23 Lectures 23

The star-prime algorithm or the bipartite matching algorithm

Based on the idea in the proof of Hall's theorem, we supply an algorithm to find a maximum matching in a bipartite graph. In graph theory, it is called 'bipartite matching algorithm'. This involves finding an augmenting path. We supply the algorithm in the form of finding a maximum independent set of zeros in a matrix.

An algorithm to find a maximum matching.

b1) Star a zero in the matrix if there is no starred zero in its column or row. Do it for all zeros. (We will read the entries of the matrix from the top-left, finish the first row, then the second row and so on, to mark these zeros. It makes it easier to discuss.)

b2) Cover each COLUMN containing a zero-star (by a line segment or dashed line segment).

b3) If all zeros of the matrix are covered then conclude that these zero-stars give us a maximum independent set.[a] STOP.

b4) There is an uncovered zero. Prime it. If there is no zero-star in its row, then go to b5). If it has a zero-star in its row, then cover the row of the zero-star and uncover the column.[b] In this process, the zero stars remain covered. Also the primes discovered till now are covered. Go to b3).

b5) We have got a zero-prime and in its row there is no 0^*. (This will lead us to a prime-star sequence or equivalently to an augmenting path.)

b51) If there is no 0^* in its column, then convert this $0'$ to a 0^*. (Thus, we have found a $0'$ in whose row or column, there is no 0^*. So we can safely add this one to our set of zero-stars. We have got a larger set of zero-stars.) Erase all primes and lines. Go to b2).

b52) There is a zero-star 0_1^* in its column. Initially this column was covered, but now it is not. So there is a zero-prime $0_2'$ in the row of 0_1^*. Suppose that there is a 0_3^* in the column of $0_2'$. As 0_2 was initially covered by the column line of 0_3^* and after that it was available as an uncovered zero, we see that the current line covering 0_3^* must be a row (as we uncover columns only). So there is a $0_4'$ in the row of 0_3^*. We continue to get a sequence $0', 0_1^*, 0_2', \cdots, 0_{2r}'$, where the column of $0_{2r}'$ does not have a 0^*. This an alternating prime-star-prime sequence. Interchange stars and primes of this sequence. Erase all primes and lines. (We have got a larger set of zero stars.) Go to b2).

[a] These zero-stars are independent by construction. Assume that we have used k lines to cover all the zeros. Consider any set of more than k zeros. Then by php, some two of them will lie on one of these lines. So they cannot be independent.

[b] Thus, this process can be taken at best n many times.

[23.1] **Theorem** The algorithm works correctly.

Proof. By construction, the set of zero-stars, at any stage is independent throughout the algorithm. Once the algorithm starts it can only stop at b3).

When the algorithm enters b3), we have only three possibilities.

Possibility 1) It has covered all the zeros by k lines, where k is the number of zero-stars at that stage. This means we cannot get more than k zero-stars. So it is maximum.

Possibility 2) We find an uncovered zero, with no zero-star in its row. In this case we go to b5) and get a larger set of independent zero-stars and start from b2) again. This step can also be used only finitely many times.

Possibility 3) We find an uncovered zero, with a zero-star in its row. In this case we change the column line segment to a row line segment. (This changes the line structures, but this can also be taken at most n times).

So essentially, when the algorithm stops, it has to stop at b3), with Possibility 1). ∎

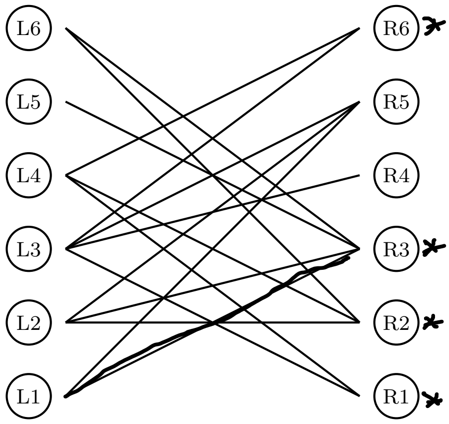
[23.2] <u>**Example**</u> `(Illustration of the star-prime algorithm.)` Consider $A = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$.

Use the star-prime algorithm to find a maximum independent set of zeros. Strictly follow the rule that the computer reads from the top-left position that is, $a_{11}, \ldots, a_{1n}, a_{21}, \ldots$. If you have used a set of lines for your purpose, you need not draw the table and lines. In stead write the lines, like, Stage k: R_1, R_3, C_1, C_4.

Answer.

$$A = \begin{bmatrix} 1 & 1 & 0^* & 1 & 0 & 1 \\ 1 & 0^* & 0 & 1 & 0 & 1 \\ 0^* & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0^* \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

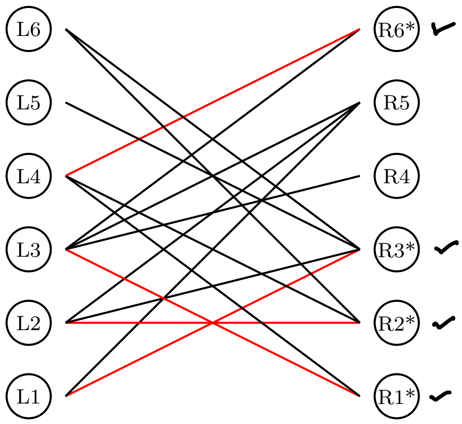
$c_1 c_2 c_3 c_6$



Initial $*$-zeros.

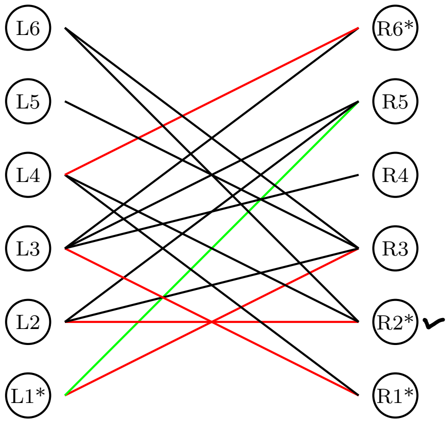
$$A = \begin{bmatrix} 1 & 1 & 0^* & 1 & 0 & 1 \\ 1 & 0^* & 0 & 1 & 0 & 1 \\ 0^* & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0^* \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

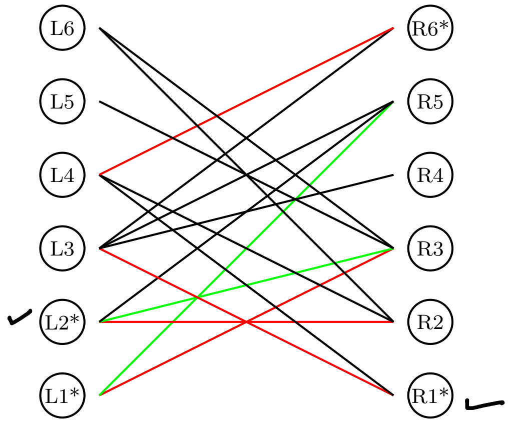
$c_1 c_2 c_3 c_6$ 9 15

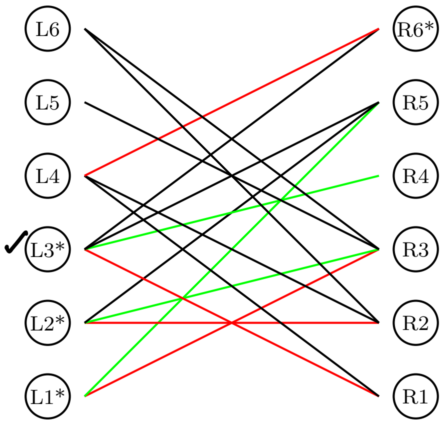


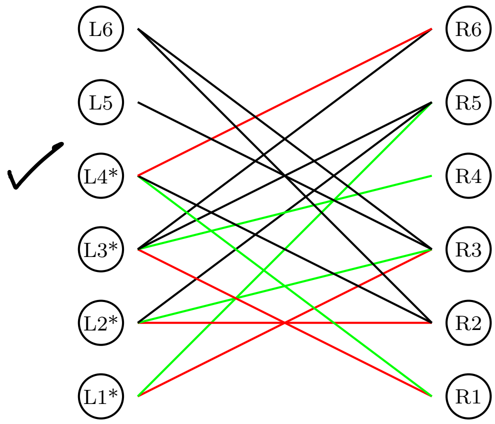
Stage	lines	′-zero
1.1	C_1, C_2, C_3, C_6	a_{15}
1.2	C_1, C_2, R_1, C_6	a_{23}
1.3	C_1, R_2, R_1, C_6	a_{34}
1.4	R_3, R_2, R_1, C_6	a_{41}
1.5	R_3, R_2, R_1, R_4	a_{53}

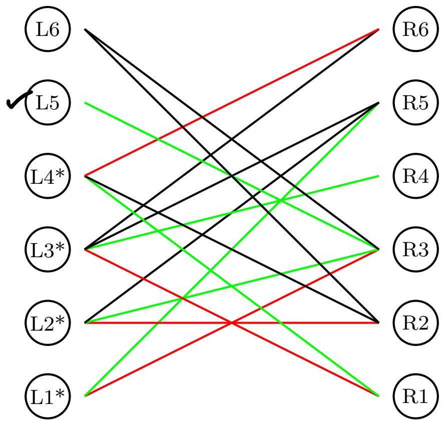
Prime an edge not covered by the stars.





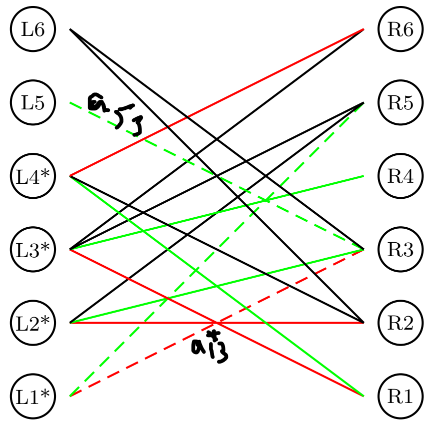






L5, R3, L1, R5

remove



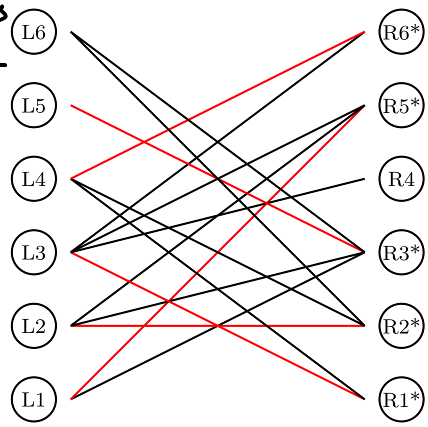
′-∗-sequence: $a_{53}', a_{13}^*, a_{15}'$

a_{53} , a_{46} , a_{31} , a_{22} , a_{15}

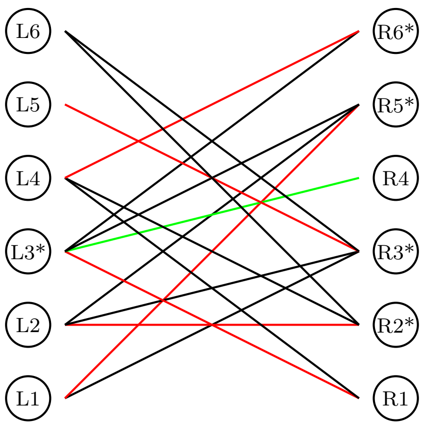
New $*$-zeros

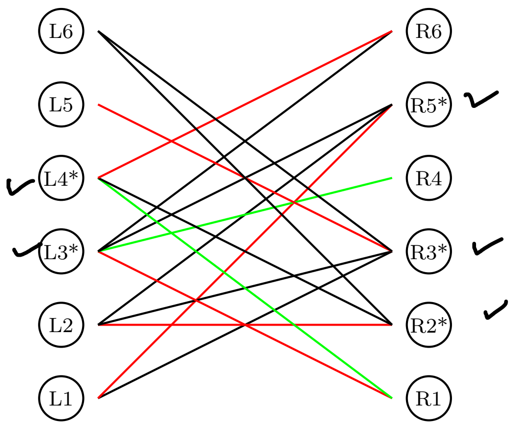
$$A = \begin{bmatrix} 1 & 1 & 0 & 1 & 0^* & 1 \\ 1 & 0^* & 0 & 1 & 0 & 1 \\ 0^* & 1 & 1 & 0' & 0 & 0 \\ 0' & 0 & 1 & 1 & 1 & 0^* \\ 1 & 1 & 0^* & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Stage 2	primes
$C_1\,C_2\,C_3\,C_5\,C_6$	a_{34}
$R_3\,C_2\,C_3\,C_5\,C_6$	a_{41}
$R_3\,C_2\,C_3\,C_5\,R_4$	none



Stage	lines	$'$-zero
2.1	C_1, C_2, C_3, C_5, C_6	a_{34}
2.2	R_3, C_2, C_3, C_5, C_6	a_{41}
2.3	R_3, C_2, C_3, C_5, R_4	none





The maximum independent set of zeros is $\{a_{15}, a_{22}, a_{31}, a_{46}, a_{53}\}$ and a minimum set of lines covering all the zeros is $\{C_2, C_3, C_5, R_3, R_4\}$.

In the graph, we have 5 vertices covering all edges and so we can at best get a matching of size 5. The 5 edges selected is a matching. Hence it is a maximum matching.

[23.3] **Corollary** Let A be a square matrix. Then the maximum number of independent zeros is the same as the minimum number of lines required to cover all the zeros of the matrix.

Proof.

Let the maximum number of independent zeros be k. Then by definition we would require at least k lines to cover all the zeros. By the bipartite matching algorithm, we can actually get k lines that cover all the zeros.■

[23.4] **Class workout** Find a maximum set of independent zeros in the following matrix.

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Answer.

$$\begin{array}{c} 0 \\ 0 \\ +1 \\ \\ \end{array}\begin{bmatrix} 0^* & 1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 0^* \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 0^* & 1 \\ 0^* & 1 & 2 \\ 1 & 0 & 0^* \end{bmatrix}$$
$$\begin{array}{ccc} 0 & -1 & -1 \end{array}$$

24 Lecture 24

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Hungarian algorithm

[24.1] **Notation** Suppose that we are given an ap with cost matrix C. For each i, replace $C(i,:)$ by $C(i,:) - \mathbf{1}^t \min_j c_{ij}$. (Subtract the minimum of ith row from each entry of ith row and do it for each row.) For each i, replace $C(:,i)$ by $C(:,i) - \mathbf{1} \min_j c_{ji}$. This is called the REDUCED COST MATRIX.

[24.2] **Discussion**

a) We already know that a minimum solution (positions in the matrix) for the new C is also a minimum solution of the old C. Only that, we have created many zeros in the new C. If we could somehow get an independent set of n many zeros, we are done. For this, we could apply the star-prime algorithm.

b) But, what would we do, if the star-prime algorithm returns that 'the maximum size of an independent set is less than $k < n$'?

c) See, by the time the star-prime algorithm stops, it would have given us a way to cover all the zeros of the matrix with k many lines: say $R_{i_1}, \ldots, R_{i_r}$ and $C_{j_1}, \ldots, C_{j_{k-r}}$. (Here R_i means row i and C_i means column i.)

d) Notice that no 0^* is at the intersection of two lines. [If a 0^* is on two lines then the remaining $k-1$ many 0^* have to be covered by the remaining $k-2$ lines, which is impossible, as no line can cover more than one 0^*.]

e) Let s be the smallest of the uncovered (by lines) numbers in the current C. (So this is positive, as all zeros are covered.) Consider subtracting s from each element of each column, (that is, from C itself) and add s to all elements of the covered rows and again to all elements of the covered columns. (So, at the intersection points, we add $2s$.)

f) The new C does not have a $-$ve entry. [This is because, from the uncovered area, you have subtracted the minimum, so you won't get a negative there. From the covered area, you first subtracted s and then added s or $2s$. So, we won't create a negative number here.] In fact, we have increased the intersection points by s. As the 0^*'s were not located at the intersection points, they would be remaining unchanged.

g) But, then we have created at least one new zero. So, we can increase the number of independent zeros. This is the main algorithm.

Hungarian algorithm to solve ap

a) Add zero rows or columns and make the cost matrix square. Compute the reduced cost matrix. Let it be C.

b) Cover all zeros in C by the minimum number of lines. (For this we shall always use the star-prime algorithm.)

c) If there is an uncovered element of the matrix then it means that there are less than n starred zeros. Since the number of lines is the same as the number of zero-stars, it follows that no zero-star is an intersection point. Let s be the smallest of the uncovered elements. Subtract s from all uncovered points and add s to all intersection points (same as subtracting s from the uncovered columns and adding s to the covered rows). This creates a new 0 in the uncovered area. Go to b).

d) If all elements of the matrix are covered then the starred elements describe the solution.

[24.3] <u>**Class workout**</u> Solve the ap for the cost matrix $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$.

Answer.

Reduced cost matrix: $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & 2 & 4 \end{bmatrix}$.

Computer assigns zero-stars and covers them by columns: $\begin{bmatrix} 0^* & 0 & 0 \\ 0 & 2 & 3 \\ 0 & 2 & 4 \end{bmatrix}$.

A zero-prime found. There is a zero-star in its row. Column changed to row: $\begin{bmatrix} 0^* & 0' & 0 \\ 0 & 2 & 3 \\ 0 & 2 & 4 \end{bmatrix}$.

Another zero-prime found. There is no zero-star in its row: $\begin{bmatrix} 0^* & 0' & 0 \\ 0' & 2 & 3 \\ 0 & 2 & 4 \end{bmatrix}$.

So a prime-start sequence found: $c'_{21}, c^*_{11}, c'_{12}$. Primes and stars in this sequence are interchanged. After that all lines and primes are removed (not the stars). $\begin{bmatrix} 0 & 0^* & 0 \\ 0^* & 2 & 3 \\ 0 & 2 & 4 \end{bmatrix}$.

Cover all zero-stars by columns: $\begin{bmatrix} 0 & 0^* & 0 \\ 0^* & 2 & 3 \\ 0 & 2 & 4 \end{bmatrix}$.

A zero-prime found. There is a zero-star in its row. Column changed to row: $\begin{bmatrix} 0 & 0^* & 0' \\ 0^* & 2 & 3 \\ 0 & 2 & 4 \end{bmatrix}$.

All zeros are covered with minimum number of lines. Let m be the minimum of the uncovered elements. Subtract m from each uncovered element and add m to each intersection points.

$$\begin{bmatrix} 2 & 0^* & 0' \\ 0^* & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

New zeros created. Find an uncovered zero and prime it. Continue.

$$\begin{bmatrix} 2 & 0^* & 0' \\ 0^* & 0' & 1 \\ 0 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 0^* & 0' \\ 0^* & 0' & 1 \\ 0' & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 0 & 0^* \\ 0 & 0^* & 1 \\ 0^* & 0 & 2 \end{bmatrix}.$$

So the optimum assignment is c_{13}, c_{22}, c_{31}.

[24.4] **<u>Theorem</u>** The Hungarian algorithm converges in finitely many steps for integer costs.

Proof. Assume that $n > 1$. Let C_1 be the matrix after Step b) and c_1, r_1 be the number of covered columns and rows, respectively. Suppose that $c_1 + r_1 < n$, so that we visit c). In c), we choose $y = \min\{$ uncovered elements $\} > 0$. The matrix C_2 is obtained by subtracting y from each uncovered element and adding y at intersections of the lines. So in this case

$$\sum_{i,j} C_1(i,j) - \sum_{i,j} C_2(i,j) = (n - c_1)(n - r_1)y - c_1 r_1 y = [n^2 - (c_1 + r_1)n]y \geq n(1)y \geq n.$$

So each time we visit c), the total cost decreases at least by n. Hence the algorithm will stop within at most $\frac{\sum c_{ij}}{n}$ iterations, where C is the reduced cost matrix with nonnegative integer entries. ∎

[24.5] **Exercise(M)** Suppose that the costs are nonnegative real numbers. I want to implement the Hungarian algorithm and get a minimum cost diagonal in finitely many steps. Can this be done?

[24.6] **Class workout** Four jobs are available and three candidates are available. The cost matrix $C = (c_{ij})$, where c_{ij} is the cost of making the ith person do the jth job. Minimize the cost while assigning exactly one job to each person. Here we introduce a virtual person.

	J1	J2	J3	J4
P1	12	9	12	9
P2	15	Unsuitable	13	20
P3	4	8	10	6

Answer.

$$\begin{bmatrix} 12 & 9 & 12 & 9 \\ 15 & \infty & 13 & 20 \\ 4 & 8 & 10 & 6 \end{bmatrix} \to \begin{bmatrix} 12 & 9 & 12 & 9 \\ 15 & \infty & 13 & 20 \\ 4 & 8 & 10 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 3 & 0 & 3 & 0 \\ 2 & \infty & 0 & 7 \\ 0 & 4 & 6 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 3 & 0 & 3 & 0^* \\ 2 & \infty & 0^* & 7 \\ 0^* & 4 & 6 & 2 \\ 0 & 0^* & 0 & 0 \end{bmatrix}.$$

Some exercises

[24.7] <u>**NoPen**</u>

a) If the cost matrix c of an assignment problem has some negative entries what do we do?

b) In case we have to maximize profit, what do we do?

c) Consider an assignment problem with reduced cost matrix $c = \begin{bmatrix} 0 & 3 & 1 & 2 & 1 \\ 0 & 3 & 2 & 3 & 1 \\ 0 & 2 & 2 & 1 & 2 \\ 0 & 2 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

Suppose that we apply the Hungarian algorithm. Is it true that we have to revise the matrix at least 6 times before we get an optimum?

[24.8] <u>**Practice**</u> Take a bipartite graph with parts of size 8 each. Consider the adjacency matrix. Interchange the ones with zeros. Check how star-prime algorithm gives you a maximum matching in the graph.

[24.9] <u>**Exercise(E)**</u> `(Wrong method)` Think of the following method to determine the smallest number of lines to cover all zeros: select a line which has the most uncovered zeros in it, cover it and repeat the process. Give an example to show that the method is not correct.

[24.10] <u>**Exercise(M)**</u> `(Wrong method)` Think of the following method to get a maximum independent set when the minimum number of lines required to cover the zeros in the matrix $A_{n\times n}$ is n: pick a row/column with minimum number of zeros. Box any arbitrary zero on it. Cross all zeros in the column/row of the boxed zero. Repeat the procedure. Give an example to show that the method is not correct.

[24.11] <u>**Exercise(E)**</u> `(Wrong method)` Think of the following method to get a maximum independent set when the minimum number of lines required to cover the zeros in the matrix $A_{n\times n}$ is n: pick a zero with minimum number of zeros in its column and row. Box it and cross all zeros in its column and row. Repeat the procedure. Give an example to show that the method is not correct.

[24.12] <u>**Exercise(E)**</u> Next semester, the department of mathematics has to offer five theoretical courses to the first year MSc students. There are four teachers available for this and they have given their top three choices to the course allocator. The "teaching value" of a particular course is 7, if the first choice is met; 4, if the second choice is met; 2, if the third choice is met; 0, else. The choices of the teachers (T1 for teacher1 and C1 for course1) are given in the following table.

Teacher	first choice	second choice	third choice
T1	C2	C4	C5
T2	C2	C3	C4
T3	C3	C1	C2
T4	C5	C1	C4

Help the allocator to allocate these 5 course to these 4 teachers in such a way that the overall teaching value is maximized and each teacher gets at least one course to teach.

[24.13] <u>**Exercise(M)**</u> `(Different cases)` There are three men who can do any of the six jobs. The time (in days) taken by each man to do various jobs is given in the table. Solve the following problems minimizing the total time taken.

i) If each man has to do only one job, find which three jobs will be left undone in an optimal assignment.

ii) If all the jobs have to be completed, find an optimal assignment so that each person gets at least one job.

iii) Find an optimal assignment if men M_1, M_2, M_3 are asked to do at least 1 job, 2 jobs, 1 job respectively, and all the jobs have to be done.

iv) Find an optimal assignment if each man gets exactly two jobs.

	J_1	J_2	J_3	J_4	J_5	J_6
M_1	10	9	7	6	8	5
M_2	7	4	8	6	5	9
M_3	3	6	9	8	4	7

[24.14] **Practice** Apply star-prime method to determine the minimum number of lines required to cover all zeros in the following matrix starting with the given position $\begin{bmatrix} 0 & 3 & 2 & 3 \\ 0 & 2 & 2 & 0^* \\ 0 & 2 & 0^* & 2 \\ 0^* & 0 & 0 & 0 \end{bmatrix}$.

[24.15] **Exercise(M)** There are 4 jobs and two men are available for it. If each of them has to be assigned exactly 2 jobs, find an optimal assignment. The amount of payment needed to assign a particular job to a particular man is shown in

Men	Job J1	J2	J3	J4
M1	5	4	5	8
M2	6	9	4	5

[24.16] **Practice** Illustrate the effect of the star-prime algorithm on the bipartite graph for the following matrix, where zeros mean the edges. Use pencils.

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Always read from top-left for searching anything.

1) First find an independent set of zero stars (a matching). In the graph draw the lines. Do not show all edges of the graph.

2) Apply the star-prime algorithm to increase the matching size by 1. Show the primes in the graph by dashed lines. Show the augmented path.

[24.17] **Practice** Solve the assignment problem with the cost matrix $\begin{bmatrix} 6 & 4 & 5 & 7 & 5 & 8 \\ 5 & 6 & 7 & 4 & 8 & 7 \\ 4 & 8 & 5 & 6 & 7 & 6 \\ 7 & 8 & 7 & 4 & 5 & 4 \\ 6 & 7 & 6 & 4 & 6 & 5 \\ 4 & 4 & 5 & 4 & 5 & 4 \end{bmatrix}$.

[24.18] **Exercise(H)** There are five candidates who can do any of the six jobs. The cost demanded by each to do various jobs is given in the table. Solve the problem minimizing the total cost giving each candidate at least one job while doing all the jobs.

	$J1$	$J2$	$J3$	$J4$	$J5$	$J6$
$C1$	6	4	5	7	5	8
$C2$	5	6	7	4	8	7
$C3$	4	8	5	6	7	6
$C4$	7	8	7	3	5	4
$C5$	6	7	6	5	5	5

[24.19] **Practice** Find the minimum cost assignment from the given matrix. The computer reads from top-left, goes to the right, then down.

$$A = \begin{bmatrix} 2 & 3 & 1 & 1 & 5 & 3 & 3 & 1 & 2 & 2 \\ 4 & 2 & 2 & 7 & 1 & 5 & 1 & 4 & 1 & 4 \\ 2 & 3 & 4 & 6 & 2 & 1 & 3 & 5 & 4 & 1 \\ 4 & 1 & 2 & 4 & 2 & 5 & 6 & 7 & 4 & 2 \\ 3 & 4 & 1 & 2 & 3 & 4 & 5 & 6 & 3 & 2 \\ 4 & 1 & 5 & 3 & 6 & 3 & 2 & 4 & 4 & 5 \\ 1 & 6 & 4 & 2 & 7 & 6 & 4 & 3 & 5 & 6 \\ 4 & 1 & 3 & 5 & 2 & 3 & 4 & 2 & 5 & 3 \\ 3 & 3 & 1 & 4 & 3 & 3 & 4 & 6 & 5 & 4 \\ 2 & 1 & 2 & 2 & 4 & 3 & 3 & 5 & 6 & 3 \end{bmatrix}$$