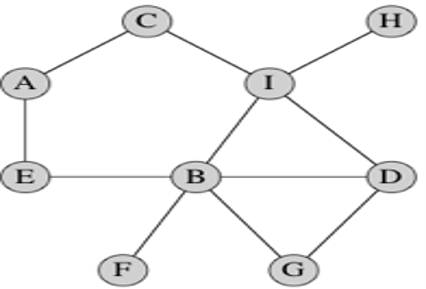
1. (4 pts) Given a Graph G below, Draw its Adjacency Array Representation.



Answer:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I |
| A | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| B | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| C | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| D | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| E | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| F | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| G | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| H | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| I | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |

1. (2 pts) A Map task has? Mappers whereas the Reducer task has? Reducers.

Answer: 1) many and 2) one or more

1. (4 pts) There are two principal costs,
   * 1. **Computation cost**of mappers, reducers, and the system. System cost depends on the operations,  
        a. ? key-value pairs by key and b. ? the key-value pairs
     2. **Communication cost**depends on transferring key-value pairs from? to?

Answer: 1) sorting, 2) merging, 3) mappers and 4) reducers

1. (2 pts) Communication cost often dominates the computation cost. Why?

Answer:

1. (2 pts) Replication rate is ?? of key-value pairs created by each mapper denoted by **r. r** represents ????

Answer: 1) the average number and 2) the communication cost per input

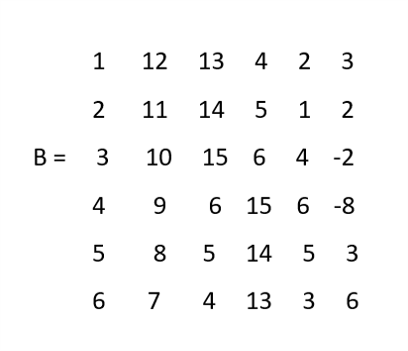
1. (2pts) If each reducer has size **q** if there are **p** reducers, and **I** (capital i) is the number of inputs for the problem, then **r=??/I.**

Answer: pq

1. (2 pts) A mapping schema for a problem with a reducer of size **q** is an assignment of inputs to sets of reducers, with two conditions:
   1. Maximum **q** inputs per reducer.
   2. For every output, there is some reducer that receives all of the inputs? with that?

Answer: 1) associated and 2) output

1. a. (8 pts) Assuming a reducer size of q = 2, trace the map-reduce algorithm discussed on module 8 to multiply the two matrices below. Show all your work.



Answer:

q = 2 and n = 6, so g = 2 \* n2 / q = 2 \* 62 / 2= 36

b. (2 pts) Specify the communication cost of this algorithm.

Answer:

1. (22 pts **extra credit**) Provide the details of the proof that a one-pass matrix-multiplication algorithm requires replication rate at least r ≥ 2n2/q, including:
2. The proof that, for a fixed reducer size, the maximum number of outputs are covered by a reducer when that reducer receives an equal number of rows of M and columns of N.
3. The algebraic manipulation needed, starting with ∑ki=1q2i ≥ 4n4.

Answer: