Gossips and Graphs

Teaching an Optimal Solution to the Gossip Problem

Nathan Gildersleeve

Introduction Activity

In your group you each have a different piece of information about a murder that happened in the game of Clue.

Try to spread the information around in as few calls as possible.

Rules:

- Only 2 people can be on a call at a time.
- Any time 2 people are on a call, they share all the information they have so far.
- You are done when everyone has the whole story.

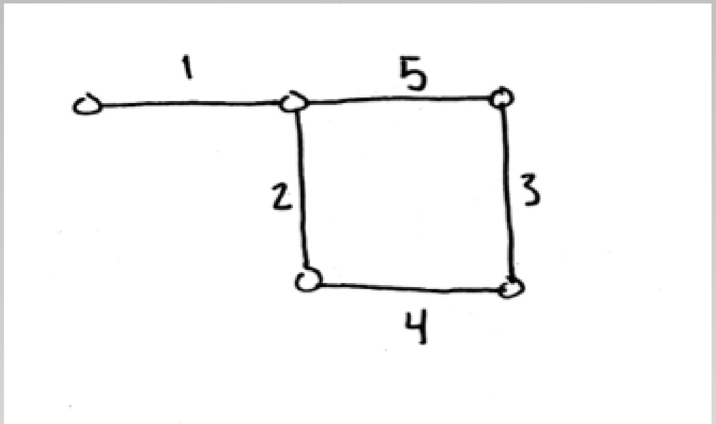
Imagine we are trying to transmit information throughout a group of people.

We assume that everyone knows one part of a story, and the goal is for everyone to know the entire story.

In order to represent this situation, we can let people be vertices in a graph and let edges between them be calls that they make.

Introduction to the Theorem

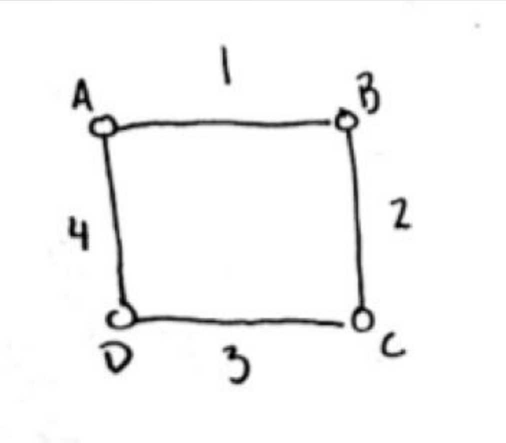
Sample **ordered graph**:

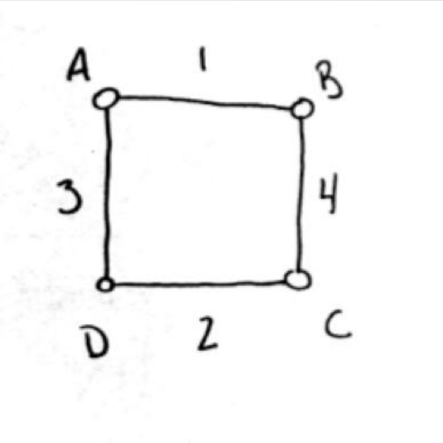


So the 1st edge represents the 1st call, the 2nd edge the 2nd call, and so on.

Introduction to the Theorem

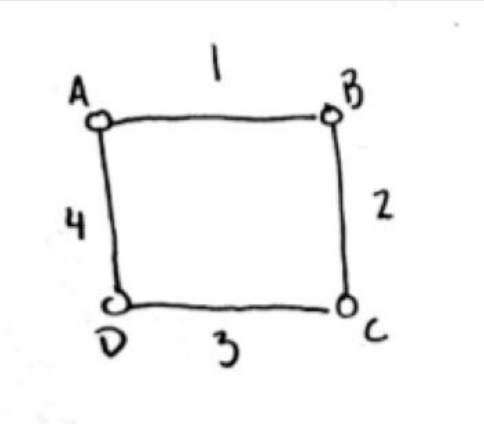
The edges between vertices are numbered to represent the order that the calls take place. This is because the order of calls matters.





Introduction to the Theorem

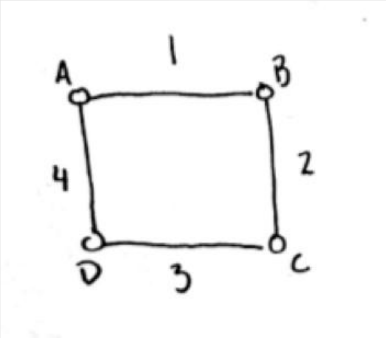
Consider the 4-cycle numbered as indicated below. Notice that there is no increasing path between two vertices D to B for this numbering.



Introduction to the Theorem

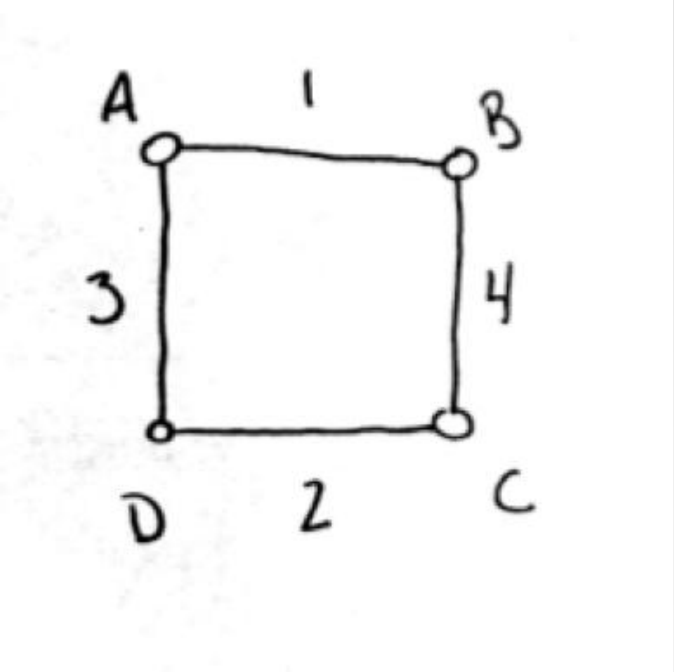
It means that B never hears D’s information.

As information moves from D, D’s call with A is A’s last call, and D’s call with C is C’s last call, so there is never a call to B telling D’s information.



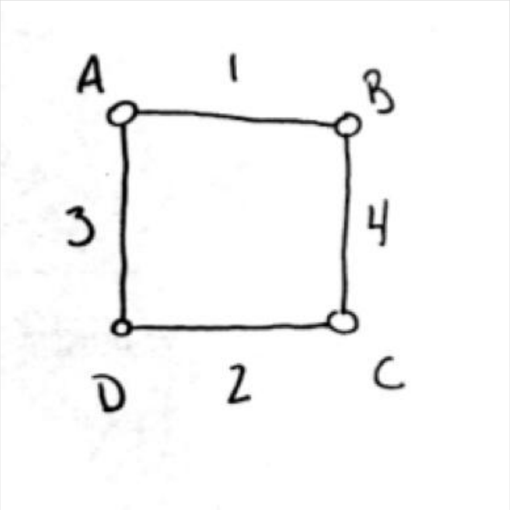
Introduction to the Theorem

The following numbering does much better:



Introduction to the Theorem

In contrast to the previous example, we can see that C’s call to B happens after her call with D, so B does hear D’s information.



Introduction to the Theorem

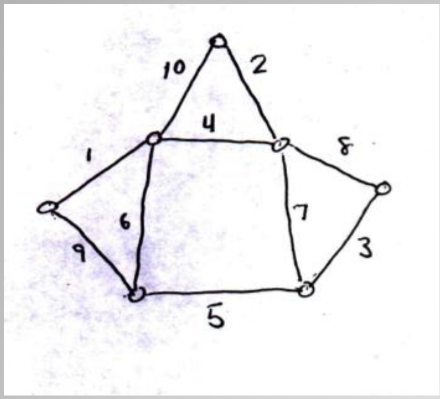
Introduction to the Theorem

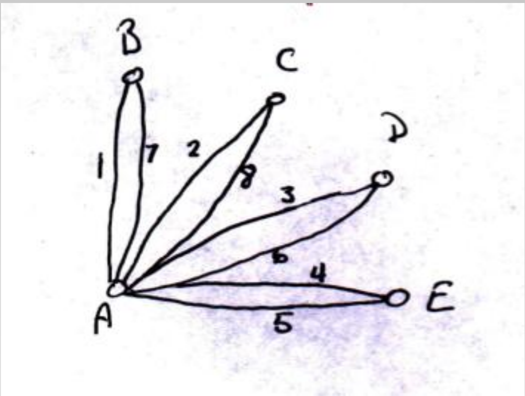
A **gossip scheme** is an ordered graph where there is an increasing path from any vertex to any other.

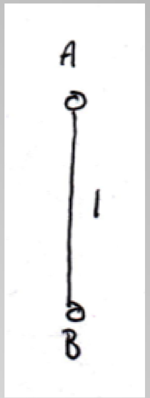
Remember, since each numbered edge is a call, such an increasing path ensures there is a sequence of calls carrying information from any person to any other.

So if we didn't have an increasing path from any point to any other, then we know that someone's information didn't make it to someone else.

Gossip Schemes:

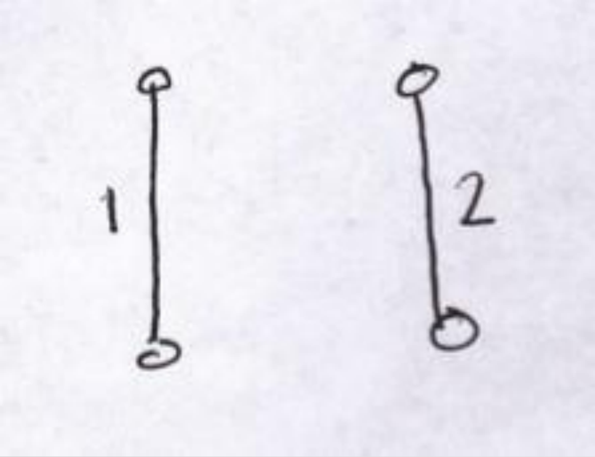


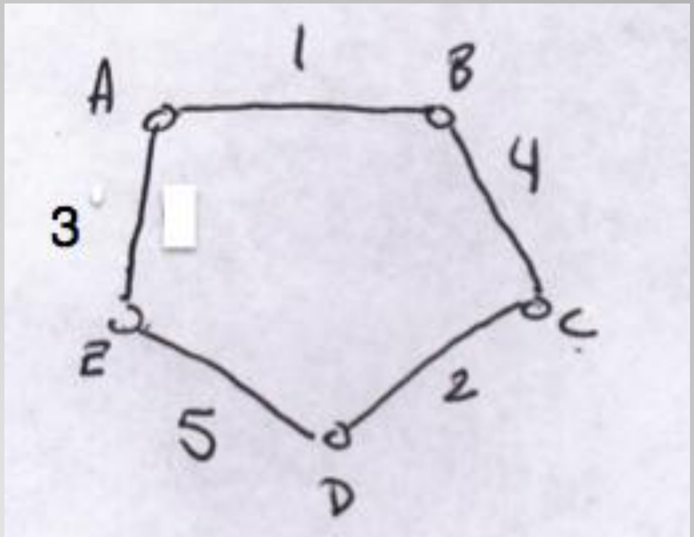




Introduction to the Theorem

Not Gossip Schemes:





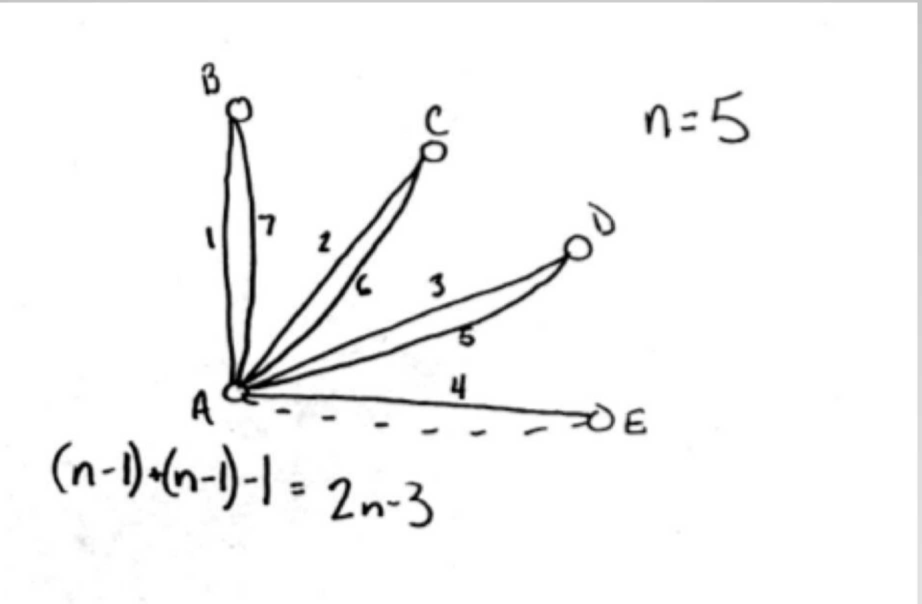
Introduction to the Theorem

One method (that we can be certain always works) to create a gossip scheme for *n* people is called the "busy body" method.

In this construction, we simply have every person call one person, a “busy body”, and then have that person call everyone back.

Introduction to the Theorem

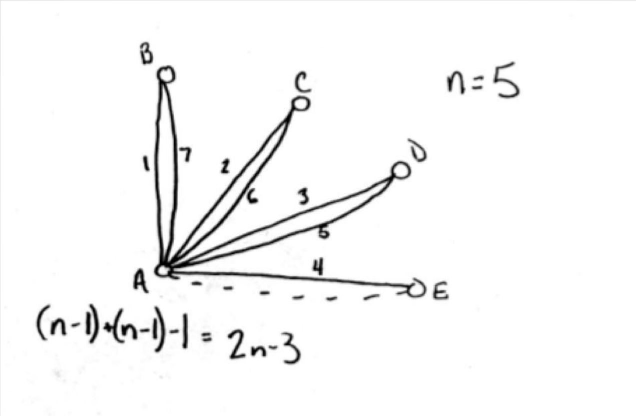
We can see a diagram of this in the graph below when we have 5 gossips:



Introduction to the Theorem

This “Busy Body” method results in 2n-3 calls.

There are n-1 calls in, and n-1 calls back out, minus another 1 since the last person to call in doesn’t need to be called back:



Introduction to the Theorem

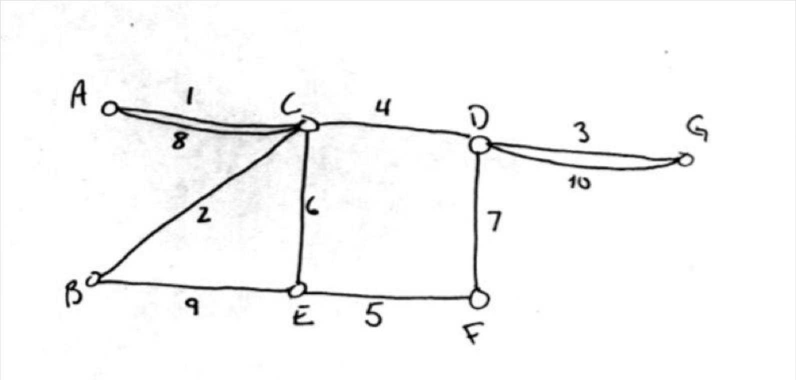
So we have a scheme that will always give us 2n-3 calls for n people.

But is this number of calls the fewest possible for n people?

Is there no gossip scheme with fewer edges?

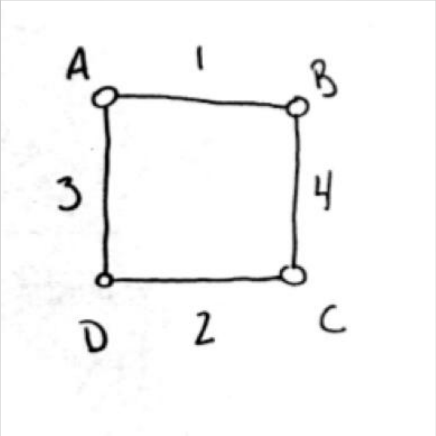
Introduction to the Theorem

It turns out we can get further savings if we have every person call one of four people, then have those people share their information with each other and finally call everyone back, as so:



Introduction to the Theorem

The key here is that this particular ordering on a 4-cycle is highly efficient:



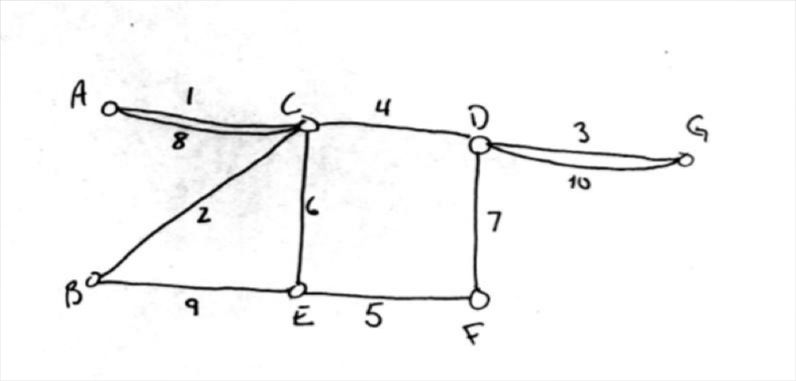
In fact a 4-cycle it is the most efficient ordering we can get.

Introduction to the Theorem

Using this "busy body task force" method, where everyone calls one of four people, we end up with 2n-4 calls.

There are n-4 calls in, n-4 calls back out, and the 4 calls within our task force:

$(n-4)+(n-4)+4 = 2n-4$



Introduction to the Theorem

Gossip Theorem: For $n \geq 4$, the minimum number of edges in a gossip scheme on n vertices is $2n$-4.

Gossip Theorem

Why this works:

- Essentially in order to have fewer calls, we would need a gossip scheme on an n-cycle (n>4) with n calls.
- We can play with a 5-cycle and see that this doesn't seem like it could work, but proving it requires many technical details.

Gossip Theorem

The structure of the proof is by contradiction:

- Suppose we have a better solution.
- We can make 3 claims we know for sure about this new solution.
- Those three claims result in a contradiction.
- Therefore 2n-4 is the optimal solution.

Gossip Theorem

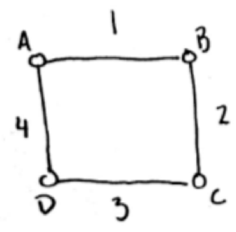
Suppose 2n-4 isn’t our optimal scheme…

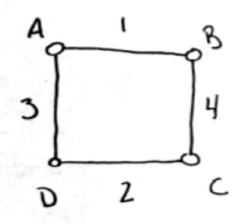
- Then we add calls to the supposed gossip scheme until we have a graph with 2n-5 edges. Call this gossip scheme G.

Gossip Theorem

Claim 1: G satisfies the **NOHO property**.

The **NOHO property** stands for **N**o **O**ne **H**ears their **O**wn information. In a graph this means no one's information is propagated back to him or her. We formalize this by saying there is no increasing closed walk in our ordered graph:



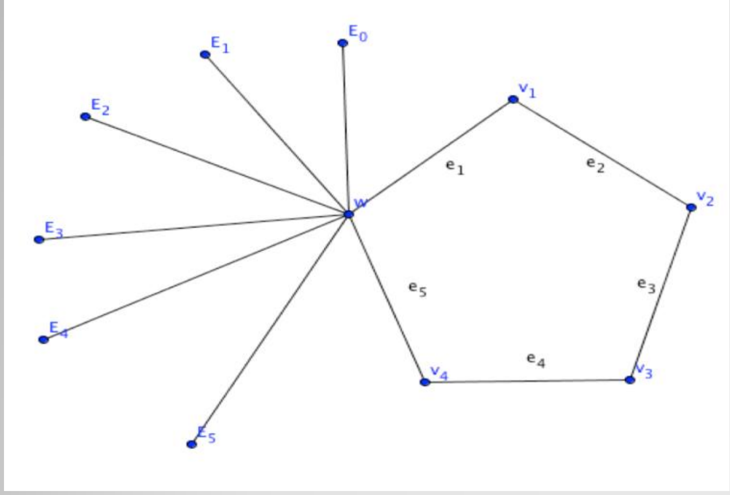


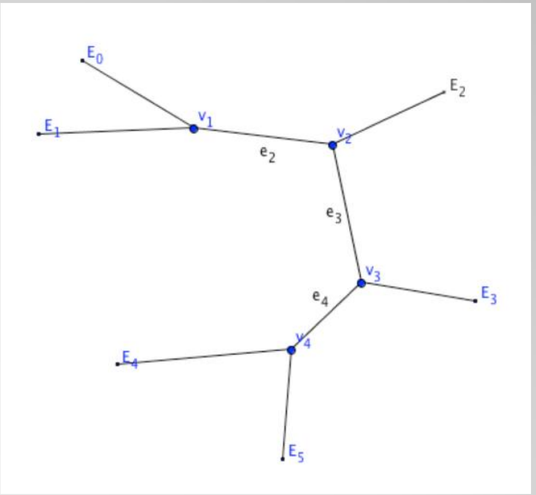
The above graph doesn't satisfy the NOHO property since we have a 1234 path around the loop. A can't hear C or D's info without hearing her own.

This graph does satisfy the NOHO property since there is no increasing path around the loop. No one has to hear their own to hear anyone else's information.

Gossip Theorem

If it doesn’t satisfy NOHO, we can show that we could create a subgraph on n-1 vertices with too few calls. It would contradict our induction hypothesis.





Gossip Theorem: Claim 1

Claim 2: All but 3 calls are useless to any given vertex u in G, hence the degree of G is greater than or equal to 3.

Degree of a vertex is the number of vertices connected to it. The degree of a graph is the minimum degree of all vertices.

Let I(u) be the spanning tree of all useful calls coming into u. Let O(u) be the spanning tree of all useful calls going out from u. They both have n-1 edges by the definition of spanning tree.

The intersection of I(u) and O(u) is all edges adjacent to u. ie the degree of u.

So if we take 2n-5 total edges, take out n-1 and another n-1 and put d(u) back in, we have d(u)-3 useless edges.

Gossip Theorem

Claim 3: The subgraph obtained by deleting the first and last call made by each vertex has at least 5 components and no isolated vertex.

A matching is a set of disjoint edges covering a graph.

A call must be the first call for both or neither of its callers, so the set of first calls is a matching. Likewise the set of last calls is a matching.

If we take out these calls we have 2n-5-(n/2)-(n/2)=n-5 edges. This subgraph must have at least 5 components. It has no isolated vertex since we only removed 2 edges at most and we have 3 connecting each vertex.

Gossip Theorem

Contradiction!

- If we find a vertex with degree 3 (which we can) we know it can have 0 useless vertices from Claim 2.
- By looking at the components from claim 3, I can show only 3 components are useful. So all the edges in the other components are useless, which is a contradiction. (Because no edges can be useless to this vertex.)

Gossip Theorem

Which is all a little fast and furious™, but the point is that if we take our intuition about trying to find better savings by finding a solution with n calls for n people, and want to show exactly why it is impossible, this is where we end up.

Gossip Theorem

In the curriculum I taught, I wanted to show students how to represent gossip and information flow using graphs, and help them find the optimal scheme in such a way that they could appreciate it.

To that end, I had them perform 6 Tasks, with discussion in between.

The Curriculum

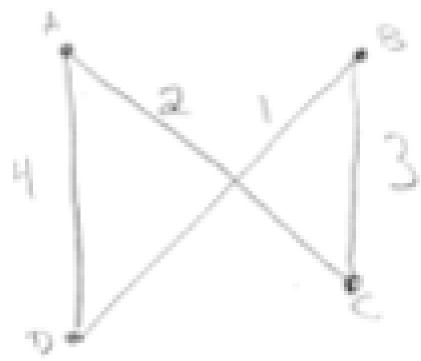
Task 1: You performed as an intro to this presentation.

The Game of Clue

The Curriculum

Example of Student Work on Task 1

Write down your group's method and draw a diagram that shows how you shared your information (maybe use a graph theory to help).



The Curriculum

Task 2:

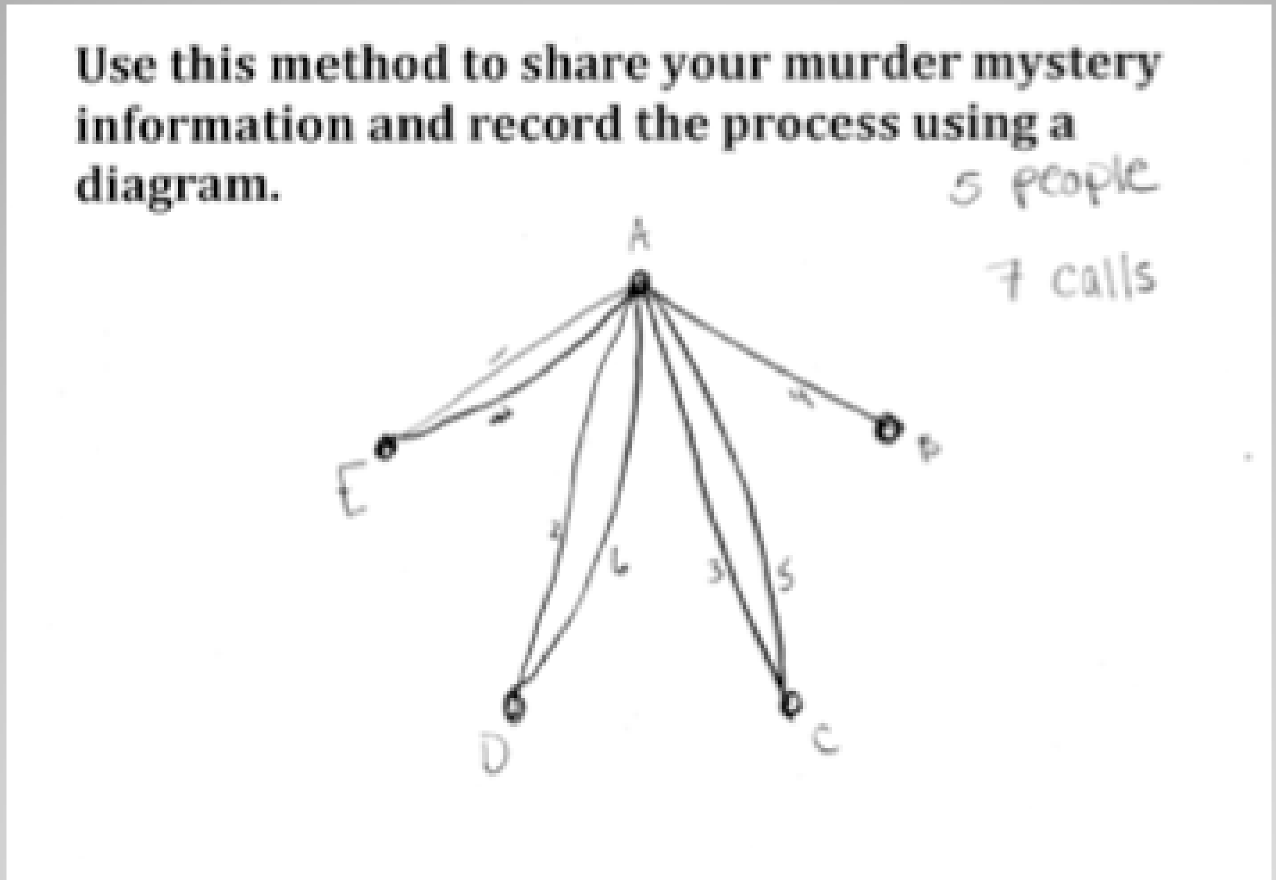
One way to solve this problem is by making one person in the group be the Busy Body.

In this method everyone calls one person in the group. After they have all the information the Busy Body calls everyone back.

Use this method to share your murder mystery information and record the process using a diagram.

The Curriculum

Example of Student Work on Task 2



The Curriculum

Task 3:

For a group of n people, the Busy Body method takes (n - 1) + (n - 1) - 1 calls.

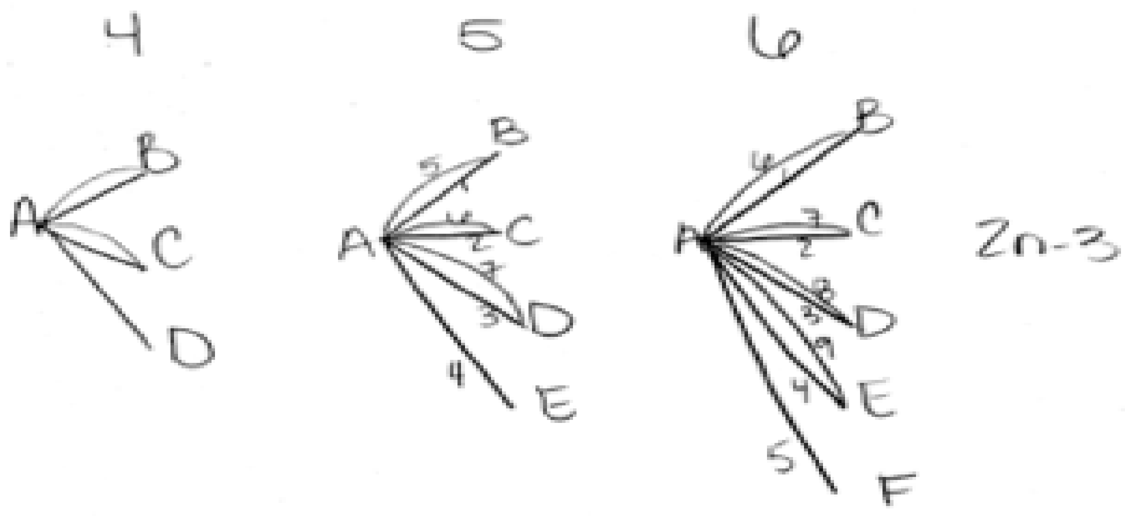
Draw a diagram to convince yourself that it will always take this many calls.

The Curriculum

Example of Student Work on Task 3

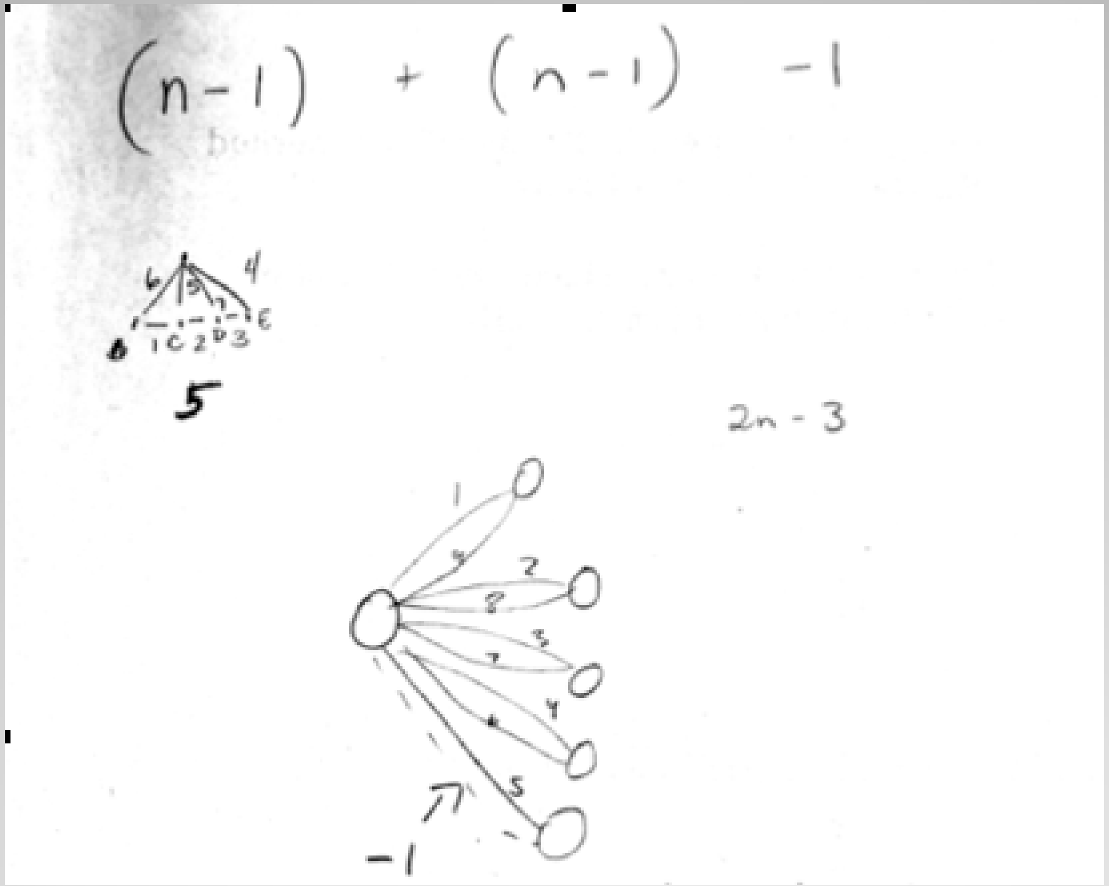
For a group of n people, the BusyBody method takes (n - 1) + (n - 1) - 1 calls.

Draw a diagram to convince yourself that it will always take this many calls.



The Curriculum

Example of Student Work on Task 3



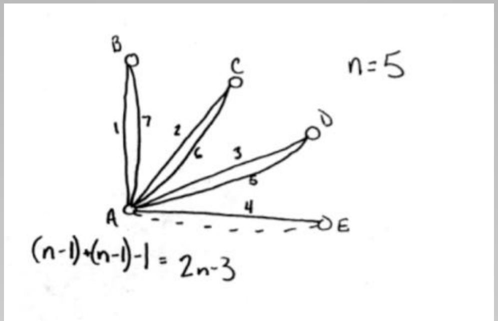
The Curriculum

Task 4:

With 4 people, the Busy Body method would take (4 – 1) + (4 – 1) – 1 = 5 calls to share all the information.

Can you beat that?

The Curriculum



In the graph above, we want to be able to show that all of our callers have the whole story.

We can use a table to track the information that each person has after any given call.

In order to have a gossip scheme we need to have abcde in every cell in the final column.

Caller	Call 0	Call 1	Call 2	Call 3	Call 4	Call 5	Call 6	Call 7
A	a	ab	abc	abcd	abcde	abcde	abcde	abcde
B	b	ab	ab	ab	ab	ab	ab	abcde
C	c	c	abc	abc	abc	abc	abcde	abcde
D	d	d	d	abcd	abcd	abcde	abcde	abcde
E	e	e	e	e	abcde	abcde	abcde	abcde

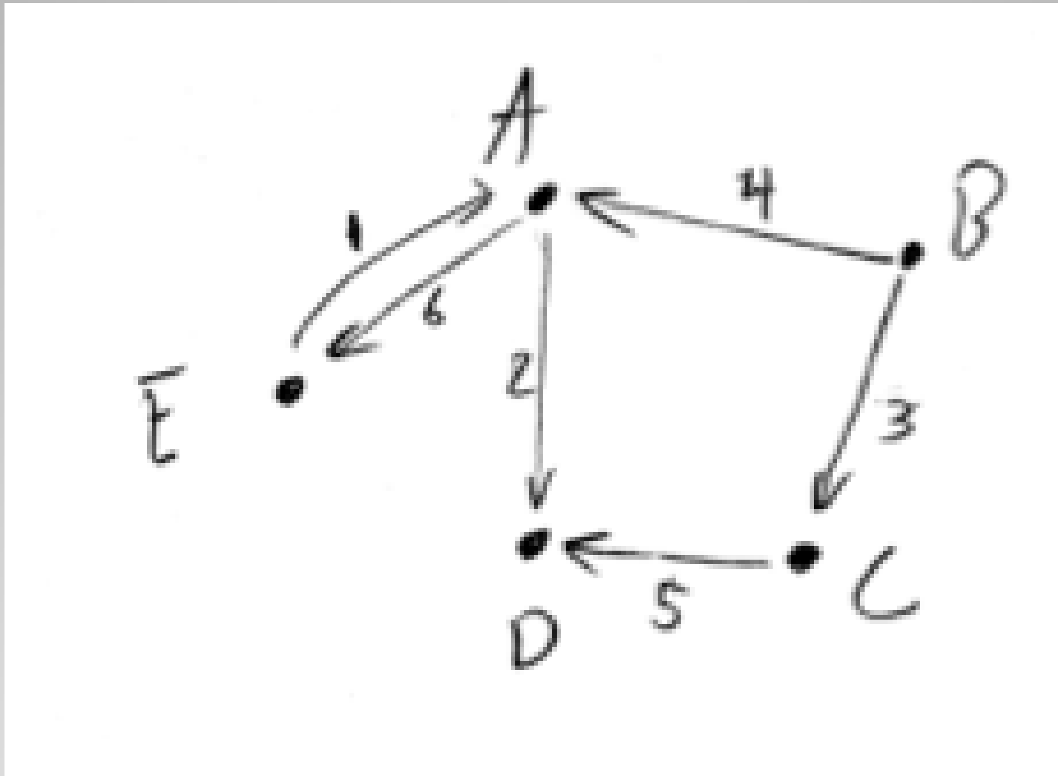
Task 5:

With 5 people, the Busy Body method would take (5 – 1) + (5 – 1) – 1 = 7 calls to share all the information.

Can you beat that? If so, show how. (Hint: Yes. Yes you can.)

The Curriculum

Example of Student Work on Task 5



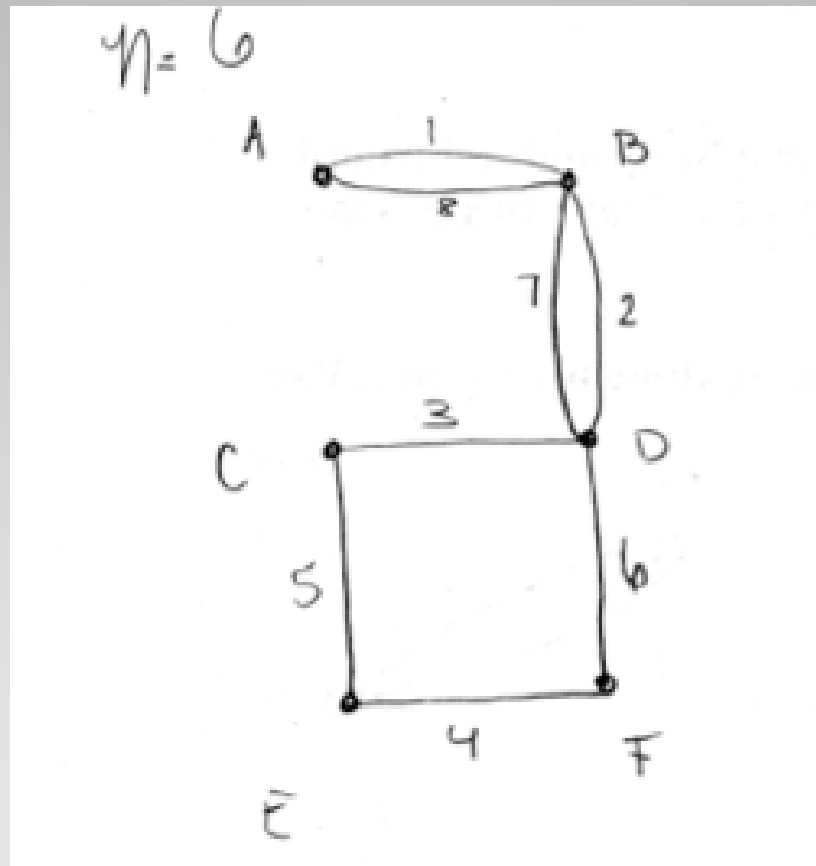
The Curriculum

Task 6:

Now that you know it is possible, try to share all information with 6 people using 2(6)-4= 8 calls.

The Curriculum

Example of Student Work on Task 6



The Curriculum

Reflections on the Curriculum

- Accomplished broad objectives.
- Need more time spent verifying we have gossip schemes.
- Follow up activity on structure and verification.
- Higher level curriculum on developing proof from basic conjectures.

The Curriculum

Thank you!