Diagram-based Input for Large Language Models to Support Accessible STEM Learning

Sarah E. Wegwerth, Alexa Urrea, and Julia E. Winter*
Alchemie Solutions, Inc., Bloomfield Hills, MI

Abstract

To meet the accessibility needs of students who are blind or have low vision (BLV), detailed textual descriptions of STEM diagrams within interactive learning tools are created in real-time and correspond to the configurations of the interactive software system. The descriptions are read by screen readers as alternative (alt) text to provide information for BLV students to compose mental representations of the diagram. These descriptions provide a unique bridge from the visual language of STEM diagrams to natural language of Large Language Models (LLMs). By interfacing with an LLM, these descriptions are used for personalized exploration by the BLV user and to guide all learners through a defined pedagogical pathway. Results from a usability study with four BLV adults are reported.

Introduction

Educational publishers have moved to delivering content through digital methods. As they ramp up investment in providing interactive and visually engaging content that has been shown to improve learning (Cooper, Stieff, and DeSutter 2017) and proven to be wanted by students (Martin-Samer, Casada, Gomez-Pozeulo 2023), publishers face the difficult technical problem of making new digital content accessible to students who are blind or have low vision (BLV) (Lieberman 2019). For STEM content, which relies heavily on diagrams, a BLV student depends on alternative (alt) text descriptions to access diagrams in the digital content. Alt text is added to the meta-data of an image, and then a screen reader delivers that description and other textual information to the BLV student, usually through an audio interface. Current solutions for automatically generated alt text for STEM diagrams do not produce enough pedagogical detail to provide a BLV student detail sufficient for creating the mental representation of the diagram. To create accessible and usable STEM interactives requires both dynamic alt text for diagrams that change with user input as well as an interface that can be accessed by a screen reader user (SRU).

An accessibility system for STEM interactives

An accessibility system was developed for five STEM interactives: three for chemistry, one for physics, and one for early math. This paper provides details on one of the chemistry interactives, the Lewis Structure interactive drawing tool. The accessibility system has three parts:

- An alt text generation engine to compose standardized descriptions in real-time
- A menu-based keyboard-accessible control panel
- An accessible AI-driven learning assistant.

Each interactive has configuration data that captures components within the diagram and tags each component's identity, position, and other pertinent details. These configurations were matched to a systematic method for describing the components within the context of specific learning objectives.

To illustrate the difference in detail between the alt text automatically generated by current word-processing software (in this case, Microsoft Word) and this new systematic method for creating real-time alt text, the alt text for the carbonate ion (Figure 1) is given below:

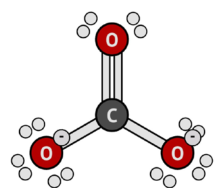

Figure 1: The diagram of a carbonate ion

Auto-generated alt text: *"A picture of a molecule."*
Alt text generated with the accessibility system: *"A Lewis structure of a molecule of carbonate with the chemical formula CO3 2-. The central atom is carbon. Carbon is single bonded to 2 oxygens, and double bonded to 1 oxygen. The single bonded oxygens have 3 lone pairs and a formal charge of -1 each. The double bonded oxygen has 2 lone pairs."*

Alt text descriptions for STEM diagrams are difficult to author because they provide communication of abstract ideas that are not easily explained with words. Effective alt text for STEM diagrams is currently constructed by specialized authors who have subject matter expertise, a knowledge of the learning context for the diagram, and who have been trained in composing alt text descriptions. Though required to meet Web Content Accessibility Guidelines in educational content (W3.org n.d.), there is no standardization for alt text authoring. Often SRUs don't know if the alt text they are accessing is complete and relevant, and poorly crafted alt text can be more detrimental than not having alt text at all (McCall and Chagnon 2022).

Typically, when a description is given, it is a long linear sequence of words with no standardized organization or method for navigation to the information of interest. Users interact differently with alt text descriptions. Some need more repetition; others need more context. We wanted to provide an interactive method for a SRU to personalize the building of the mental representation of the diagram.

To meet these individual needs, it was hypothesized that that we could use available Large Language Models (LLMs) to create an interactive learning assistant whereby the SRU could ask for specific details regarding the alt text descriptions. We tested this concept by integrating the Lewis Structure (chemical drawings) interactive with the OpenAI GPT API.

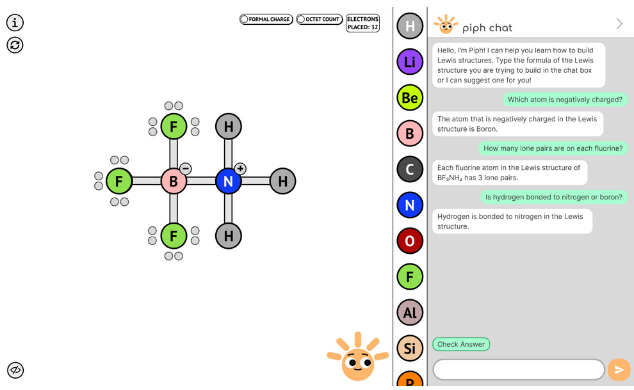

Figure 2: The Lewis Structure for BF_3NH_3 and an example of the alt text learning assistant dialog

Initial tests included a prompt to instruct the LLM that it was working with a BLV student studying Lewis structures and its job was to answer the student's questions based on the given description of the molecule, which was provided by the dynamically generated alt text. The LLM was then asked questions like "what is bonded to carbon?" or "what is the charge on nitrogen?" With the short alt text, the learning assistant showed promising results, but there were still errors. Providing more detailed descriptions and context allowed for more complex questions to be asked and answered accurately as shown in Figure 2. Finally, a knowledge base that included information like "the oxygen atom is a red circle" was incorporated into the LLM so that SRUs could also ask aesthetic questions. Through the learning assistant, the extra details can be accessed in a straightforward manner that reflects how a BLV user would interact with a sighted person when asking for assistance in understanding a diagram.

Usability Study with BLV adults

Four BLV adults who regularly use a screen reader and rely on alt text to access non-text information were recruited for a usability study of the accessibility system. This number of participants was deemed sufficient as prior research demonstrated it appropriate for detecting most usability problems (Neilson and Landauer 1993). Prior to recruitment, all procedures were reviewed and approved for ethical human subjects research by HML IRB (study #2146). The three-part study gathered feedback on the alt text descriptions, the use of the learning assistant, and the keyboard accessible control panel. With respect to the alt text descriptions and the control panel, all participants used the dynamically generated alt text to identify the correct tactile diagram from three choices and could readily navigate with the control panel to draw chemical structures.

Of the four participants, three used the LLM-driven learning assistant while reviewing the alt text descriptions. The fourth was satisfied with the alt text descriptions provided. One participant preferred to use the learning assistant over listening to the alt text numerous times and explained, "I liked it because then I could ask the questions without having to arrow through all the information… getting the stuff I wanted answered and then asking follow up questions to get the other things I wanted answered was easier for me."

Expansion beyond accessibility

The systematic, textual descriptions of the alt text generation system provided a unique bridge between the natural language of LLMs and the visual language of STEM diagrams. The next step in the project was to create a pedagogical rubric system for the Lewis Structure interactive, in which the learning objective (creating a structure from a given molecular formula) was divided into a structured set of requisite skills (such as counting valence electrons and adding bonds) to produce a defined learning pathway. This rubric was added to the knowledge base in the LLM interface and connected through the alt text descriptions with additional context.

With this expanded learning assistant, both sighted and BLV users, could either draw the Lewis structure to obtain feedback on where they were along the learning pathway or ask specific text-based questions about the structure they had drawn. The development of this diagram-based learning assistant has created a use case where all users benefit from the inclusion of alt text. This diagram-based input method for LLMs has potential to deliver inclusive and accessible personalized learning for STEM subjects from elementary through higher education, and into workforce training.

Future research includes expanding the system beyond chemistry and conducting more comprehensive feasibility studies in authentic learning environments.

Acknowledgments

The authors wish to thank Nicole Kada for her assistance in feedback on design from her perspective as a recent BLV STEM graduate, and Deb Nischik for her interview support in the usability study. This material is based upon work supported by the National Science Foundation under Grant No. 2221722. The corresponding author, Julia E. Winter, can be reached at julia@alchem.ie.

Conflict of Interest Statement

The authors have received compensation for work performed as employees of Alchemie Solutions, Inc, the producer of the Lewis Structure interactive learning tool.

References

Cooper, M.M., Stieff, M. and DeSutter, D. 2017. Sketching the Invisible to Predict the Visible: From Drawing to Modeling in Chemistry. *Topics in Cognitive Science*, 9: 902-920. https://doi.org/10.1111/tops.12285.

Lieberman, M. Inside Higher Ed, Helping Institutions Reach Accessibility Goals. 2019. https://www.insidehighered.com/digital-learning/article/2019/02/20/guide-accessibility-practices-aims-help-institutions-develop. Accessed: 2023-11-25.

Martin-Samer, M., Casada, C., Gomez-Pozeulo, G. 2023. Utilising Interactive Applications as Educational Tools in Higher Education: Perspectives from Teachers and Students, and an Analysis of Academic Outcomes. *Education for Chemical Engineers*. 46, 1-9. 10.1016/j.ece.2023.10.001.

McCall, K., Chagnon, B. 2022. Rethinking Alt Text to Improve Its Effectiveness. In: Miesenberger, K., Kouroupetroglou, G., Mavrou, K., Manduchi, R., Covarrubias Rodriguez, M., Penáz, P. (eds) Computers Helping People with Special Needs. ICCHP-AAATE 2022. Lecture Notes in Computer Science, vol 13342. Springer, Cham. https://doi.org/10.1007/978-3-031-08645-8_4.

Nielsen, J.; Landauer, T. K. 1993. A Mathematical Model of the Finding of Usability Problems. In Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems; CHI '93; Association for Computing Machinery: New York, NY, USA, 1993; pp 206–213.

W3.org, Web Content Accessibility Guidelines 2.0, W3C World Wide Web Consortium, (http://www.w3.org/TR/WCAG20/). Accessed: 2023-11-25.