Modeling Perceptional Fluency with Visual Representations

Blake Mason, Martina Rau, Lalit Jain, Robert Nowak BMASON3,MARAU,LJAIN,RDNOWAK@WISC.EDU
University of Wisconsin-Madison, Madison, WI 53706 USA

Abstract

Visual representations are ubiquitous instructional tools in science, technology, engineering, and math (STEM) domains. The goal of our ongoing research is to develop a new methodology for cognitive modeling of perceptual learning processes so as to create adaptive technologies that support perceptual fluency. We are using metric learning methods to assess which visual features novice students and experts focus on when presented with visual representations. Comparing novice to expert perceptions will establish which visual features perceptual support should help students attend to (e.g., because experts focus on them but novices do not). Hence, metric learning will provide a skill model of student perceptions (i.e., analogous to what verbalization techniques provide in traditional cognitive modeling). In this paper, we apply metric learning to identify salient features in the visual perception of molecular diagrams used in chemistry education.

1. Visual Representations in Chemistry

Instructors use the visual representations shown in Figure 1 to help students learn chemical bonding. Yet, to a novice student, these visual representations may not be helpful because the student may not know how to interpret the representations. First, they typically focus on one set of representational competencies: students' conceptual understanding of representations (e.g., the ability to explain how visual features depict concepts). This focus mimics education psychology research's focus on conceptual learning (Ainsworth, 2006; Seufert, 2003). However, research suggests a second type of representational competency is crucial for students' learning success: perceptual knowledge (Kellman & Massey, 2013; Massey et al., 2011), the ability to rapidly and effortlessly perceive information based on visual features of the representations. This ability results from implicit forms of learning. For example, expert chemists simply 'see' that the molecules depicted in Figure 1 have a local negative charge by the Oxygen atom, without having to make a conceptual inference. In contrast, novice students may wonder: does the red color in the ball-and-stick figure (Figure 1-b) mean the same thing as in the electrostatic potential map (EPM; Figure 1-d)? (It does not.) Instructors often ask students to use visual representations that they have never seen before to make sense of concepts that they have not yet learned about (Airey & Linder, 2009; Wertsch & Kazak, 2011), an issue known as the representation dilemma (Dreher & Kuntze, 2015). Hence, to succeed in STEM, students need representation skills that enable them to use visual representations to make sense of and solve domain-relevant problems (Ainsworth, 2006; Gilbert, 2005).

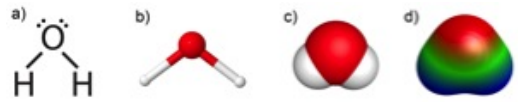

Figure 1. Representations of water. a: Lewis structure; b: ball-and-stick figure; c: space-filling model; d: electrostatic potential map (EPM).

Educational technologies are suitable to support representation skills because they can provide instructional support that adapts to individual needs. Adaptive capabilities of educational technologies such as intelligent tutoring systems (ITSs) rely on a cognitive model of student learning. Cognitive models infer whether the student has learned target skills through interactions with technology. Current cognitive models exist only for verbally accessible, explicit knowledge and cannot capture perceptual fluency resulting from implicit learning processes. Hence, ITSs cannot adequately adjust to students perceptual fluency, which limits capabilities of ITSs to adaptively sequence perceptual learning tasks and provide useful feedback on student interactions. The proposed research is to develop a methodology for cognitive modeling of perceptual learning processes so as to create adaptive support for perceptual learning tasks.

Proceedings of the 33^{rd} International Conference on Machine Learning, New York, NY, USA, 2016. JMLR: W&CP volume 48. Copyright 2016 by the author(s).

2. Metric Learning of Perceptual Similarities

For our experiment, we selected visual representations of chemical molecules are common in undergraduate chemistry instruction. Lewis structure representations are the most commonly used type of visual representations in undergraduate chemistry textbooks. To decide which molecules to include, we reviewed textbooks and online instructional materials and listed all occurring molecules. Further, we searched online instructional materials for the frequency of chemical names (e.g., H_2O) and common names (e.g., water) for these molecules. For our experiment, we chose the 50 most common molecules.

We created feature vectors that describe which visual features each representation contains (e.g., bond angles, the numbers of specific atoms, or the numbers of different atoms present). A total of 110 features were obtained for each molecule representation. The features for molecule i are denoted by the vector $\boldsymbol{x}_i \in \mathbb{R}^{110}$. The goal of our experiment is to identify which features chemistry students attend to while making judgments about similarities or differences between molecule representations.

To this end, we aim to learn a similarity function that describes student perceived similarities between molecule representations. We model the perceptual similarity between molecules i and j by

$$S_{ij} \;=\; \boldsymbol{x}_i^T \boldsymbol{A} \boldsymbol{x}_j \;,$$

where $\boldsymbol{A} \in \mathbb{R}^{110\times 110}$ matrix that parameterizes the model. The matrix $\boldsymbol{A}$ is learned by collecting perceptual judgments from chemistry students. Specifically, students are presented with a target molecule representation i and two other molecule representations, say j and k, and they are asked to select which of these two is most similar to the target representation. The metric learning problem is to find the $\boldsymbol{A}$ that minimizes the number of disagreements between the ranking predictions for each triple (i.e., either $S_{ij} > S_{ik}$ or vice-versa) and the comparative judgments collected from the students.

This learning problem can be posed as follows. Suppose that N comparative judgments are collected from students. For each judgment, $n = 1, \ldots, N$, let i_n denote the target in judgment n, j_n and k_n denote the two alternatives, and $y_n = \{-1, +1\}$ denote the response ($+1$ if j_n is selected and -1 otherwise). Then we solve the convex optimization

$$\widehat{\boldsymbol{A}} \;=\; \arg\min_{\boldsymbol{A}} \sum_{n=1}^{N} \Big(y_n - \boldsymbol{x}_{i_n}^T \boldsymbol{A} (\boldsymbol{x}_{j_n} - \boldsymbol{x}_{k_n}) \Big)^2 \;.$$

The relative importance of different features to the perceptual judgments can be quantified in terms of the norm of each row/column in $\widehat{\boldsymbol{A}}$. Alternatively, sparse regression techniques, such as the Lasso, can be employed to automatically select an predictive subset of the features.

Feature	Avg weight
Distinct letters	4.50%
Single bonds between Oxygen and Hydrogen	3.45%
180-degree angle in Hydrogen-Carbon-Fluorine	3.16%
Double bonds between Oxygen and Nitrogen	3.03%
Number of Nitrogen atoms	2.99%
Double bonds between Carbon and Oxygen	2.78%
120-degree angle in Hydrogen-Carbon-Hydrogen	2.73%
Number of Oxygen atoms	2.64%
180-degree angle in Carbon-Carbon-Oxygen	2.62%
Single bonds between Carbon and Oxygen	2.37%

Figure 2. Top ten features predictive of similarity judgments of novice chemistry students.

Another way to analyze and interpret the comparative judgment data is to create a metric embedding using non-metric multidimensional scaling techniques. The idea here is to create a two-dimensional image where each molecule representation is located such that distances between molecules in the image agree as well as possible with the judgment data. The results presented here are part of a larger study we conducted, detailed in (Rau et al., 2016).

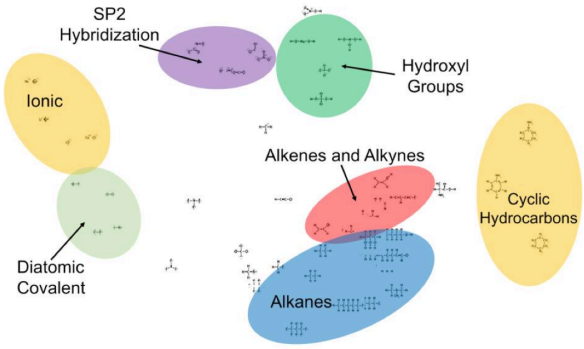

Figure 3. Two-dimensional similarity embedding of molecule representations. Distances between molecule representations correspond to student perceptions of dissimilarity (i.e., molecule representations that are depicted close to one another are perceived to be similar).

2.1. Collecting Perceptual Judgments from Students

Using NEXT, described further in Section 3, we collected perceptual judgments from students in an undergraduate chemistry course regarding the similarity between different Lewis structure representations of molecules in our corpus. We uploaded the images in our corpus to NEXT and used random sampling to generate the queries in this experiment.

Students completed similarity judgment tasks in the form of triplet comparisons (see Figure 4). Given a represen-

tation of a molecule (the"target-molecule"), students were asked to choose which of two other molecule representations (the "choice-molecules") was most similar to the given one. For each task, the student chose between one of the two choice-molecules that he/she perceived to be more similar to the target-molecule. After each task, another triplet was generated uniformly at random from our corpus of molecule representations.

In NEXT, students first received a brief description of the study and then worked through a sequence of 50 similarity judgment tasks. In total, we were able to collect 26,180 responses from 563 possibly non-unique participants through NEXT.

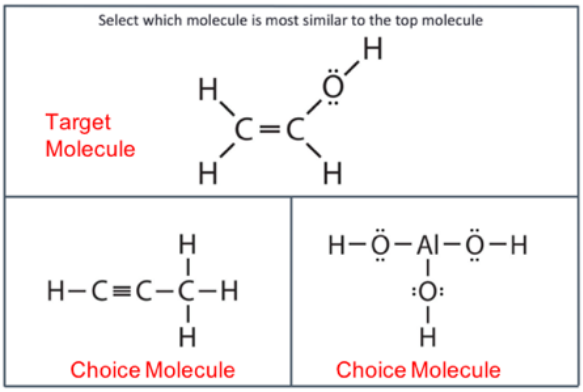

Figure 4. A sample query from NEXT of the type presented to students in the similarity judgment task. Students clicked on which of the two choice molecules they deemed most similar to the target.

3. NEXT: A Open-Source Software System for Machine Learning

As mentioned, the research study above was carried out using the cloud-based NEXT software system (`http://nextml.org`). The goal of NEXT is to provide experimenters with an accessible, user-friendly interface to run reproducible, large scale experiments, with a suite of potential algorithms for adaptive data collection in addition to being a platform for machine learning researchers to implement and test new algorithms.

Adaptive data collection can lead to learning an embedding with much fewer questions by asking the most informative triplet query at any given time. This can potentially provide huge gains over over random collection (Jamieson & Nowak, 2011; Tamuz et al., 2011). Crowdsourcing platforms such as Amazon's Mechanical Turk provide access to potentially thousands of users, and tapping into these resources combined with adaptive learning can lead to studies with thousands of responses being completed in a matter of hours.

NEXT is a cloud-based tool that allows researchers to utilize adaptive data collection for their own experiments. Note that NEXT is not just a tool for the triplet embedding problem. It can be used for ranking, classification, clustering and various other machine learning tasks that need to be distributed to a large group of participants and rely on adaptive methods. In addition to providing algorithms that ask adaptive questions and serving questions to participants, NEXT provides extensive dashboards on experiment performance, model specific data, and the ability to run validation studies. Most large scale experiments require a high overhead to reproduce, especially in the educationally psychology realm. NEXT alleviates this issue by providing a way to duplicate an adaptive data collection process simply by using the recycling experiment parameters.

Finally NEXT is general purpose and extensible. This enables researchers to implement their own experiment types and test various algorithms for data analysis. Note that this extensibility differentiates NEXT from other machine learning libraries such as MLib (Meng et al., 2015) and GraphLab (Low et al., 2012) that work on fixed data sets rather than providing a data collection tool. NEXT is built to be deployed on Amazon's EC2 service and experiments can be launched Currently NEXT is being used by our team, other researchers at the University of Wisconsin-Madison, and the Air Force Research Laboratory for adaptive data collection needs.

4. Initial Results

The initial results of the experiment with Lewis Structures are part of a larger effort of ours to understand and leverage student and expert perception (Rau et al., 2016). By characterizing the difference between student and expert perceptual judgments, we aim to build an ITS capable of providing adaptive support for perceptual learning tasks.

By solving the minimization shown in section 2 and using the resulting matrix $\widehat{A}$, we were able to predict the outcomes of held out triplets with 69% accuracy. Furthermore, from $\widehat{A}$ we can also compute the importance of each feature and rank them accordingly, as shown in figure 2 (Rau et al., 2016). These results show that the most highly ranked feature is the number of distinct letters. Specific visual features that are relevant to organic molecules were also ranked highly (e.g., the number of single bonds between Oxygen and Hydrogen atoms, the number of bonds between Carbon and Oxygen, the number of Nitrogen and Oxygen atoms). These features indicate the presence of chemical functional groups that are relevant to predicting molecule's reactive behavior. This ranking agrees visually with the embedding generated by NEXT, which has been labelled to show that clusters form around specific classes of hydrocarbon and organic molecules.

References

Ainsworth, Shaaron. Deft: A conceptual framework for considering learning with multiple representations. *Learning and instruction*, 16(3):183–198, 2006.

Airey, John and Linder, Cedric. A disciplinary discourse perspective on university science learning: Achieving fluency in a critical constellation of modes. *Journal of Research in Science Teaching*, 46(1):27–49, 2009.

Dreher, Anika and Kuntze, Sebastian. Teachers facing the dilemma of multiple representations being aid and obstacle for learning: evaluations of tasks and theme-specific noticing. *Journal für Mathematik-Didaktik*, 36(1):23–44, 2015.

Gilbert, John K. Visualization: A metacognitive skill in science and science education. In *Visualization in science education*, pp. 9–27. Springer, 2005.

Jamieson, Kevin G and Nowak, Robert D. Low-dimensional embedding using adaptively selected ordinal data. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pp. 1077–1084. IEEE, 2011.

Kellman, Philip J and Massey, Christine M. Perceptual learning, cognition, and expertise. *The psychology of learning and motivation*, 58:117–165, 2013.

Low, Yucheng, Bickson, Danny, Gonzalez, Joseph, Guestrin, Carlos, Kyrola, Aapo, and Hellerstein, Joseph M. Distributed graphlab: a framework for machine learning and data mining in the cloud. *Proceedings of the VLDB Endowment*, 5(8):716–727, 2012.

Massey, Kellman, P, Roth, Z, and Burke, T. Perceptual learning and adaptive learning technology. pp. 235–249, 2011.

Meng, Xiangrui, Bradley, Joseph, Yavuz, Burak, Sparks, Evan, Venkataraman, Shivaram, Liu, Davies, Freeman, Jeremy, Tsai, DB, Amde, Manish, Owen, Sean, et al. Mllib: Machine learning in apache spark. *arXiv preprint arXiv:1505.06807*, 2015.

Rau, M., Mason, B., and Nowak, R. D. How to model implicit knowledge? Similarity learning methods to assess perceptions of visual representations. In *Proceedings of the 9th International Conference on Educational Data Mining*, 2016.

Seufert, Tina. Supporting coherence formation in learning from multiple representations. *Learning and instruction*, 13(2):227–237, 2003.

Tamuz, Omer, Liu, Ce, Belongie, Serge, Shamir, Ohad, and Kalai, Adam Tauman. Adaptively learning the crowd kernel. *arXiv preprint arXiv:1105.1033*, 2011.

Wertsch, James V and Kazak, Sibel. Saying more than you know in instructional settings. In *Theories of learning and studies of instructional practice*, pp. 153–166. Springer, 2011.