u,v))\mathrm{dudv} = \mathrm{n!}\frac{u^{i-1}}{(i-1)!}\frac{(u-v)^{j-i-1}}{(j-i-1)}\frac{(1-u)^{n-j}}{(n-j)!}\mathrm{dudv}$$

which is O(dudv) the probability that i − 1, 1, j − 1 − i, 1 and n − j sample elements fall in the intervals (0,u), (u, u+du), (u+du,v), (v,v+dv), (v+dv,1) respectively.

2.4 Generating order statistics of exponential distributions

For $\mathrm{X_1, X_2, \dots\dots, \ X_n}$ random samples from an exponential distribution with parameter λ the order statistics X(i) for i = 1,2,3, ..., n each have distribution

$$X_{(i)} \overset{d}{=} \frac{1}{\lambda}\left(\sum_{j=1}^{i} \frac{Z_j}{n-j+1}\right)$$

where the Zj are iid standard exponential random variables (i.e. with rate parameter 1).