CS 70 Discrete Mathematics and Probability Theory
Spring 2015 Vazirani HW 3

Due Monday February 9

1. **The celebrity problem**

A *celebrity* at a party is someone whom everyone knows, yet who knows no one. Suppose that you are at a party with n people. For any pair of people A and B at the party, you can ask A if they know B and receive an honest answer. Give a recursive algorithm to determine whether there is a celebrity at the party, and if so who, by asking at most $3n-4$ questions. (Note: for the purpose of this question you are just visiting the party to ask questions. What you are trying to determine is whether the n people actually attending the party include a celebrity).

Prove by induction that your algorithm always correctly identifies a celebrity iff there is one, and that the number of questions is at most $3n-4$.

Answer: Here is the recursive algorithm that we use to find the celebrity.

```
function FindCelebrity(S)
    if Size(S) = 2 then                                          ▷ Base case
        return Ask two questions and deduce if there's a celebrity
    Let A and B be arbitrary elements of S.                      ▷ Recursive case
    if Knows(A,B) then
        E ← A                                  ▷ A is not celebrity, A is eliminated
    else
        E ← B                                  ▷ B is not celebrity, B is eliminated
    X ← FindCelebrity(S − {E})
    if X = "No celebrity found" OR Knows(X,E) OR not Knows(E,X) then
        return "No celebrity found"
    else
        return X
```

Let $T(n)$ be the number of questions we ask when there are n people at the party. We ask three questions and then recursively find a celebrity amongst $n-1$ people, so $T(n) = T(n-1)+3$ questions. Since $T(2) = 2$, we have $T(n) = 3n-4$ by induction. Let us prove this formally.

Base case The base case $n = 2$ is clearly correct, since we ask 2 questions and $T(2) = 2 = 3\times 2-4$.

Induction Hypothesis Suppose that $T(n) = 3n-4$.

Induction Step We have $T(n+1) = T(n)+3 = 3n-4+3 = 3(n+1)-4$.

Next we will prove that our algorithm is correct by simple induction on n.

Base case The base case $n = 2$ is clearly correct, since we've asked all possible questions.

Induction Hypothesis Suppose that this algorithm either finds the celebrity or tells that it doesn't exist for $n = k$.

Induction Step First notice that E can't be a celebrity. Moreover, any celebrity in the original group is still a celebrity when we restrict our attention to $S - \{E\}$. So by the induction hypothesis we know that if there is a celebrity in the group, our recursion will find some celebrity in $S - \{E\}$. Moreover, there can be at most one celebrity in $S - \{E\}$ (since otherwise they would have to know each other). So the celebrity X we find amongst $S - \{E\}$ is the only possible celebrity in S. Now X is a celebrity in S iff E knows X and X does not know E. This is exactly the condition the algorithm checks.

2. **True or False**

For each of the following statements about the traditional stable marriage algorithm with men proposing, indicate whether the statement is True or False and justify your answer with a short 2-3 line explanation:

(a) (True/False) In a stable marriage algorithm execution which takes n days, there is a woman who did not receive a proposal on the $(n-1)$th day.

(b) (True/False) In a stable marriage algorithm execution, if a woman receives a proposal on day k, she receives a proposal on every subsequent day until termination.

(c) (True/False) There is a set of preferences for n men and n women, such that in a stable marriage algorithm execution every man ends up with his least preferred woman.

(d) (True/False) There is a set of preferences for n men and n women, such that in a stable marriage algorithm execution every woman ends up with her least preferred man.

(e) (True/False) In a stable marriage algorithm execution, if woman W receives no proposal on day i, then she receives no proposal on any previous day j which is less than i.

Answer:

(a) **True**: The algorithm will terminate once there is exactly one man proposing to each woman. This means that there will be at least one woman who is not proposed to every day before the algorithm terminates. This includes day $(n-1)$.

(b) **True**: This is true by the improvement lemma, since once a woman has a suitor on a string, she always has some suitor on a string and therefore must receive a proposal each day.

(c) **False**: If this were to occur it would mean that at the end of the algorithm, every man would have proposed to every woman on his list and has been rejected $n-1$ times. This would also require every woman to reject $n-1$ suitors. We know this is impossible though if we consider what we learned in parts (a) and (b). There must be at least one woman who is not proposed to until the very last day.

(d) **True**: One example occurs when every woman has a different least favorite man, who happens to prefer her over all other women. Then the stable marriage algorithm will end in one day.

(e) **True**: This is the contrapositive of part (b).

3. **Hard examples for stable marriage algorithm**

(a) Run the traditional propose and reject algorithm on the following example.

Men's preference list:

1	A	B	C	D
2	B	C	A	D
3	C	A	B	D
4	A	B	C	D

Women's preference list:

A	2	3	4	1
B	3	4	1	2
C	4	1	2	3
D	1	2	3	4

(b) In class we showed that the propose and reject algorithm must terminate after at most n^2 proposals. Prove a sharper bound showing that the algorithm must terminate after at most $n(n-1)+1$ proposals. Conclude that the above example is a worst case instance for $n=4$. How many days does the algorithm take on this instance? *Hint: Question 2b might be useful for this part*

(c) **Extra credit:** Generalize the above example to arbitrary n and prove rigorously that the algorithm makes $n(n-1)+1$ proposals on your example. How many days does the algorithm take?

Answer:

(a) Here is the overall structure of the proposals: On the first day, each man proposes to the first woman on his list. In the next four days, the men propose to the women from their second column, because each man in turn gets rejected by his first choice. In the next four days the men propose to the women from the third column. On the last (10-th) day, man 1 proposes to D, the last woman on his list, and the algorithm halts.

The table below shows the execution of the algorithm. Every entry shows which woman holds that man on a string. A bold letter indicates a proposal. A $-$ sign indicates the man is not on a string.

	Day 1		Day 2		Day 3		Day 4		Day 5		Day 6		Day 7		Day 8		Day 9		Day 10	
	M	A	M	A	M	A	M	A	M	A	M	A	M	A	M	A	M	A	M	A
1	**A**	-	**B**	B	B	B	B	B	B	-	**C**	C	C	C	C	C	C	-	**D**	D
2	**B**	B	B	-	**C**	C	C	C	C	C	C	-	**A**	A	A	A	A	A	A	A
3	**C**	C	C	C	C	-	**A**	A	A	A	A	A	A	-	**B**	B	B	B	B	B
4	**A**	A	A	A	A	A	A	-	**B**	B	B	B	B	B	B	-	**C**	C	C	C

(b) Let us prove that there is at most one man who proposes to the last woman in his list.

On the day when a man, M, proposes to the last woman in his list W, we claim that every other woman must have some man on the string. This is because M proposed to each of these women, and by the Improvement Lemma, once a woman has been proposed to she always has a man on the string. Since there are $n-1$ other women, they must be paired with all $n-1$ remaining men. Thus there is only one proposal on this day, and since it is accepted, the algorithm halts.

Therefore, at most one man proposes to his last choice, and thus there are at most $n^2-(n-1) = n(n-1)+1$ proposals.

In the example above there was $13 = 4(4-1)+1$ proposals in 10 days.

(c) **Extra credit:**

Let $M_0,...M_{n-1}$ denote the men and $W_0,...,W_{n-1}$ denote women. Let $a\ mod\ b$ denote the remainder a divided by b. Let a preference list for M_i be

$$[W_{i\ mod\ n-1}, W_{i+1\ mod\ n-1}, W_{i+2\ mod\ n-1}, \dots, W_{i+(n-2)\ mod\ n-1}, W_{n-1}].$$

Let a preference list for W_i be

$$[M_{i+1\ mod\ n}, M_{i+2\ mod\ n}, M_{i+3\ mod\ n}, \dots, M_{i+n\ mod\ n}].$$

For example when $n = 5$, men's preference lists are:

M_0	W_0	W_1	W_2	W_3	W_4
M_1	W_1	W_2	W_3	W_0	W_4
M_2	W_2	W_3	W_0	W_1	W_4
M_3	W_3	W_0	W_1	W_2	W_4
M_4	W_0	W_1	W_2	W_3	W_4

Women's preference lists:

W_0	M_1	M_2	M_3	M_4	M_0
W_1	M_2	M_3	M_4	M_0	M_1
W_2	M_3	M_4	M_0	M_1	M_2
W_3	M_4	M_0	M_1	M_2	M_3
W_4	M_0	M_1	M_2	M_3	M_4

Let us enumerate everything - men, women, days, and columns - starting from zero.

On the 0-th day, every man proposes to the first woman on his list, therefore there are n proposals.

Now let us prove by induction a lemma saying that each entry in the men's table is crossed from above to below, column by column. Let $k = in + j$ for some $0 \le j < n$ and $i \in \mathbb{N}$ in the lemma below.

Lemma: At the end of k-th day only crossed entries are in first $i-1$ columns and top j entries in the i-th column and only the last entry was crossed on k-th day.

Base Case: $k = 0 = 0 \cdot n + 0$. All men propose to different woman except M_0 and M_{n-1} who proposes W_0. M_0 is the last in her preference list but M_{n-1} the second last so she rejects M_0.

Inductive Hypothesis: At the end of k-th day, where $0 < k < n(n-2)$, the only crossed entries are in first $i-1$ columns and top j entries in the i-th column and the last entry was the only one that got crossed out at the end of that day.

Inductive step: What happens in $k+1$-th day? The only new proposal comes from the man who got rejected yesterday, and by the inductive hypothesis it is M_j. He proposes to the next uncrossed woman in his list, which is $W_{j+i+1\ mod\ n-1}$. Only he and men who are proposing to this woman can be rejected.

Men above M_j are M_l for $0 \le l < j$. By inductive assumption they are proposing to $W_{l+i+1\ mod\ n-1}$. Since $(i+j+1)\ mod\ (n-1) = (l+i+1)\ mod\ (n-1)$, if and only if $j\ mod\ (n-1) = l\ mod\ (n-1)$, which happens if and only if $j = n-1$ and $l = 0$; so it is only in the case $j = n-1$, that a man above M_j (i.e. M_0) proposes to the same woman as M_j. This woman is $W_{n-1+i+1\ mod\ n-1} = W_{i+1}$. Since $i < n-2$, this woman rejects M_0. Therefore, M_0 crosses the $i+1$-th entry in his list, which was what we had to prove in the case $j = n-1$ because $k+1 = (in+j)+1 = in+(n-1)+1 = (i+1)n+0$.

Men below M_j are M_l for $n > l > j$. They are proposing to $W_{l+i \ mod \ n-1}$, but $(j+i+1) \ mod \ (n-1) = (l+i) \ mod \ (n-1)$ if and only if $j+1 \ mod \ n-1 = l \ mod \ n-1$, which happens if and only if and $l = j+1$ and that is possible only if $j < n-1$. This woman $W_{j+i+1 \ mod \ n-1}$ rejects M_{j+1}, because for successive pairs of men, M_j, M_{j+1}, every women prefers M_j. This is what we had to prove in the case $j < n-1$ because $k+1 = (in+j)+1 = in+j+1 = in+(j+1)$. □

The $(n(n-2)+1)$-th day is the last one, because by the lemma, at the end of the previous day only M_0 gets rejected and therefore he now proposes to W_{n-1}. Now all women are proposed to, because other men propose to the first $n-1$ women.

Since there was one proposal in every day except the zeroth day, totally, there are $n + n(n-2) + 1 = n(n-1) + 1$ proposals.

4. **Procrastinating men**

Suppose we relax the rules for the men, so that each unpaired man proposes to the next woman on his list at a time of his choice (some men might procrastinate for several days, while others might propose and get rejected several times in a single day). Can the order of the proposals change the resulting pairing? Give an example of such a change or prove that the pairing that results is the same as with the standard algorithm.

Answer:

We shall prove that the pairing P' that results from the men procrastinating is the same as the pairing P found by the regular algorithm. Suppose this is not true, so there must either be a woman W who receives a proposal from a man M' who she prefers to the man M she was paired with in P, or a woman W who is not proposed to by her partner M in P. The second event is clearly impossible, since if W did not get a proposal from M, then it is because M was not rejected by some other woman W' who his higher on his preference list and who rejects him in P. But this contradicts the fact that P is male optimal. So let us assume the first event takes place. By the well-ordering principle, there is a first time when it happens.

If W rejected M, then it is because she got a proposal from the other man M' who is higher in her preference list. But M' proposed to her only because he was rejected by another woman W' who is higher on his preference list and did not rejects him in P. In other words, even before W rejected M, W' rejected her previous pair M'. Contradiction.

Therefore, the resulting pairing P' is the same as P.

5. **Better of the two**

Suppose that in the stable marriage problem with n men and n women, we have found two (possibly different) stable matchings S and T. We will show how to combine S and T into two new stable matchings W and M, which are the best of both worlds for the women and men respectively.

(a) Each woman is given the name of the man she is matched with in S and the one she is matched with in T (note that they might be the same man). Of the two names each woman receives, she picks the one she prefers the most.

Prove that no two women would end up picking the same man. Conclude that if women are matched with the men they pick, the result is a matching. Call this matching W.

(b) Prove that W is a stable matching.

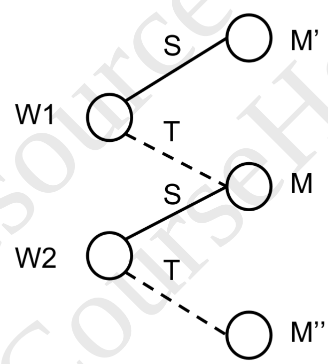
(c) Another way of combining the matchings S and T, is to give each man the names of the women he is matched with in S and T and force each man to pick the woman he prefers the least amongst the two. Prove that this results in the same stable matching W as before.

(d) How would you combine the matchings S and T to create stable matching M, which is the best of both worlds for the men? Just think about how M related to W. No need to write a proof or justification for your answer.

(e) **(Extra Credit)** Now suppose that instead of two matchings we have m matchings. Again each woman is given m names (the men she is matched with in these m matchings) and she sorts them according to her preferences. Note that the m names might have repetitions. Now let $1 \leq k \leq m$, and have each woman pick the k-th name in her list of m names. Prove that this still results in a stable matching. Also prove that if we had given each man the m names they were matched with, and had them pick the $(m-k+1)$-th name according to their preferences, then we would arrive at the same matching.

Answer:

(a) We will use proof by contradiction. Suppose two women W_1 and W_2 choose the same man M. This means that (W_1, M) comes from one stable pairing and (W_2, M) comes from the other. Without loss of generality we have the pairings $(W_1, M'), (W_2, M)$ in S and $(W_1, M), (W_2, M'')$ in T. Now suppose again without loss of generality that M prefers W_1 to W_2. We now see that in the matching S, W_1 prefers M to M' and M prefers W_1 to W_2, hence we have a rogue couple, W_1 and M, in S, contradicting that it should be a stable matching.



(b) Again we will use proof by contradiction. Assume that the matching produced is not stable, and therefore there is a rogue couple which we call M_r and W_r. Note that W_r ended up with the man she liked the best out of the two proposals and yet she does not like him as much as M_r. This means that W_r was not matched to M_r in either of S or T. We know both S and T are stable, and therefore M_r and W_r couldn't have been rogue in either.

Since W_r does like M_r more than her match in S or T (since she likes M_r more than the best of them), so the other side of the definition for rogue couples must not hold. i.e. it must be that M_r likes both of his matches more than W_r.

But in W, M_r must end up with one of his matches in the two proposals, both of whom he likes better than W_r. So he likes his final match more than W_r which contradicts the fact that M_r and W_r are a rogue couple in the first place.

(c) Suppose for the sake of contradiction that the two matchings are not the same. Then there exists a man M paired with W_1 in T and W_2 in S, who prefers W_1 less than W_2 without loss of generality, but W_1 prefers her partner M' in S more than her partner M in T (refer to the diagram above to visualize this), so that M and W_1 are paired differently in both ways of combining. Note that W_2's favorite must be M, since M must be picked by either W_1 or W_2 as their favorite partner (already shown that we have a matching if women picked their favorites), and W_1 is picking a different person.

However, we can see that in T, W_2 and M would be a rogue couple, since W_2's favorite out of the two is M and M prefers W_2 to W_1. Therefore we get a contradiction.

(d) Because of symmetry between men and women, we would ask each man to pick his best partner from S and T, and this will result in a stable pairing that is the best of both worlds for men.

(e) Consider the m matchings, and how men and women write down their m partners. Assume that we have given a number to the matchings, and when people write down their partners, they also write down from which matching they come from (so that repeated names can be distinguished). Men and women both write down their m partners from most preferred to least preferred as before. Assume that when there is a tie (which only happens when a name is repeated), for women the man who comes from the lower-numbered matching is considered preferred, but for men, the woman from the higher-numbered matching is considered preferred.

Consider a woman W and look at the man M who she is matched with in one of these matchings (let's call that matching P). Let W be the i-th person on M's list and let M be the j-th person on W's list. We will prove that $i+j=m+1$. Consider any of the m matchings other than P, like Q. In Q, M,W cannot both have partners that they prefer to each other, because according to the previous parts if we merge P and Q so that it's the best of both worlds for women, then it is the worst of both worlds for men. Similarly they cannot both have partners that are they prefer less than each other. Again if we merge P and Q so that it is the best of both worlds for women, W picks M, but W is also best of both worlds for M which is a contradiction. So there are 3 cases for how M and W are matched in Q:

- M is matched to someone he prefers over W and W is matched to someone she does not prefer over M. In this case M's match comes before W on his list, and W's match comes after M. So the pairing Q adds contributes exactly 1 to the sum $i+j$ (the contribution is to i).
- M is matched to someone he does not prefer over W and W is matched to someone she prefers over M. Again in this case the pairing Q contributes 1 to the sum $i+j$ (this time the contribution is to the j part).
- M and W are matched together in Q. In this case depending on whether Q's number is lower or higher than P, Q contributes 1 to either i or j. So no matter what, Q contributes 1 to $i+j$.

So every matching besides P contributes 1 to $i+j$. Note that P itself contributes 2 to $i+j$. Therefore $i+j$ equals $(m-1)+2=m+1$. Another way to read this is that if a man is k-th on a woman's list, then that woman is $(m-k+1)$-th on the man's list.

This immediately proves that the result of women picking their k-th partners is a matching. Because if two women picked the same man, they both have to be $(m-k+1)$-th on that man's list which is impossible.

It is also easy to see that if men picked their $(m-k+1)$-th choices, the result would be the same.

It just remains to show that the resulting matching is also stable. Suppose for the sake of contradiction that there was a rogue pair (M,W). Note that M,W are not matched. Let the number of men on W's list whom she prefers at least as much as M be j and let the number of women on M's list whom he prefers at least as much as W be i. Then $j<k$ and $i<m-k+1$, because W prefers M over her resulting partner (who is the k-th) and M prefers W over his resulting partner (who is the $(m-k+1)$-th). Therefore $i+j \le (m-k)+(k-1)=m-1$. So to derive a contradiction we just need to show that $i+j \ge m$.

Look at any of the m matchings, like P. Since M,W are not a rogue couple in P, it must be that either M,W are matched in P, or that one of them has a partner who s/he prefers over the other. In the first case we get a contribution of 1 to both i and j, and in the second case we get a contribution 1 to either i or j. So in all cases we get a contribution of at least 1 to $i+j$. But we have m matchings, so $i+j \geq m$.

6. **Highly connected website**
Suppose we have n websites such that for every pair of websites A and B, either A has a link to B or B has a link to A. Prove by induction that there exists a website that is reachable from every other website by clicking at most 2 links.

Answer: We prove this by induction on the number of websites n.

Base case For $n = 2$, there's always a link from one website to the other.

Induction Hypothesis When there are k websites, there exists a website w that is reachable from every other website by clicking at most 2 links.

Induction Step Let A be the set of websites with a link to w, and B be the set of websites two links away from w. The induction hypothesis states that the set of k websites $W = \{w\} \cup A \cup B$. Now suppose we add another website v. Between this website and every website in W, there must be a link from one to the other. If there is at least one link from v to $\{w\} \cup A$, w would still be reachable from v with at most 2 clicks. Otherwise, if all links from $\{w\} \cup A$ point to v, v will be reachable from every website in B with at most 2 clicks, because every website in B can click one link to go to a website in A, then click on one more link to go to v. In either case there exists a website in the new set of $k+1$ websites that is reachable from every other website by clicking at most 2 links.

7. **Color the graph**
Suppose that the degrees of the vertices in a graph are all at most d. Prove, using the well-ordering principle, that one can color the vertices of the graph using at most $d+1$ colors so that no two adjacent vertices end up having the same color.

Answer:

We prove this by the well-ordering principle. We call a graph *colorable* with k colors if it can be colored with at most k colors such that no two adjacent vertices end up having the same color. Suppose the statement is not true. Let G be the graph with the smallest number of vertices (all of degree at most d) that is not colorable with $d+1$ colors. We know that G exists by the well-ordering principle. Note that the number of vertices in G is greater than 1 since a graph of 1 vertex is always colorable with 1 color. Now take any vertex v in G and remove it. Then we know that the resulting graph is colorable with $d+1$ colors because G is the smallest graph that is not colorable. However, we can then add v back to the graph and assign it a color. Since v is connected to at most d edges, we can choose a color different from the colors of the vertices it is connected to. But this means that G is colorable with $d+1$ colors. We have arrived at a contradiction, and therefore proved that all graphs with degrees at most d are colorable with $d+1$ colors.

8. **Eulerian Walks and Tours**
We proved that a connected undirected graph has an Eulerian tour if and only if every vertex has even degree.

Prove that if a connected graph G on n vertices has exactly $2d$ vertices of odd degree, then there are d walks that *together* cover all the edges of G (i.e., each edge of G occurs in exactly one of the d walks; and each of the walks should not contain any particular edge more than once).

Answer:

We split the $2d$ odd-degree vertices into d pairs, and join each pair with an edge, adding d more edges in total. Notice that now all vertices in this graph are of even degree. Now by Euler's theorem the resulting graph has an Eulerian tour. Removing the d added edges breaks the tour into d walks covering all the edges in the original graph, with each edge belonging to exactly one walk.