CS 6150: A refresher

Christopher Mertin
(Worked with Milinda Fernando)

Due Date: Sep 8, 2015

This assignment has 5 questions, for a total of 50 points. Unless otherwise specified, complete and reasoned arguments will be expected for all answers.

Question	Points	Score
Recurrences I	10	
Recurrences II	15	
Probability I	10	
Probability II	9	
Pigeons and Probability	6	
Total:	50	

Question 1: Recurrences I .... **[10]**

Solve each of the following recurrences. You may use any method you like, but please show your work. In each recurrence, you may assume convenient starting values for $T(0)$ or $T(1)$ unless otherwise specified.. Note that c is an undetermined constant.

Solving a recurrence means that you provide a bound of the form $T(n) = O(f(n))$ for a specific f. Tight bounds get full credit: for example, if the recurrence is $T(n) = 2T(n/2) + cn$, the answer $T(n) = O(n^2)$ is correct but not precise enough, and you will not get full credit for it.

(a) [**2**] $T(n) = 7T(n/2) + cn^2$

Solution:

This solution uses the first case of the Master's Theorem

$$\log_2(7) \approx 2.81 \tag{1}$$

Since $n^{\log_2(7)} > f(n)$

$$\Rightarrow n^{2.81} > cn^2 \tag{2}$$

$$T(n) = \Theta\left(n^{\log_2(7)}\right) \tag{3}$$

(b) [**2**] $T(n) = 3T(n/2) + cn$

Solution:

This solution uses the first case of the Master's Theorem

$$\log_2(3) = 1.58 \tag{4}$$

Since $n^{1.58} > f(n)$

$$\Rightarrow n^{1.58} = cn \tag{5}$$

$$T(n) = \Theta\left(n^{log_2(3)}\right) \tag{6}$$

(c) [**2**] $T(n) = T(n/5) + c$

Solution:

This solution uses the second case of the Master's Theorem

$$\log_5(1) = 0 \tag{7}$$

Since $n^c = n^{\log_b(a)}$

$$\Rightarrow n^0 = n^0 \tag{8}$$

$$T(n) = \Theta\left(\log(n)\right) \tag{9}$$

(d) [**2**] $T(n) = T(n/2) + T(3n/10) + cn$

Solution:

Longest path on recursion tree:

$$\frac{n}{2^i} = 1 \tag{10}$$
$$i = \log_2(n) \tag{11}$$

From the recursion tree, guess $n\log(n)$ from there being $\log(n)$ terms of n at the longest recursion path

$$T(n) \leq T\left(\frac{n}{2}\right) + T\left(\frac{3n}{10}\right) + cn \tag{12}$$
$$T(n) \leq c\left(\frac{n}{2}\right)\log\left(\frac{n}{2}\right) + c\left(\frac{3n}{10}\right)\log\left(\frac{3n}{10}\right) + cn \tag{13}$$
$$T(n) \leq \left[c_1\left(\frac{n}{2}\right)(\log(n) - \log(2)) + \left(c_2\left(\frac{3n}{10}\right)\right)(\log(3n) - \log(10))\right] + cn \tag{14}$$
$$T(n) \leq [c_1 nlog(n) + c_2] + [c_3 nlog(3n) + c_4] + c_5 n \tag{15}$$
$$T(n) \leq c_1 nlog(n) + c_2 nlog(3n) + c_3 n + c_4 \tag{16}$$
$$T(n) \leq \mathcal{O}\left(n\log(n)\right) \checkmark \tag{17}$$

(e) [**2**] $T(n) = 2T(n-1) + 10n^2$

Solution:

$$T(n-1) = 2T(n-2) + 10(n-1)^2 \tag{18}$$
$$T(n) = 2\left[2T(n-2) + 10(n-1)^2\right] + 10n^2 \tag{19}$$
$$T(n) = 2^2T(n-2) + 2\cdot 10(n-1)^2 + 10n^2 \tag{20}$$
$$T(n-2) = 2T(n-3) + 10(n-2)^2 \tag{21}$$
$$T(n) = 2^2\left[2T(n-3) + 10(n-2)^2\right] + 2\cdot 10(n-1)^2 + 10n^2 \tag{22}$$
$$T(n) = 2^3T(n-3) + 10\left[2^2(n-2)^2 + 2(n-1)^2 + n^2\right] \tag{23}$$
$$\vdots$$
$$T(n-k) = 2^{k+1}T(n-k) + 10\sum_{i=0}^{k} 2^k(n-k)^i \tag{24}$$
$$T(n) = \mathcal{O}\left(2^n\right) \tag{25}$$

Question 2: Recurrences II .... [**15**]

Solve each of the following recurrences. The same note as above applies here.

(a) [**3**] $T(n) = T(n-1)T(n-2)$ (hint: you can express your answer in terms of the n^{th} Fibonacci number $F(n)$)

Solution:

$$T(n) = T(n-1)T(n-2) \tag{26}$$
$$T(n-1) = T(n-2)T(n-3)$$
$$T(n-2) = T(n-3)T(n-4)$$
$$T(n) = [T(n-2)T(n-3)]\,[T(n-3)T(n-4)] = T(n-2)\,[T(n-3)]^2\,T(n-4) \tag{27}$$
$$T(n-2) = T(n-3)T(n-4)$$
$$T(n-3) = T(n-4)T(n-5)$$
$$T(n-4) = T(n-5)T(n-6)$$
$$\begin{aligned} T(n) &= T(n-3)T(n-4)\,[T(n-4)T(n-5)]^2\,T(n-5)T(n-6) \\ &= T(n-3)\,[T(n-4)]^3\,[T(n-5)]^3\,T(n-6) \end{aligned} \tag{28}$$
$$T(n-3) = T(n-4)T(n-5)$$
$$T(n-4) = T(n-5)T(n-6)$$
$$T(n-5) = T(n-6)T(n-7)$$
$$T(n-6) = T(n-7)T(n-8)$$
$$T(n) = T(n-4)\,[T(n-5)]^4\,[T(n-6)]^6\,[T(n-7)]^4\,T(n-8) \tag{29}$$
$$T(n-4) = T(n-5)T(n-6)$$
$$T(n-5) = T(n-6)T(n-7)$$
$$T(n-6) = T(n-7)T(n-8)$$
$$T(n-7) = T(n-8)T(n-9)$$
$$T(n-8) = T(n-9)T(n-10)$$
$$T(n) = T(n-5)\,[T(n-6)]^5\,[T(n-7)]^{10}\,[T(n-8)]^{10}\,[T(n-9)]^5\,T(n-10) \tag{30}$$

The powers of the recurrences are every line of Pascal's Triangle, thus the k^{th} iteration would be as follows:

$$T(n) = \prod_{i=0}^{k} T(n-k-i)^{\binom{k}{i}}; \quad \binom{k}{i} = \frac{k!}{i!(k-i)!}; \quad k \geq 1 \tag{31}$$

The recurrence that is going to take the most time is going to be in the center of the sequence, which will be the $\lfloor \frac{k}{2} \rfloor$ k^{th} term. Quite possibly can lower bound with using the k^{th} term with Ω and maybe upperbound somehow???

(b) **[3]** $T(n) = n^{2/3}T(\sqrt{n}) + n^{1.2}$

Solution: *Solutions are hard*

(c) **[3]** $T(n) = T(n-3) + 8^n$

Solution:

Creating a recurion tree, it will stop at $8^n - (8^n + 3)$, where this can be used to find the height of the recurion tree by:

$$8^n + 3 = 3^i \tag{32}$$
$$i = \log_3(8^n + 3) \tag{33}$$

From this, we can change the time complexity to be:

$$T(n) = \sum_{i=0}^{\log_3(8^n+3)} 8^n - 3^i \tag{34}$$
$$T(n) = 8^n - \underbrace{\sum_{i=0}^{\log_3(8^n+3)} 3^i}_{\text{Geometric Series}} \tag{35}$$
$$\sum_{i=0}^{\log_3(8^n+3)} 3^i = \frac{3^{\log_3(8^n+3)+1} - 1}{1-3} \tag{36}$$
$$= \frac{3(8^n+3) - 1}{-2} = \frac{1 - 3 \cdot 8^n + 9}{2} \tag{37}$$
$$T(n) = 8^n - \frac{1}{2} + \frac{3}{2}8^n - 9 \tag{38}$$
$$T(n) = \frac{5}{2}8^n - \frac{19}{2} \tag{39}$$
$$T(n) = \mathcal{O}(8^n) \tag{40}$$

(d) [**3**] $T(n) = \sum_{k=0}^{n-1}(k \cdot T(k-1))$, where $T(0) = 1$.

Solution: *Omitted by Dr. Suresh Venkatasubramanian*

(e) [**3**] $T(n) = n + \sum_{i=1}^{\log n} T(n/2^i)$.

Solution:

From the recursion tree, we can solve for the height of it to be:

$$\frac{n}{2^i} = 1 \tag{41}$$
$$i = \log_2(n) \tag{42}$$

As there is only one thing going on in this recursion, we can sum the time complexity recursion tree to get the closed form of the solution. From the recursion tree:

$$T(n) = n + \frac{n}{2} + \frac{n}{4} + \cdots + 1 \tag{43}$$
$$T(n) = n\sum_{i=0}^{m-1}\left(\frac{1}{2}\right)^i = n\frac{1 - \left(\frac{1}{2}\right)^n}{\frac{1}{2}} = 2\left[n - \left(\frac{1}{2}\right)^n n\right] \tag{44}$$

Therefore, the sum total is:

$$T(n) = \log(n)\left[2n - \left(\frac{1}{2}\right)^n n\right] \tag{45}$$

Which reduces to the following since $\lim\limits_{n\to\infty} \left(\frac{1}{2}\right)^n n = 0$

$$T(n) = \mathcal{O}\left(n\log(n)\right) \tag{46}$$

Question 3: Probability I .... **[10]**

(a) **[2]** `d10` is a decahedron die used in Dungeons and Dragons with the ten faces labeled with the numbers 1 through 10. What is the expected value of a roll? What is the variance?

Solution:

$$\langle X\rangle = \sum_{i=0}^{\infty} xP\left(x\right) \tag{47}$$

Probability of any roll is $\frac{1}{10}$

$$\langle X\rangle = \frac{1}{10}\left(\sum_{x=1}^{10} x\right) = \frac{55}{10} = 5.5 \tag{48}$$

$$\left\langle X^2\right\rangle = \frac{1}{10}\left(\sum_{x=1}^{10} x^2\right) = \frac{385}{10} = 38.5 \tag{49}$$

The variance is defined by:

$$\sigma^2\left(X\right) = \left\langle X^2\right\rangle - \langle X\rangle^2 = 38.5 - 30.25 = 8.25 \tag{50}$$

(b) **[2]** Show that $\binom{n}{k} \le n^k/k!$. Using Stirling's approximation for the factorial function

$$n! \simeq \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

show that $\binom{n}{k} \simeq \left(\frac{ne}{k}\right)^k$

Solution:

- Show that $\binom{n}{k} \le n^k/k!$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{51}$$

<u>For $n = k$</u>

$$\frac{\cancel{n!}}{\cancel{n!}(0)!} = 1 < \frac{n^k}{k!} \checkmark \tag{52}$$

For $n > k$

$$\frac{n \cdot (n-1) \cdots (n-k+1)\cancel{(n-k)!}}{k!\cancel{(n-k)!}} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} < \frac{n^k}{k!} \checkmark \tag{53}$$

- Using Stirling's approximation for the factorial function

$$n! \simeq \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

show that $\binom{n}{k} \simeq \left(\frac{ne}{k}\right)^k$

The problem seems ill-posed as there is no way to actually get $\binom{n}{k} \simeq \left(\frac{ne}{k}\right)^k$, but what we can do is show that they are approximately in the same order of magnitude by calculating a lower bound and upper bound.

Upper Bound

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \approx \frac{n!e^k}{(n-k)!k^k\sqrt{2\pi k}} \tag{54}$$

Since $n! \leq n^k$

$$\frac{n!e^k}{(n-k)!k^k\sqrt{2\pi k}} \leq \frac{n^k e^k}{(n-k)!k^k\sqrt{2\pi k}} \leq \left(\frac{ne}{k}\right)^k \checkmark \tag{55}$$

Lower Bound

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \geq \frac{n!}{k^k n^k} \geq \frac{n^n\sqrt{2\pi n}}{e^n k^k n^k} \geq \frac{n^{n-k}}{e^n k^k} \checkmark \tag{56}$$

Therefore, we can say that $\frac{n^{n-k}}{e^n k^k} \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$ ✓

(57)

(c) [**2**] Use a Taylor expansion to justify the approximation

$$e^{-x} \simeq 1 - x$$

for x near 0.

Solution:

The definiton of the Taylor Series Expansion about some point a is:

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \tag{58}$$

with $f^{(n)}(a)$ being the n^{th} derivative of $f(a)$. Taking the Taylor Series expansion of e^{-x} results in:

$$f(a) = e^{-a} \tag{59}$$
$$f^{(0)}(a) = \left.\frac{e^{-a}}{0!}(x-a)^0\right|_{a=0} = 1 \tag{60}$$
$$f^{(1)}(a) = -\left.\frac{e^{-a}}{1!}(x-a)^1\right|_{a=0} = -x \tag{61}$$
$$f^{(2)}(a) = \mathcal{O}\left(x^2\right) \tag{62}$$

Therefore, the first-order approximation of e^{-x} about $x = 0$ is

$$e^{-x} = 1 - x + \mathcal{O}\left(x^2\right) \simeq 1 - x \tag{63}$$

(d) [**4**] A set of random variables $X_1, X_2, \ldots, X_k$ is said to be *mutually independent* if for any subset $S = \{j_1, j_2, \ldots, j_k\} \subset \{1, 2, \ldots, k\}$ it holds that

$$\Pr(X_{j_1} = v_i, X_{j_2} = v_2, \ldots, X_{j_k} = v_k) = \prod_i \Pr(X_{j_i} = v_i)$$

The same set of random variables is said to be *pairwise independent* if this holds for subsets of size $k = 2$, but not necessarily for larger subsets. Construct a set of three variables that are *not* mutually independent but are pairwise independent.

Solution:

To rephrase what this is asking for:

$$P\left(x_1 \cap x_2 \cap x_3\right) \neq P\left(x_1\right) \cdot P\left(x_2\right) \cdot P\left(x_3\right) \tag{64}$$

but...

$$P\left(x_1 \cap x_2\right) = P\left(x_1\right) \cdot P\left(x_2\right) \tag{65}$$
$$P\left(x_1 \cap x_3\right) = P\left(x_1\right) \cdot P\left(x_3\right) \tag{66}$$
$$P\left(x_2 \cap x_3\right) = P\left(x_2\right) \cdot P\left(x_3\right) \tag{67}$$

Values

x_1	x_2	x_3
0	0	0
1	0	1
0	1	1
1	1	0

Each combination has a probability of $\frac{1}{4}$ succeeding in producing 1, meaning this fits the defintion from above.

Question 4: Probability II .... **[9]**

(a) **[4]** A strange clause in a version of Dungeons and Dragons says: roll a `d6` (a six-sided die with faces from 1 to 6). If the value rolled is 3 or less, roll a `d8` else roll a `d10`. Add the two values obtained. Let X be the value of the `d6` roll and let Y be the value of the other die roll.

- Is it true that $E[X+Y] = E[X] + E[Y]$?
- Is it true that $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$?

Compute the expectation and variance of the overall value obtained.

Solution:

- Is it true that $E[X+Y] = E[X] + E[Y]$?

$$
\begin{aligned}
\langle j \rangle &= \sum_{j=0}^{\infty} jP(j) \\
\langle X+Y \rangle &= \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} P_{XY}(x,y) \\
&= \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} xP_{XY}(x,y) + \sum_{y=0}^{\infty}\sum_{x=0}^{\infty} yP_{XY}(x,y) \\
&= \sum_{x=0}^{\infty} xP_X(x) + \sum_{y=0}^{\infty} yP_Y(y) \\
&= \langle X \rangle + \langle Y \rangle \ \checkmark \\
\langle X \rangle &= \sum_{x=1}^{6} \frac{x}{6} = 3.5 \\
\langle Y_1 \rangle &= \sum_{y_1=1}^{8} \frac{y_1}{8} = 4.5 \\
\langle Y_2 \rangle &= \sum_{y_2=1}^{10} \frac{y_2}{10} = 5.5 \\
\langle Y \rangle &= \frac{1}{2}\left(\langle Y_1 \rangle + \langle Y_2 \rangle\right) = 5.0 \\
\langle X+Y \rangle &= \sum_{x=1}^{6} \frac{x}{6} + \frac{1}{2}\left(\sum_{y_1=1}^{8} \frac{y_1}{8} + \sum_{y_2=1}^{10} \frac{y_2}{10}\right) = 3.5 + \frac{1}{2}(4.5+5.5) = 8.5
\end{aligned}
$$

Solution:

- Is it true that $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$?

$$\sigma^2(j) = \langle j^2 \rangle - \langle j \rangle^2$$

$$\langle XY \rangle = \sum_{x=1}^{n}\sum_{y=1}^{n} xyP(x)P(y) = \sum_{x=1}^{n}\left(xP(x) \underbrace{\sum_{y=1}^{n} yP(y)}_{\langle Y \rangle} \right)$$

$$= \left(\langle Y_1 \rangle \sum_{x=1}^{3} xP(x) + \langle Y_2 \rangle \sum_{x=4}^{6} xP(x) \right)$$

$$= 4.5(1.0) + 5.5\left(\frac{5}{2}\right) = \frac{73}{4}$$

$$\langle X^2 \rangle = \sum_{x=1}^{6} \frac{x^2}{6} = \frac{91}{6}$$

$$\langle Y^2 \rangle = \left\langle \left(\frac{1}{2}(Y_1+Y_2)\right)^2 \right\rangle = \frac{1}{4}\left(\langle Y_1^2 \rangle + \langle Y_2^2 \rangle + 2\langle Y_1 Y_2 \rangle\right)$$

$$= \frac{1}{4}\left(\langle Y_1^2 \rangle + \langle Y_2^2 \rangle + 2\langle Y_1 \rangle \cdot \langle Y_2 \rangle\right)$$

$$= \frac{1}{4}\left(\sum_{y_1=1}^{8}\frac{y_1^2}{8} + \sum_{y_2=1}^{10}\frac{y_2^2}{10} + 2(4.5)5.5\right) = \frac{227}{8}$$

$$\left\langle (X+Y)^2 \right\rangle = \langle X^2 + 2XY + Y^2 \rangle = \langle X^2 \rangle + 2\langle XY \rangle + \langle Y^2 \rangle$$

$$= \frac{91}{6} + 2\left(\frac{73}{4}\right) + \frac{227}{8} = \frac{1921}{24}$$

$$\sigma^2(X+Y) = \left\langle (X+Y)^2 \right\rangle - \langle X+Y \rangle^2 = \frac{1921}{24} - 8.5^2 = 7.791\bar{6}$$

$$\sigma^2(X) = \langle X^2 \rangle - \langle X \rangle^2 = \frac{91}{6} - 3.5^2 = 2.91\bar{6}$$

$$\sigma^2(Y) = \langle Y^2 \rangle - \langle Y \rangle^2 = \frac{227}{8} - 5.0^2 = \frac{27}{8}$$

$$\sigma^2(X) + \sigma^2(Y) = 2.91\bar{6} + \frac{27}{8} = 6.291\bar{6}$$

(b) **[5]** Your favorite game store is constantly sold out of the latest edition of the Dungeons and Dragons Monster Guide. In fact, the probability of you finding the book when you visit the store is p. What is the expected number of times you'll need to visit to purchase the guide and continue your DnD adventures?

Solution:

For the game to be there, you want an expectation value of 1. Thus

$$\langle \text{buy} \rangle = 1 = \sum_{i=0}^{\infty} d_i p \tag{68}$$

As p is independent of the sum, you can divide by it and thus solves for the total number of days

$$\sum_{i=0}^{\infty} d_i = \frac{1}{p} \tag{69}$$

Question 5: Pigeons and Probability . . . . **[6]**

(a) **[2]** The pigeonhole principle says that if you have m objects and $n < m$ buckets to put them in, then at least one bucket contains two objects. Prove that if you have a random variable X taking positive integer values such that $E[X] = k$, then there must be at least one integer $r > k$ such that $\Pr(X = k) > 0$.

Solution: *Insert solution here.*

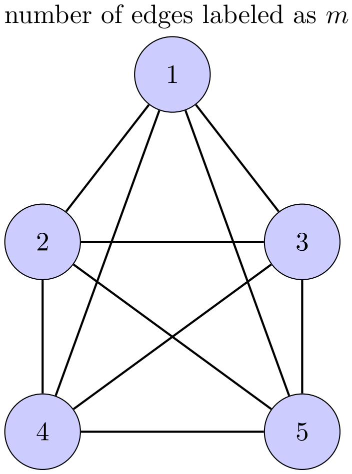
(b) **[4]** We can think of the above as a probabilistic proof of existence. We're using probabilistic techniques to show that an object (deterministically) exists. Let's apply it to a problem.

Let $G = (V, E)$ be a graph and let $H \subset V$. The *induced subgraph* $G[H]$ is the graph with vertex set H and whose edges are all edges of G whose endpoints are in H (i.e are elements of $H \times H \cap E$).

A graph $G = (V, E)$ is *bipartite* if V can be partitioned into two sets A, B (note that A, B are disjoint and $A \cup B = V$) such that all edges go between A and B (i.e $E \subset A \times B$).

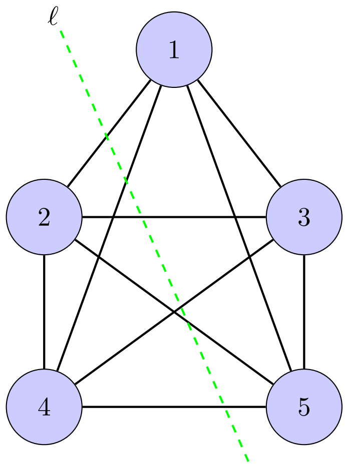
Show that any graph G with m edges contains a ~~induced~~ bipartite subgraph with $m/2$ edges (Hint: define a random process such that the desired quantity is its expectation).

Solution: Random Process: *Take a random cut through G*
Proof by Contradition: Assume that a random cut through G would *not* have the possiblity of creating a bipartite subgraph with $\frac{m}{2}$ edges. Thus, consider the following graph with the number of edges labeled as m



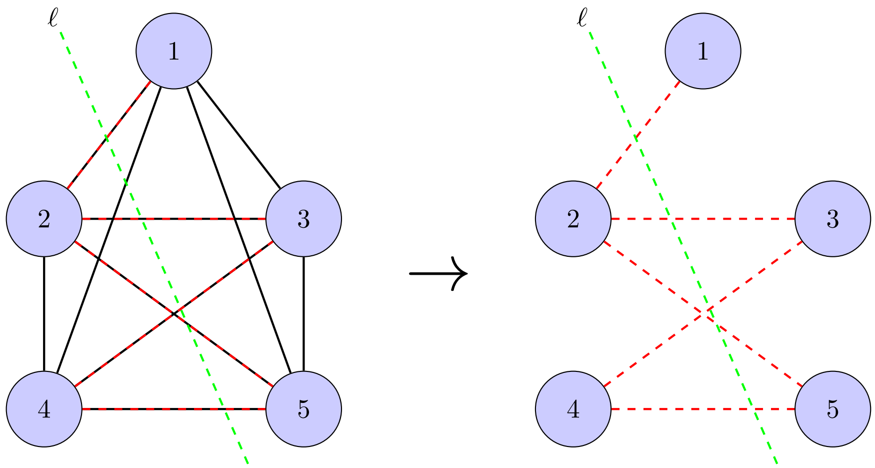
$m = 10$

If this assumption is to hold true, then there shouldn't be any random line ℓ that would cause any subgraph with $\frac{m}{2}$ edges to become bipartite. Looking back at the graph above, the requirements would be for any subgraph to be bipartite with only 5 edges. By drawing a random cut through G



$m = 10$

By separating these nodes into two sets and removing extra edges, the graph should be able to be separated into two sets of nodes which are bipartite and subgraphs of G



$m' = 5$

By removing the edges in G except for those that contain the subgraph that was decided by the random cut ℓ, it is easy to see that this graph contains a bipartite subgraph due to the random cut from ℓ. It is bipartite because the nodes $\{2, 4\}$ are not connected and the nodes $\{1, 3, 5\}$ are not connected. Therefore, by proof of contradiction, a random cut will produce a bipartite graph in any given graph.