Chord Diagrams as a Visual Tool in Chemistry Education Research

Emily Kable and Jordan Harshman

Chemistry education presents challenges to students in large part due to content relationships across multiple semesters of instruction. For instructors to improve their assessment, they need a tool that allows them to visualize how their students are thinking. Time-and resource-intensive qualitative research is the gold standard for these types of analyses, leaving many instructors without the possibility of in-depth assessment of their large classes for short answer questions.

Historically, word clouds have provided a solution to this problem and are used to represent key connections between concepts. This graphical representation of qualitative data displays the frequency of concepts by the size of the words under the assumption that the larger words are ideas more commonly raised. However, word clouds lack context between question and student responses and are only useful for one-word responses. If words need to be grouped together, the responses need to be hyphenated, making it harder for the reader to comprehend.

A novel alternative to the word clouds are chord diagrams. Chord diagrams arrange data around a circular axis showing links across different related entities. These connections are shown through different colored arcs that connect pieces of the data together. The chord diagram improves the visual representation of the data by showing the context and relevance of the data through data point connections. Chord diagrams can also accommodate more than one-word responses on the radial graph.

To compare these graphical methods, we surveyed a sample of students taking Fundamentals Chemistry I. The survey involved 10 questions covering electrolytes and the quantum mechanical model of the atom, and required students to respond in short phrases. Each question was exported and presented as a word cloud and a chord diagram to determine overall effectiveness of each.

Once the survey responses were recorded, the data were cleaned and exported through Microsoft Excel® to R studio® to generate a chord diagram and word cloud. Data cleaning entailed spellchecking, removing general outliers, stemming of words that had the same contextual meaning, and removing words that provided no contextual importance to the chord diagram or word cloud. Once the data were cleaned, the chord diagram was generated using R, and the word cloud was generated using www.wordclouds.com.

Figure 1 shows the word cloud and Figure 2 shows the chord diagram generated from the question: "What is the difference between a strong and weak electrolyte?" From analyzing the word cloud, the largest terms seen are "complet," "strong," "water," and "dissoci." The word cloud provides no real context for how the students were answering the question. When looking at the chord diagram, the thickest arcs connecting between two points are "strong" to "completely" and "weak" to "partial." Much thinner arcs connect "strong" to "partial" and "weak" to "completely," providing context for an instructor to understand that majority of the students are responding to this question in a similar and accurate way of thinking.

Our results showed that chord diagrams drew easier-to-understand connections between the data collected from the survey to analyze student responses. Chord diagrams are a useful graphical tool to represent student thought processes in how they answer a question, while they also highlight common misconceptions. Word clouds lack the ability to show the connections between the data sets for a particular question and cannot highlight the frequency of correct versus wrong in one graph.

Statement of Research Advisor

Emily showed a passion for taking on a complex statistical programming language and was successfully able to understand the code well enough to manipulate basic characteristics of the code. Her contributions will hope-

fully lead to innovation in the ways that instructors can receive immediate feedback for student submissions to open-response items.
–Jordan Harshman, Chemistry and Biochemistry

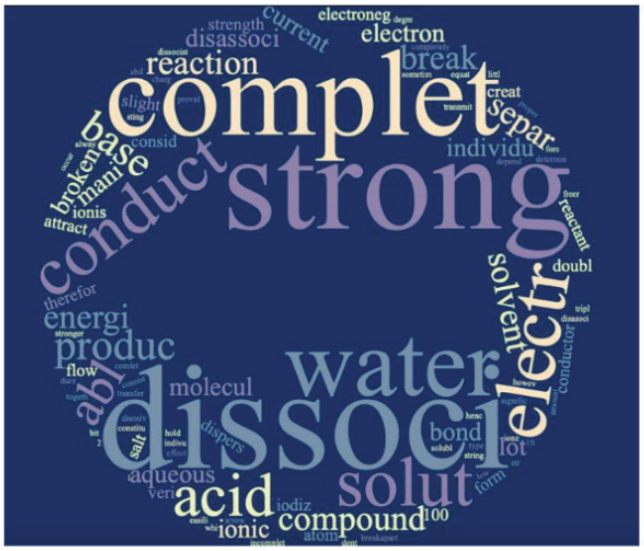

Figure 1. Question 1 Word Cloud. The larger the size of the terms, the higher the frequency of the terms used in responses.

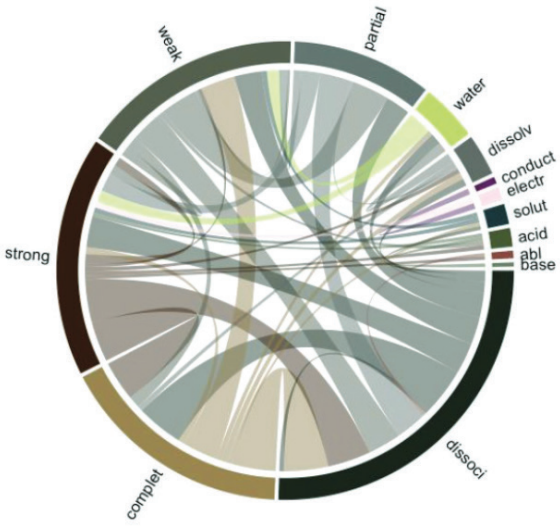

Figure 2. Question 1 Chord Diagram. The thicker the arc between two terms, the higher the frequency the students responded connecting those sequences of terms.