# London Taught Course Centre

**2018 examination**

# Graph Theory

## Answers

**1** (a) This is standard, and not very hard as long as you use the equivalent definition of NP as the class of languages for which there exists a polynomial time verifier (it is still not very hard with the original definition, but the proof is a bit more cumbersome to phrase).

Let $\mathcal{L}$ be a language in NP, and let $M$ be a Turing Machine which verifies $\mathcal{L}$: that is, there is a polynomial $p$ with the following property. For each word $x$ of $L$ there is a word $y$ of length at most $p(|x|)$ such that $(x, y)$ is accepted by $M$ in time at most $p(|x| + |y|)$, and for each word $x \notin L$ there is no word $y$ such that $M$ accepts $(x, y)$.

We let $M'$ be a Turing Machine which does the following. On input $x$, it computes $p(x)$ and iterates through all words $y$ of length at most $p(x)$. For each such word, it computes $p(|x| + |y|)$ and simulates the running of $M$ on input $(x, y)$ for $p(|x| + |y|)$ time steps or until halting, whichever comes first. If the simulated $M$ at some stage accepts, then $M'$ accepts as well, otherwise $M'$ rejects.

It is clear from the definition that $M'$ decides $\mathcal{L}$. It is only necessary to check that there is $k$ such that the running time is at most exponential in $|x|^k$. The number of strings $y$ over which $M'$ iterates is at most $3^{p(|x|)}$, and to check each string entails a polynomial in $|x|$ amount of time since $M'$ can simulate $M$ efficiently, proving the result. (For this purpose efficiency is not really necessary; the most naive simulation will be fast enough)

In terms of the original definition, one should iterate over all the possibilities for nondeterministic branching; this is also easily enough checked to be a set of at most exponential in a power of $|x|$ size since the nondeterministic Turing Machine runs for at most polynomial in $|x|$ time and since the maximum degree of branching is independent of $|x|$ (it depends on the number of states).

(b) No, you cannot; or at least, if you can it should be published not submitted as an exam answer (and you should find this out using Google very easily). It's not known (and many complexity theorists do not believe it is true) that EXP = NEXP. The $st -$ CONNECTIVITY algorithm which is at the heart of Savitch's theorem is capable of deciding connectivity of $s$ and $t$ using only polylogarithmic space, but its running time is still at least linear in $n$ (as any connectivity algorithm for general graphs rather trivially must be: you cannot decide connectivity without reading the entire input, otherwise you may fail to see a connecting path). You can formalise deciding a language $\mathcal{L}$ in NEXP in terms of asking for a path from the initial state to the accepting state in the tape/state graph of a nondeterministic Turing Machine which witnesses $\mathcal{L} \in$ NEXP. This graph has a doubly exponential number of vertices, so that a polylog-space connectivity algorithm can decide connectivity using exponential space. But it still takes doubly exponential time, unless you find a way to make use of the (rather special!) structure of the tape/state graph. No-one managed to do such a thing yet, and it's not clear that it should even be possible.

So you can modify the proof of Savitch's theorem to show that EXPSPACE = NEXPSPACE (with the obvious definitions) but, unless you just published a breakthrough paper in computational complexity, you cannot do it for the time versions asked in the question.

**2** (a) This is a modification of a theorem given in the lectures, and the following solution is similarly a (very small) modification of the proof.

To begin with, observe that any 2-degenerate graph is 3-colourable. It follows that if $H$ is not 3-colourable, then removing successively vertices of degree at most 2 (in the current graph) until no more exist we find $H' \subseteq H$ with minimum degree at least three. We can construct a four-vertex path in $H'$ greedily.

We can assume $\alpha \le 1$, trivially. We choose $\varepsilon = \alpha/100$ (which is easily small enough for the proof; there is no good reason to try to optimise constants), and we set $K = \lfloor \log_{1+\varepsilon} 2/\alpha \rfloor + 1$. We choose

$$C(\alpha) = 100^K \alpha^{-K-10} \,.$$

If $n \le C(\alpha)$, we simply colour each vertex of $G$ with a different colour. So suppose $n > C(\alpha)$. We begin by taking a maximum cut $(X, Y)$ of $G$; that is, $Y = V(G) \setminus X$ where $X$ is chosen to maximise the number of edges between $X$ and $Y$.

Without loss of generality, we can assume $\chi(G[X]) \ge \chi(G[Y])$. So it is enough to show $\chi(G[X]) \le C/2$. Observe that for each $x \in X$ we have $d(x;Y) \ge \alpha n/2$ (otherwise we could move $x$ to $Y$ and increase the number of edges crossing). Here $d(x;Y)$ means the number of edges from $x$ to $Y$.

Given $X' \subseteq X$ and $Y' \subseteq Y$, and a partition $Y' = Y_1 \cup Y_2 \cup Y_3$, for $i = 1, 2, 3$ let

$$X_i := \Big\{ x \in X' : x \notin X_j \text{ for } j < i \text{ and } \frac{d(x;Y_i)}{|Y_i|} \ge (1+\varepsilon)\frac{d(x;Y')}{|Y'|} \Big\} \,.$$

Let $X_4 := X' \setminus (X_1 \cup X_2 \cup X_3)$. If $|Y_i| \ge \frac{\alpha}{10}|Y'|$ for each $i = 1, 2, 3$, and $G[X_4]$ is 4-colourable then we have an *$\varepsilon$-booster for* $(X', Y')$.

We generate partitions $\mathcal{X}$ and $\mathcal{Y}$ of $X$ and $Y$ respectively, together with a relation 'in correspondence', as follows. We begin with $\{X\}, \{Y\}$ and we say that $X$ and $Y$ are in correspondence. We now iteratively do the following. Pick $X' \in \mathcal{X}$ and $Y' \in \mathcal{Y}$ which are in correspondence. If there is a $\varepsilon$-booster $Y_1, Y_2, Y_3$ for $(X', Y')$, we replace $Y'$ with $Y_1, Y_2, Y_3$ and we replace $X'$ with $X_1, X_2, X_3, X_4$ (if some of these sets are empty we simply do not add them); we say $X_i$ and $Y_i$ are in correspondence for each $i = 1, 2, 3$. So $X_4$ is not in correspondence with any set of the new $\mathcal{Y}$. We repeat this procedure until any remaining $(X', Y')$ which are in correspondence do not have $\varepsilon$-boosters.

There is a natural way to draw a tree representing this process: we start with the root labelled $X$, then add children labelled with the sets into which $X$ is split by finding an $\varepsilon$-booster for $(X, Y)$, and then to each of those, children for their $\varepsilon$-boosters, and so on.

We claim that when this process terminates we have $\mathcal{X}$ with at most $4^K$ parts. Indeed, suppose that we have in the above tree a path from the root with $K+1$ vertices. Then we have $Y = Y'_0, Y'_1, \ldots, Y'_K$, where each $Y'_i$ is obtained by finding an $\varepsilon$-booster in $Y'_{i-1}$. Let $x$ be a vertex in the corresponding set of $\mathcal{X}$ (the end of the path). Then we have

$$\frac{d(x;Y'_K)}{|Y'_K|} \ge (1+\varepsilon)\frac{d(x;Y'_{K-1})}{|Y'_{K-1}|} \ge \cdots \ge (1+\varepsilon)^K \frac{d(x;Y'_0)}{|Y'_0|} \ge (1+\varepsilon)^K \frac{2}{\alpha} > 1 \,.$$

But $x$ cannot have more than $|Y'_K|$ neighbours in $Y'_K$; this is a contradiction. So our tree, in which each node has at most four children, also has depth at most $K$; it has at most $4^K$ leaves, and the leaves are precisely the elements of $\mathcal{X}$.

This also justifies that any set $Y' \in \mathcal{Y}$ has size at least $\left(\frac{\alpha}{10}\right)^{K+1} n$; we start with $|Y| \geq \frac{\alpha}{2}n$, and by definition each time we find a booster our sets decrease in size by at most a $\frac{\alpha}{10}$ factor.

At this point we do something slightly different to the proof from the notes.

Consider a set $X'$ in the final $\mathcal{X}$. If it does not have a corresponding $Y'$, it is 4-colourable by construction. If it does have a corresponding $Y'$, then $(X', Y')$ does not have any $\varepsilon$-booster.

Suppose that $G[X']$ is not 4-colourable. Then (by the observation above) we can find a four-vertex path $af, fg, gb$ in $G[X']$. Each of $a, b$ has at least $\alpha|Y'|/2$ neighbours in $Y'$ by construction. Let $Z_1$ be a set of $\alpha|Y'|/4$ vertices in $Y'$ which are neighbours of $a$; and let $Z_2$ be a disjoint set of $\alpha|Y'|/4$ neighbours of $b$ in $Y'$. Let $Z_3 := Y' \setminus (Z_1 \cup Z_2)$. Note that each of $Z_1, Z_2, Z_3$ contains more than $\frac{\alpha}{10}|Y'|$ vertices. Since there is no booster for $(X', Y')$, in particular this partition of $Y'$ does not give an $\varepsilon$-booster. So in the corresponding partition $X'_1, X'_2, X'_3, X'_4$ of $X'$ the graph $G[X'_4]$ is not 4-colourable.

Since $G[X'_4]$ is not 4-colourable, it contains at least five vertices; we can take $c$ to be some vertex of $X'_4$ other than $a, b, f, g$. Now, if $c$ has neighbours $d$ in $Z_1$ and $e \in Z_2$, we get a copy of $C_7$ on vertices $a, d, c, e, b, g, f$. But suppose $c$ has no neighbours in $Z_1$. Then we have $d(c; Z_2 \cup Z_3) = d(c; Y')$, and so

$$\frac{d(c; Z_2 \cup Z_3)|Y'|}{|Z_2 \cup Z_3| d(c; Y')} = \frac{|Z_2 \cup Z_3|}{|Y'|} \geq \tfrac{1}{1-\alpha/10} > 1 + \varepsilon\,,$$

and by averaging for at least one of $i = 2, 3$ we have

$$\frac{d(c; Z_i)|Y'|}{|Z_i| d(c; Y')} > 1 + \varepsilon\,,$$

which is a contradiction to the assumption $c \in X'_4$. The same calculation gives a contradiction if $c$ has no neighbours in $Z_2$; again we reached a contradiction.

In conclusion, if in the final $\mathcal{X}$ we have a set $X'$ in correspondence with some $Y'$ then $X'$ is 4-colourable. So we can colour $G[X]$ using at most $4 \cdot 4^K$ colours. By the same argument we can colour $Y$ with at most $4 \cdot 4^K$ colours. Putting these together, we can colour $G$ with at most $4^{K+2} < C(\alpha)$ colours, as desired.

It would be enough for a student to point out how to modify the proof. We use the same decomposition approach, but allow the non-corresponding set to be 4-colourable rather than independent. We claim that a final set $X'$ in correspondence with $Y'$ is also 4-colourable. If not, we can find in it a path on four vertices (following the hint) $a, f, g, b$. As in the notes, we generate a candidate $\varepsilon$-booster using the neighbourhoods of $a$ and $b$ in $Y'$; because we do not obtain an $\varepsilon$-booster we conclude that the corresponding $X'_4$ is not 4-colourable. We have the same case distinction as in the notes; because $X'_4$ has at least five vertices we can choose $c$ not equal to $a, b, f, g$ in $X'_4$ and in either case we obtain $C_7$.

It is also enough for a student to provide a clear reference to the theorem in the literature, together with an explanation of how the literature result applies. The reference is to the (so far unpublished, but on arXiv) paper 'Coloring dense graphs via VC-dimension' by Tomasz Łuczak and Stéphan Thomassé (and the same is true for the next question). The result is generalised in Allen, Böttcher, Griffiths, Kohayakawa, Morris 'The chromatic thresholds of graphs' and this could also be cited (but really it should not be; the Łuczak-Thomassé paper has priority on these two results).

(b) This is rather harder. Again it is proved using a modification of the proof in notes, but the modification is more subtle.

We can assume $\alpha \le 1$, trivially. We choose $\varepsilon = \alpha^4/10000$ (which is easily small enough for the proof; there is no good reason to try to optimise constants), and we set $K = \lfloor \log_{1+\varepsilon} 2/\alpha \rfloor + 1$. We choose

$$C(\alpha) = 100^K \alpha^{-K-10} .$$

If $n \le C(\alpha)$, we simply colour each vertex of $G$ with a different colour. So suppose $n > C(\alpha)$. We begin by taking a maximum cut $(X, Y)$ of $G$; that is, $Y = V(G) \setminus X$ where $X$ is chosen to maximise the number of edges between $X$ and $Y$.

Without loss of generality, we can assume $\chi(G[X]) \ge \chi(G[Y])$. So it is enough to show $\chi(G[X]) \le C/2$. Observe that for each $x \in X$ we have $d(x; Y) \ge \alpha n/2$ (otherwise we could move $x$ to $Y$ and increase the number of edges crossing). Here $d(x; Y)$ means the number of edges from $x$ to $Y$.

Given $X' \subseteq X$ and $Y' \subseteq Y$, and a partition $Y' = Y_1 \cup Y_2 \cup Y_3 \cup Y_4 \cup Y_5$, for $i = 1, 2, 3, 4, 5$ let

$$X_i := \left\{ x \in X' : x \notin X_j \text{ for } j < i \text{ and } \frac{d(x; Y_i)}{|Y_i|} \ge (1+\varepsilon) \frac{d(x; Y')}{|Y'|} \right\} .$$

Let $X_6 := X' \setminus (X_1 \cup X_2 \cup X_3 \cup X_4 \cup X_5)$. If $|Y_i| \ge \frac{\alpha^2}{100}|Y'|$ for each $i = 1, 2, 3$, and $G[X_6]$ is 6-colourable then we have an *$\varepsilon$-booster for* $(X', Y')$.

We generate partitions $\mathcal{X}$ and $\mathcal{Y}$ of $X$ and $Y$ respectively, together with a relation 'in correspondence', as follows. We begin with $\{X\}, \{Y\}$ and we say that $X$ and $Y$ are in correspondence. We now iteratively do the following. Pick $X' \in \mathcal{X}$ and $Y' \in \mathcal{Y}$ which are in correspondence. If there is a $\varepsilon$-booster $Y_1, Y_2, Y_3, Y_4, Y_5$ for $(X', Y')$, we replace $Y'$ with $Y_1, Y_2, Y_3, Y_4, Y_5$ and we replace $X'$ with $X_1, X_2, X_3, X_4, X_5, X_6$ (if some of these sets are empty we simply do not add them); we say $X_i$ and $Y_i$ are in correspondence for each $i = 1, 2, 3, 4, 5$. So $X_6$ is not in correspondence with any set of the new $\mathcal{Y}$. We repeat this procedure until any remaining $(X', Y')$ which are in correspondence do not have $\varepsilon$-boosters.

There is a natural way to draw a tree representing this process: we start with the root labelled $X$, then add children labelled with the sets into which $X$ is split by finding an $\varepsilon$-booster for $(X, Y)$, and then to each of those, children for their $\varepsilon$-boosters, and so on.

We claim that when this process terminates we have $\mathcal{X}$ with at most $6^K$ parts. Indeed, suppose that we have in the above tree a path from the root with $K+1$ vertices. Then we have $Y = Y'_0, Y'_1, \dots, Y'_K$, where each $Y'_i$ is obtained by finding an $\varepsilon$-booster in $Y'_{i-1}$.

Let $x$ be a vertex in the corresponding set of $\mathcal{X}$ (the end of the path). Then we have

$$\frac{d(x;Y'_K)}{|Y'_K|} \geq (1+\varepsilon)\frac{d(x;Y'_{K-1})}{|Y'_{K-1}|} \geq \cdots \geq (1+\varepsilon)^K \frac{d(x;Y'_0)}{|Y'_0|} \geq (1+\varepsilon)^K \frac{2}{\alpha} > 1\,.$$

But $x$ cannot have more than $|Y'_K|$ neighbours in $Y'_K$; this is a contradiction. So our tree, in which each node has at most six children, also has depth at most $K$; it has at most $6^K$ leaves, and the leaves are precisely the elements of $\mathcal{X}$.

This also justifies that any set $Y' \in \mathcal{Y}$ has size at least $\left(\frac{\alpha^2}{100}\right)^{K+1} n$; we start with $|Y| \geq \frac{\alpha}{2}n$, and by definition each time we find a booster our sets decrease in size by at most a $\frac{\alpha^2}{100}$ factor.

Consider a set $X'$ in the final $\mathcal{X}$. If it does not have a corresponding $Y'$, then $G[X']$ is 6-colourable. If it does have a corresponding $Y'$, then $(X', Y')$ does not have any $\varepsilon$-booster; we will now show that in this case also $G[X']$ is 6-colourable.

To begin with, if $G[X']$ is not 4-colourable it contains an edge $ab$. As before we can take two disjoint subsets $Z_1, Z_2$ of the neighbourhoods in $Y'$ of $a$ and $b$ respectively, each of size $\alpha|Y'|/5$. Choose arbitrarily two further sets $Z_3, Z_4$ of $Y'$ such that $Z_1, Z_2, Z_3, Z_4$ are disjoint, and set $Z_5$ to be whatever is left of $Y'$. These sets are all large enough to form an $\varepsilon$-booster; so the corresponding $X_6$ is not 6-colourable. In particular it contains an edge $cd$ which is disjoint from $ab$.

Now, by essentially the same calculation as in the previous proof, $c$ and $d$ each have at least $\alpha|Z_i|/5$ neighbours in $Z_i$ for each $i = 1, 2$. We can thus find four pairwise disjoint vertex sets $Y_1, Y_2, Y_3, Y_4$ such that $Y_1, Y_2 \subseteq Z_1$ and $Y_3, Y_4 \subseteq Z_2$, with all vertices of $Y_1$ and $Y_3$ adjacent to $c$ and all vertices of $Y_2$ and $Y_4$ adjacent to $d$.

Finally, let $Y_5 = Y' \setminus (Y_1 \cup Y_2 \cup Y_3 \cup Y_4)$. Again, these sets are big enough to form a candidate $\varepsilon$-booster; so the corresponding $X_6$ is not 6-colourable. In particular, it contains an edge $ef$ which is disjoint from $ab$ and from $cd$; and (by the same calculation again) both $e$ and $f$ have at least two neighbours in each of $Z_1, Z_2, Z_3, Z_4$. We choose one vertex fro each set adjacent to $e$ and a different one from each set adjacent to $f$; this gives a copy of $F$ in $G$, a contradiction.

So each of the at most $6^K$ sets in the final $\mathcal{X}$ is 6-colourable; we conclude that $G$ is $12 \cdot 6^k$-colourable as desired.

Again it would be enough to sketch the modification, or to provide a clear reference to the theorem in the literature.

(c) There are a couple of routes to this. The easier is as follows.

Let $H$ be a fixed graph which has girth at least 9 and chromatic number at least $C$; such graphs exist by a result of Erdős covered in the lectures. Without loss of generality we may assume $\delta(H) \geq 1$ (because we can remove isolated vertices without affecting either of our desired properties).

Let $G$ be obtained by blowing up each vertex of $H$; that means replacing each vertex with an independent set, and each edge with a complete bipartite graph between the corresponding independent sets. We blow up each vertex to size $n/v(H)$, rounded up or down to ensure $v(G) = n$. Now $G$ contains no odd cycle on less than 9 vertices, so in particular it cannot contain $F$; and since $H$ is contained in $G$ we have $\chi(G) \geq C$. If $n$ is large enough, then $f(n) < n/v(H)$ by definition, and so $\delta(G) \geq f(n)$.

Another way (which works in this specific case) is to take a random graph $H$ on $m$ vertices with edge probability $p = m^{-0.99}$. The expected number of copies of $F$ in $H$ is at most $m^8 \cdot p^{15}$ which tends (rather fast) to zero as $m \to \infty$, so with probability tending to one $H$ does not contain any copy of $F$. The probability that $H$ has any independent set on $\frac{m}{\log m}$ vertices is at most

$$m^{m/\log m}(1-p)^{-m^2/\log^2 m} \leq 2^m e^{-pm^2/\log^2 m} = 2^m e^{-m^{1.01}/\log^2 m},$$

which also tends to zero as $m$ tends to infinity, so with probability tending to one $H$ has no independent set on $m/\log m$ vertices and hence has chromatic number at least $\log m$.

We choose $m$ such that $\log m \geq C$, and $n$ large enough so that $n > (m+1)(f(n)+1)$. We can construct $G$ as follows. Take first a copy of $H$. Now add $m$ independent sets $X_1, \ldots, X_m$ of size $f(n)$, and for each $i$ put each vertex of$X_i$ adjacent to the $i$th vertex of $H$. Finally add $n - m(f(n)+1) > f(n)$ vertices $Y$ adjacent to all vertices in each $X_i$.

By construction $G$ has $n$ vertices and minimum degree at least $f(n)$, and it contains $H$ which has chromatic number $C$. It remains only to check that $G$ does not contain $F$. To begin with, observe that we cannot have any two side vertices of $F$ in $Y$; because these two vertices are adjacent to distinct endpoints of some middle edge, both of whose end-vertices can only be in $\bigcup_i X_i$. But that set is independent.

It follows that no middle vertex can lie in $\bigcup_i X_i$, since at most one of the four side vertices to which it is adjacent can then be in the copy of $H$, and we would have three side vertices in $Y$. But then also no middle vertex lies in $Y$, since then the adjacent middle vertex is in $\bigcup_i X_i$. So all middle vertices lie in the copy of $H$. Now only vertices in the copy of $H$ have more than one neighbour in the copy of $H$: so all side vertices must also lie in $H$. But $H$ is $F$-free.