SKETCHMINER: Mining Learner-Generated Science Drawings with Topological Abstraction

Andy Smith[1], Eric Wiebe[2], Bradford Mott[1], James Lester[1]

[1]Department of Computer Science [2]Department of STEM Education
North Carolina State University, Raleigh, NC, USA
{pmsmith4, wiebe, bwmott, lester}@ncsu.edu

ABSTRACT
Mining learner-generated sketches holds significant potential for acquiring deep insight into learners' mental models. Drawing has been shown to benefit both learning outcomes and engagement, and learners' sketches offer a rich source of diagnostic information. Unfortunately, interpreting learners' sketches—even sketches comprised of semantically grounded symbols—poses significant computational challenges. In this paper we describe SKETCHMINER, an educational sketch mining framework that automatically maps learners' symbolic sketches to topology-based abstract representations that are then analyzed with graph similarity metrics to perform automated assessment and misconception discovery. SKETCHMINER has been used to mine a corpus of symbolic science sketches created by upper elementary students in inquiry-based drawing episodes as they interact with an intelligent science notebook in the domain of physical science. Results of a study with SKETCHMINER suggest that it can correctly assess learners' symbolic sketches.

Keywords
Student modeling, Sketch analysis

1. INTRODUCTION
Diagrams and sketching are fundamental to science education. From primary through post-secondary education, students use drawings and graphical representations to make sense of complex systems and as a tool to organize and communicate their ideas to others. Studies have shown that learning strategies focusing on learner-generated sketches can produce effective learning outcomes, such as improving science text comprehension and student engagement [12], facilitating the writing process [11], and improving the acquisition of content knowledge [3]. Furthermore, spatial ability has been recognized as a predictor of STEM success even when accounting for mathematical and verbal ability [17].

Unlike the well studied areas of how people learn from writing text, viewing graphics, and reading, relatively little is known about how the generation of scientific drawings affects learning. Van Meter and Garner [9] posit that students asked to draw a picture engage in three cognitive processes: selecting relevant information, organizing the information to build up an internal verbal model, and constructing an internal nonverbal representation to connect with the verbal representation. Others suggest that drawing can be a meaningful learning activity requiring both essential and generative processing to mentally connect multiple knowledge representations [14].

The benefits of learner-generated sketching can best be realized by thoughtfully designing activities within a well-designed curriculum, as the positive effects of drawing strongly depend on the quality of the learner-generated products and scaffolding [10]. The act of generating a visual representation is a cognitively demanding task and, as such, requires scaffolds to guard against excessive and extraneous cognitive load [16]. Effective scaffolds for drawing include providing cutout figures, guiding questions, and targeted drawing prompts [7,19].

From a computational perspective, learner-generated drawings pose significant challenges. Even in an environment with predefined symbolic elements, the generative nature of the task yields a very large solution space of unique drawings and configurations. The work presented here describes initial efforts to mine learner-generated science drawings. To automatically cluster and compare drawings, the proposed framework uses a multi-step process of translating trace sketch behavior data of student drawings into topological representations. This process consists of converting the drawn elements into a graph representation based on a topology derived from the domain and using a modified edit distance methodology for comparing the topological graphs. We show how these comparisons can be used to analyze drawings to detect misconceptions, as well as to cluster student solutions in a manner that exhibits high fidelity with respect to human categorization.

This paper is structured as follows. Section 2 discusses other approaches that have been used to analyze student sketching. Section 3 describes the tablet-based learning environment that was used to collect the symbolic sketch dataset from elementary students. Section 4 introduces SKETCHMINER, a sketch data mining systems that automatically analyzes and compares student drawings using topological graphs. Finally, Section 5 describes an application of SKETCHMINER to cluster student drawings compared to a human clustering.

2. RELATED WORK

Sketch analysis poses significant computational challenges, with a majority of prior work focused on sketch recognition. For example, sketch recognition frameworks have been designed for domains such as organic chemistry and circuits in which free-hand drawing is translated into domain-specific symbols [1]. Another system, Mechanix, combines free-hand recognition capabilities with error checking to create feedback for undergraduate engineering students enrolled in a statics course [15].

Bollen and van Joolingen's SimSketch merges sketching with modeling and simulation of science phenomena [2]. In SimSketch, user free-hand drawings are segmented into objects by the system, and then annotated by the user with a variety of behaviors and attributes. Students can then run a simulation based on their drawing and see the results before revising their sketch. SimSketch has been evaluated in a planetarium setting and been shown to be both a functionally useable and enjoyable system for visitors.

Another promising line of investigation for studying learner-generated drawing in educational settings centers on the CogSketch system [5]. CogSketch has been developed as an open-domain sketch understanding system. Sketch worksheets were built within CogSketch, and used in a study to collect and cluster undergraduate geology student sketches by an analogical generalization engine [4].

3. LEONARDO CYBERPADS

Recent years have witnessed growing interest in introducing science notebooks into elementary science classrooms [13]. Science notebooks capture students' inquiry-based activities in both written and graphical form, potentially providing a valuable source of both diagnostic and prognostic information. However, because elementary teachers have limited training in science pedagogy, they often struggle with effectively using science notebooks in classroom learning activities [18].

For the past three years our laboratory has been developing a digital science notebook, the LEONARDO CyberPad (Figure 1), which runs on tablet computing platforms. LEONARDO integrates intelligent tutoring systems technologies into a digital science notebook that enables students to graphically model science phenomena. With a focus on the physical and earth sciences, the LEONARDO PadMate, a pedagogical agent, supports students' learning with real-time problem-solving advice. LEONARDO's curriculum is based on that of the Full Option Science System [8] and is aligned with the Next Generation Science Standard goals in elementary school science education [20].

Throughout the inquiry process, students using the LEONARDO CyberPad are invited to create symbolic sketches, including electrical circuits. Given the challenges of machine analysis of freehand sketching, as well as concerns of excessive cognitive demand for elementary students working in such an unstructured space [18], LEONARDO supports symbolic drawing tasks. To preserve the generative processing hypothesized to be of great benefit for learner-generated drawings strategies, each activity begins with a blank page so that the representations must be created from scratch. Students then choose from a variety of semantically grounded objects and place them at various points in the drawing space. For example, objects for the electricity unit include light bulbs, motors, switches, and batteries. Students then place wires on the drawing space, connecting the various objects to simulate proper electrical behavior. This focuses the learning activity on choosing the appropriate circuit elements and creating the appropriate circuit topology. Drawing tasks vary in complexity from copying a picture of a circuit held up by the PadMate, to recreating a circuit made during a physical investigation, to creating more complex circuits designed to increase their understanding of series and parallel circuits.

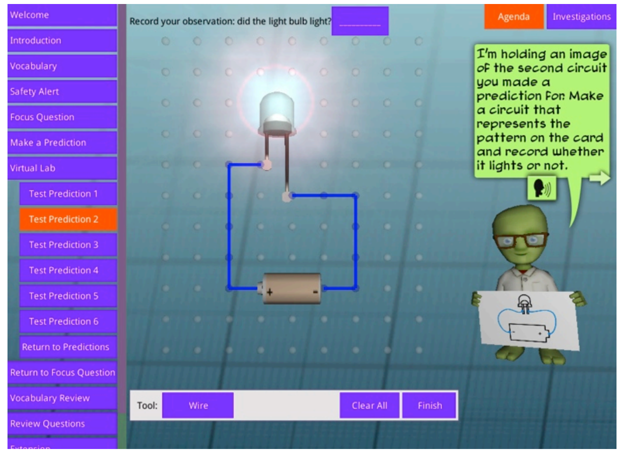

Figure 1. Screenshot of the LEONARDO CyberPad

4. TOPOLOGY-BASED SKETCH MINING

To analyze student drawings, SKETCHMINER first translates them into a more abstract representation. It takes as input trace logs from students' work in the CyberPad. From the trace logs it extracts student actions at a level of granularity capable of producing replay-quality representations of the drawing activities. From these actions it extracts the state of the student drawing at each point in the activity. For the analyses reported in this paper, we focus only on the final submitted sketches rather than the multiple drawings generated during the sketching process. The set of objects and locations are then utilized by a simulation engine that supports the querying of topological features of the drawing.

SKETCHMINER uses these topological features to generate a labeled graph representation of the drawing. Topological graphs provide two key representational benefits. First, they are very flexible and can be used across many domains. For the domain of circuits, our representation focuses on the electrical topology of the circuit drawing, which could be replaced or augmented by other features such as two-dimensional spatial topology. Second, graphs are easily visualized and interpretable by humans, which facilitates the interpretation of patterns and features extracted by automated analysis.

The first step in the translation from drawings to topological graphs is encoding the non-wire circuit elements. Circuit elements

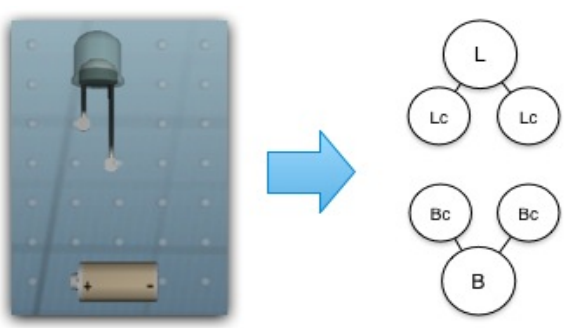

Figure 2. Circuit elements and corresponding topology

are represented as nodes in the graph. Because there are only two points where each node can interact with other objects in the drawing space, each node is connected to two nodes representing its contact points (Figure 2).

After creating the nodes of the graph, SKETCHMINER then generates the edges between them. For each contact point in the graph, the simulation engine uses a depth first search to return all other contact points reachable with a zero resistance path. If one or more paths exist between contact points, they are then connected with a single edge in the graph (Figure 3).

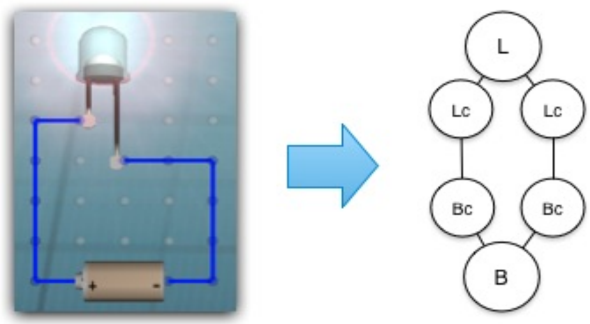

Figure 3. Connections encoded as edges

While multiple methods can be used to compare the similarity of graphs and trees, SKETCHMINER uses a method capable of numerically summarizing the difference between topographical states that also provides a description of how to transition from one state to the other. In particular, it uses a modified form of *edit distance*. Edit distance has been used to characterize errors in a variety of domains and is perhaps best known for its application in natural language spelling correction. Edit distance captures the difference between two representations as a series of edit operations. Additionally, these edit operations can be weighted, with the sum of necessary operations equaling the edit distance.

SKETCHMINER uses edit distance to measure the number of element additions, element deletions, edge additions and edge deletions needed to match two topologies. While traditional string edit distances tend to also utilize substitution, we chose to treat this instead as deleting an element, then adding a new one because this is the path a student would have to take to modify his or her drawing.

To determine the sequence of edit operations necessary to match two topologies, SKETCHMINER utilizes a guided search of possible actions to determine the lowest cost path through the operation space. While there are more efficient algorithms for graph edit distance, (e.g., see [6] for a survey), the greater complexity of these is not justified for the size of topological graphs generated from student sketches in this work.

Another design decision for SKETCHMINER considered how to weigh different edit operations for calculating the edit distance. An unweighted edit distance produces some undesirable effects. In particular, an unweighted score does not differentiate well between different types of errors. Consider a target drawing of a complete circuit featuring a battery and a motor. A blank submission and a complete circuit with the motor contacts short-circuited will both produce the same edit distance.

One approach to correcting for this is to adjust the weighting of actions. A subset of the student answers was analyzed with subject matter experts in an attempt to determine how the edit distance was aligning with curricular goals and assessment of different types of errors. A weighting scheme was generated to penalize missing elements at a cost of 4, extra elements at a cost of 2, and extra/missing edges at a cost of 1. SKETCHMINER uses this weighting scheme.

5. CORPUS ANALYSIS

For the analyses of SKETCHMINER reported here, a corpus of fourth grade symbolic drawings was collected with the LEONARDO CyberPads running on iPads in elementary classrooms in North Carolina and California. After data cleaning, drawing activities from 132 students were used for the analysis. Student drawings were scored in comparison to normative models constructed by the research team. Because there may be multiple correct solutions to a given exercise, student submissions were scored against multiple "correct" solutions and assigned the score of the closest match. These scores were then used to qualitatively analyze the student drawings as a basis for the distance metric for unsupervised clustering and for misconception detection.

To evaluate SKETCHMINER's edit distance's value as an assessment metric, we clustered student drawings using both the weighted and unweighted topographical edit distance as the distance metric. In order to evaluate the clusters, two independent coders from the project's education team developed a rubric (described in Table 2) and scored the student responses for a circuit involving a switch, motor, and battery connected in series. Based on the rubric, the drawings were independently classified into 4 clusters by the two coders ($\kappa = .9$), creating a gold standard clustering to validate our clusters against.

After the hand coding, we then ran an automated cluster analysis on the student drawings based on the SKETCHMINER generated codings. To cluster the drawings we utilized the WEKA toolkit implementation of k-means clustering with k=4 to align with the human coding. Because k-means can be dependent on initialization, the analysis was run 10 times with different random seeds and the results averaged.

Table 1. Classification accuracy

Distance Metric	Accuracy	Precision	Recall
Unweighted	.73	.56	.63
Weighted	.86	.74	.76

As shown in Table 1 above, SKETCHMINER produced strong alignment with the human classifications, with the weighted edit distance producing better results than unweighted. The improved accuracy is a result of the weighted edit distance outperforming the unweighted edit distance at separating the three error classes.

Table 2. Classification by class for weighted edit distance

Class	Accuracy	Precision	Recall
1 (Blank)	.89	.61	1
2 (No Structure)	.87	.66	.5
3 (Some Structure)	.86	.92	.6
4 (Correct)	.98	1	.96

Further analysis of the weighted edit distance classification reveals that the process produced near-perfect accuracy on correct answers (Class 4). Inspection of the misclassified correct student sketches showed one example where the student had created the correct circuit, and a smaller unrelated circuit on a different part of the drawing space which inflated its edit distance. The other human-coded correct answer misclassified by SKETCHMINER was due to the student creating the correct topology but using a light bulb instead of a motor.

For classifying errors, the clustering showed strong alignment with empty entries, but had difficulty separating Class 2 errors (elements present but with no structure) from empty submissions. One possible way of improving this in the future could be to treat absence-of-circuit elements as a special case error.

6. CONCLUSION

Understanding how students learn from drawing is a foundational problem in learning analytics. Tablet-based science notebooks, such as the one provided by the LEONARDO CyberPad, offer an excellent "laboratory" for instrumenting the drawing process and afford significant opportunity for educational data mining techniques. In this paper we have introduced SKETCHMINER, which utilizes a graph-based representation of drawing topologies to automatically interpret learner-generated symbolic sketches. In an analysis of SKETCHMINER's application to a corpus of fourth grade student symbolic sketches, it was found that its assessment of student drawings aligns with human-provided assessments.

The results show promise as a means of automatically assessing learner drawings and suggest several lines of investigation for future research. First, while "distance to solution" is a valuable metric, SKETCHMINER's edit distance could also be used to compare errors to each other. Preliminary analysis using this technique has shown promise for identifying common error states that could be used in curriculum redesign or to generate targeted scaffolding for students.

Another area for future research is applying SKETCHMINER to more topologically complex domains. Because the topographical relations in the domain of circuits are somewhat sparse, SKETCHMINER's representations would need to be evaluated on more complex student drawings containing more diverse sets of elements and relationships with more complex topologies.

Perhaps the most promising area for analysis is investigating the drawing process itself. Topographical representations can be created at any point in the drawing process, allowing for analysis of sequences and patterns in student drawing. Models learned from corpora of learner drawing processes can be used to create more accurate models of learners' conceptual representations, as well as the basis for providing customized scaffolding to support a broad range of learner populations.

7. ACKNOWLEDGMENTS

This research is based upon work supported by the National Science Foundation under Grant No. DRL-1020229. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

8. REFERENCES

[1] Alvarado, C. and Davis, R. SketchREAD: a multi-domain sketch recognition engine. *Proceedings of the 17th annual ACM symposium on User interface software and technology*, ACM (2004), 23–32.

[2] Bollen, L. and Joolingen, W. van. SimSketch: Multi-Agent Simulations Based on Learner-Created Sketches for Early Science Education. *IEEE Transactions on Learning Technologies 6*, 3 (2013), 208–216.

[3] Britton, L. and Wandersee, J. Cutting up Text to Make Moveable, Magnetic Diagrams: A Way of Teaching & Assessing Biological Processes. *The American Biology Teacher 59*, (1997), 288–291.

[4] Chang, M. and Forbus, K. Clustering Hand-Drawn Sketches via Analogical Generalization. *Proceedings of the 25th Annual Conference on Innovative Applications of Artificial Intelligence*, (2013).

[5] Forbus, K., Usher, J., Lovett, A., Lockwood, K., and Wetzel, J. CogSketch: Sketch Understanding for Cognitive Science Research and for Education. *Topics in Cognitive Science 3*, 4 (2011), 648–666.

[6] Gao, X., Xiao, B., Tao, D., and Li, X. A survey of graph edit distance. *Pattern Analysis and Applications 13*, 1 (2009), 113–129.

[7] Lesgold, A.M., Levin, J.R., Shirmon, J., and Guttman, J. Pictures and young children's learning from oral prose. *Journal of Education Psychology*, 67 (1975), 636–642.

[8] Mangrubang, F. Preparing elementary education majors to teach science using an inquiry-based approach: The Full Option Science System. *American Annals of the Deaf 149*, 3 (2004), 290–303.

[9] Van Meter, P. and Garner, J. The Promise and Practice of Learner-Generated Drawing: Literature Review and Synthesis. *Educational Psychology Review 17*, 4 (2005), 285–325.

[10] Van Meter, P. Drawing construction as a strategy for learning from text. *Journal of Educational Psychology 93*, 1 (2001), 129–140.

[11] Moore, B. and Caldwell, H. Drama and drawing for narrative writing in primary grades. *The Journal of Educational Research 87*, (1993), 100–110.

[12] Rich, R.Z. and Blake, S. Using pictures to assist in comprehension and recall. *Intervention in School and Clinic 29*, 5 (1994), 271–275.

[13] Ruiz-Primo, M. a., Li, M., Ayala, C., and Shavelson, R.J. Evaluating students' science notebooks as an assessment tool. *International Journal of Science Education 26*, 12 (2004), 1477–1506.

[14] Schwamborn, A., Mayer, R.E., Thillmann, H., Leopold, C., and Leutner, D. Drawing as a generative activity and drawing as a prognostic activity. *Journal of Educational Psychology 102*, 4 (2010), 872–879.

[15] Valentine, S., Vides, F., Lucchese, G., et al. Mechanix: A Sketch-Based Tutoring System for Statics Courses. *Proceedings of the Twenty-Fourth Conference on Innovative Applications of Artificial Intelligence*, (2012).

[16] Verhoeven, L., Schnotz, W., and Paas, F. Cognitive load in interactive knowledge construction. *Learning and Instruction 19*, 5 (2009), 369–375.

[17] Wai, J., Lubinski, D., and Benbow, C.P. Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology 101*, 4 (2009), 817–835.

[18] Wiebe, E.N., Bedward, J.C., and Madden, L.P. Graphic Representations in Science Notebooks : A Vehicle for Understanding Science Inquiry in the Elementary Classroom. *Presented at AERA*, (2009).

[19] Zhang, H. and Linn, M. Using drawings to support learning from dynamic visualizations. *Proceedings of the 8th International Conference of the Learning Sciences*, (2008), 161–162.

[20] Achieve. In *Next Generation Science Standards - First Public Draft*. Achieve, Inc, 2012.