Using module analysis to identify patterns in students' responses when adding and subtracting vectors in arrow format

Emily Frederick

University of North Florida

Introduction

- A fundamental component of the mathematics used in physics involves vectors.
- Students struggle with vectors in their beginning physics classes.
- We look at students mistakes in an algebra-based physics class on an assessment involving vector addition and subtraction.
- We want to find any modules within the non-normative (incorrect) responses to identify common student mistakes.
- Using network analysis, we analyzed these students' responses to a vector addition and subtraction assessment.
- We followed the Module Analysis for Multiple Choice Responses (MAMCR)[1] method for data analysis.

Two datasets

- Separate datasets for the 1D and 2D questions
- 1D question set had twelve questions.
  - Each had four responses
- The twelve questions were six different vector sets (A & B) with either the addition or subtraction operation (A + B or A – B)
- 2D question set had sixteen questions.
  - Each had eight responses
- Both questions sets had two questions with the same vectors $\vec{A}$ and $\vec{B}$ ($\vec{A}+\vec{B}$ or $\vec{A}-\vec{B}$)
- Vectors $\vec{A}$ and $\vec{B}$ had varying relative orientations

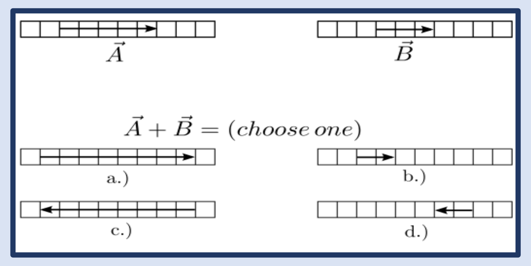

FIG. 1 Example of 1D question

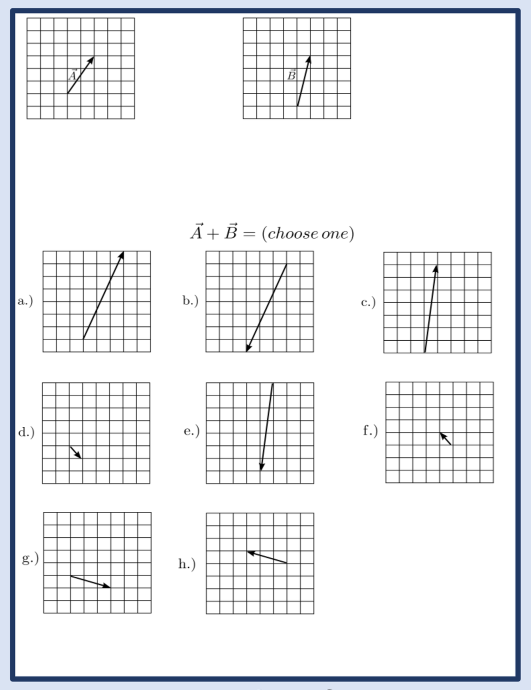

FIG. 2 Example of 2D question

Network analysis and bipartite network

- We create our network with two different types of information (students , responses) called nodes
- Nodes are connected with edges or links
- The two node types (circles and squares) are connected to each other based on what the student chose as their response for a particular question (Fig. 3)
- We create a single partite network from the bipartite netowrk
- We want to connect responses to each other and give the edges weight to count how often a connection is made

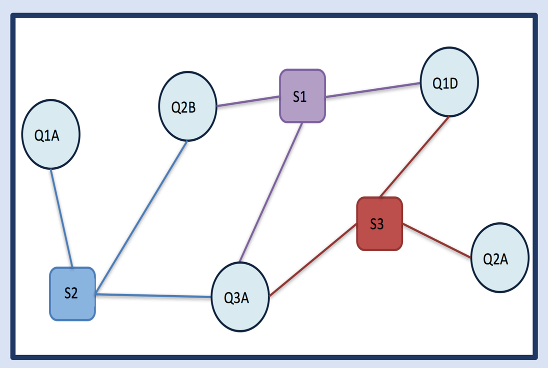

FIG. 3 Example of a bipartite network. Students and responses are shown as square and circle nodes respectively

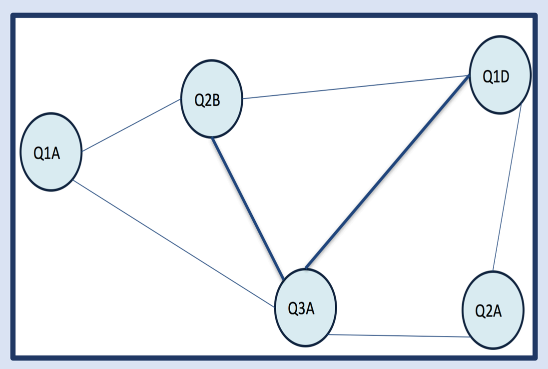

FIG. 4 Example of a single partite network. Student nodes were removed to create edges with weights

One dimensional dataset

Network characteristics: 1D dataset

- We had 52 student responses for our dataset
- We initially had 48 response nodes
- Once we remove normative responses from the network we had 36 nodes in the network.
- After we removed the isolated nodes from the network we had 23 response nodes left in the network.
- Therefore there are 13 responses that weren't chosen by any students.

Backbone extraction

- We extracted the backbone of the network following the Locally Adaptive Network Sparsification (LANS) algorithm.
- In our network, we kept any edges with weight greater than the 99% quantile for the individual nodes
- We also kept any edge that is significant to at least one node.

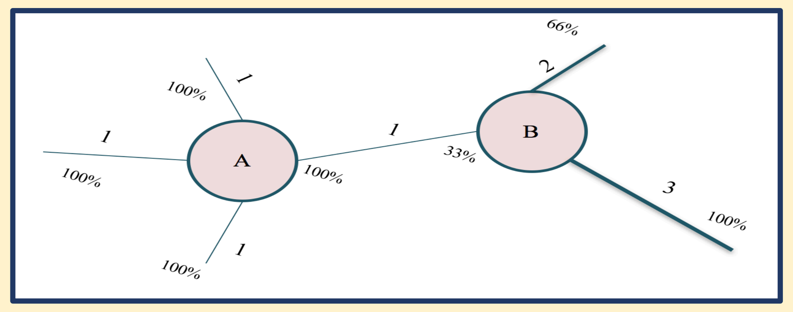

FIG. 5 Since all but two edges are below the 99% quantile, we would remove the two edges unless its significant to at least one node.

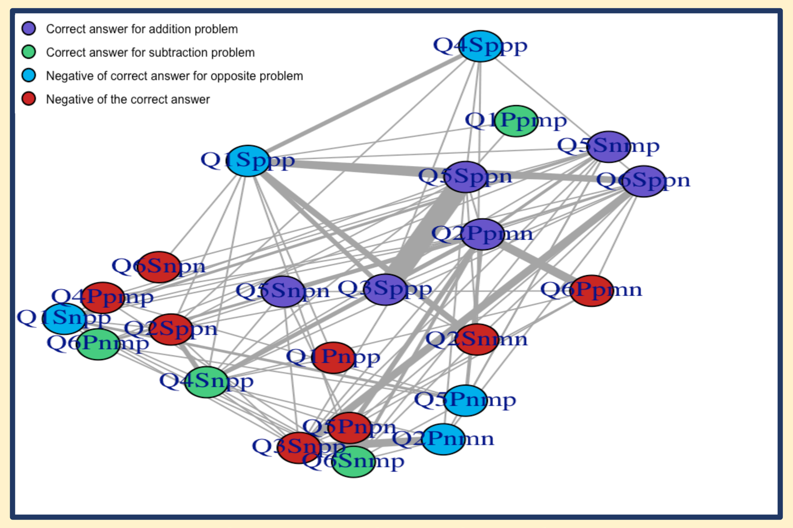

FIG. 6 Initial network with isolated nodes removed

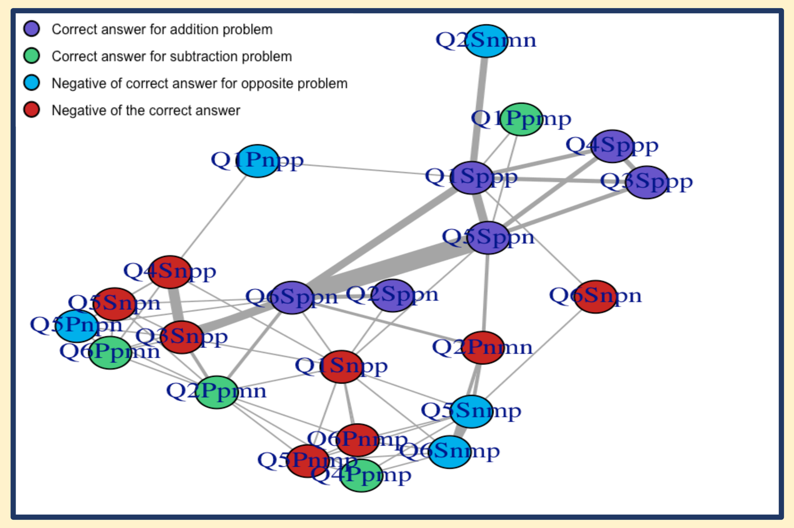

FIG. 7 Backbone extracted from initial network

- Extracting the backbone removes edges that are not significant to any nodes.
- This gives us the underlying structure of the network to analyze

Identifying modules within the backbone network

We used a program, infomap, to detect subgroups with a network. The program takes random variations of "walks" on the edges of the network to determine which nodes are often connected based on the weights of the edges.

- Three groups were detected within the backbone network.
- The first subgroup, colored in purple, had eight responses in the group.
- Within those eight responses, seven of the responses are the correct answer for the opposite problem type.
- The second subgroup, colored in red, have eight responses in the group.
- Within those responses, five of them are the negative of the correct answer.

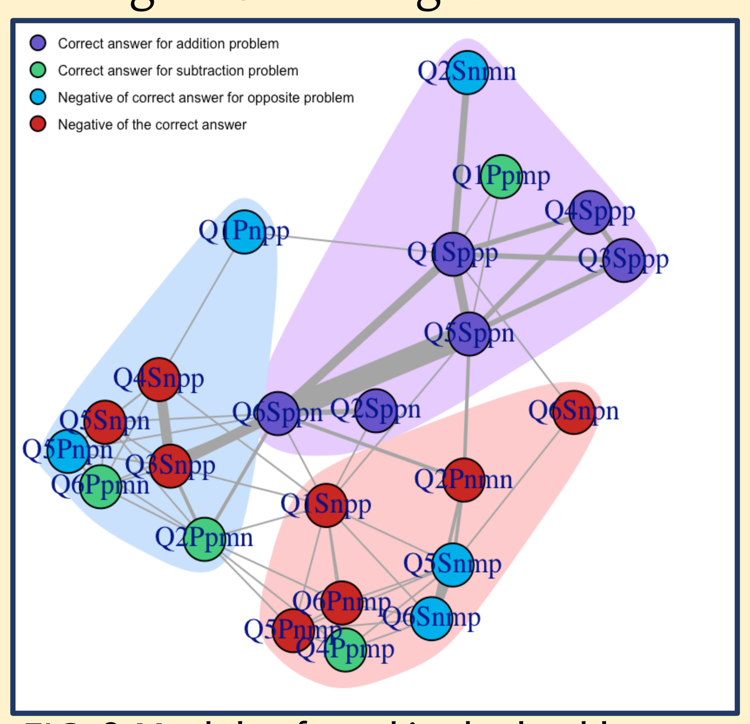

FIG. 8 Modules found in the backbone

- The third subgroup, colored in blue, is the "left over" group. These responses weren't grouped in the other two groups.
- Since they were all different types of responses, we couldn't make any conclusions on this specific group.

Two dimensional network analysis

- The assessment designed for the 2D analysis had many errors with a number of the non-normative responses.
- The errors and the lack of a larger dataset made it difficult to analyze data from the 2D network.

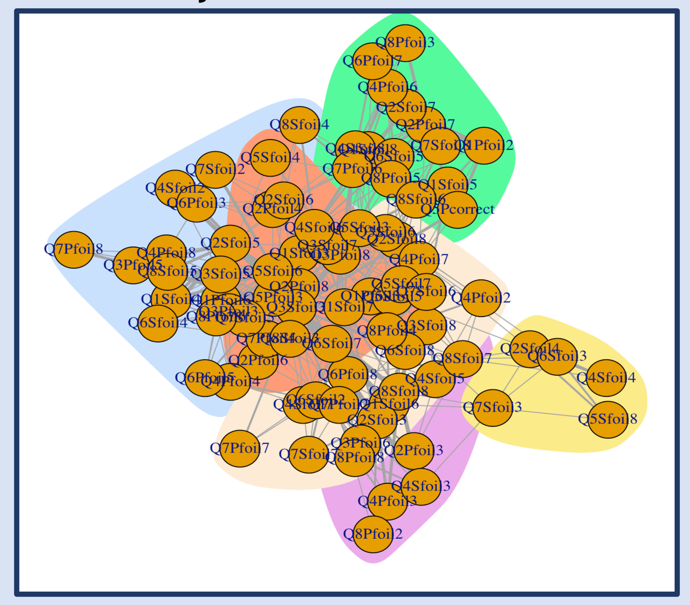

FIG. 9 InfoMap detected six different groups for the 2D network

- The 2D network originally had 128 nodes.
- The backbone had 83 nodes with 382 edges.
- There is too many connections between the nodes to make any justified claim on the groups
- We hope to recollect data with the 2D questions.

Limitations

- MAMCR analysis cannot determine why students chose specific responses together
- The assessment was designed to compare *ijk* format style questions to arrow format questions
- Some common arrow format errors that are noted in the literature are not present in the assessment
- 2D vector questions requires a large data set in order to verify any conclusions

Next steps

- Redesign assessment to include common responses and remove distractors that weren't often chosen.
- Collect student work (written or interview) to identify how students are arriving at his/her answer

Conclusions

- In the one dimensional questions, student responses fell into two main subgroups
- Group one, students often performed the wrong operation. This can also be attributed to students performing a "tail-to-tail" method with the correct operation.
- Group two, students gave the negative of the correct answer, thus they performed *B - A* instead of *A – B*.
- Though there are some consistency in a few groups within the 2D network, we need to collect more data to verify the modules.

References

[1]Eric Brewe, Jesper Brunn and Ian G. Bearden. Physical Review Physics Education Research 12, 020131 (2016).

Acknowledgements

NSF Grant: DUE 1560142

North Dakota State University

John B Buncher