

# Sketch Worksheets in Science, Technology, Engineering, and Mathematics Classrooms: Two Deployments

*Kenneth D. Forbus, Bridget Garnier,  
Basil Tikoff, Wayne Marko, Madeline Usher, Matthew McLure*

■ *Sketching is a valuable but underutilized tool for science education. Sketch worksheets were developed to help change this, by using artificial intelligence technology to give students immediate feedback and to give instructors assistance in grading. Sketch worksheets use automatically computed visual representations combined with conceptual information to give feedback to students, by computing analogies between students' sketches and an instructor's solution sketch. This enables domain experts to develop sketch worksheets, to facilitate dissemination. We describe our experiences in deploying them in geoscience and artificial intelligence classes. The geoscience worksheets, authored by geoscientists at University of Wisconsin–Madison, were used at both Wisconsin and Northwestern University. The artificial intelligence worksheets were developed and used at Northwestern. Our experience indicates that sketch worksheets can provide helpful on-the-spot feedback to students, and significantly improve grading efficiency, to the point where sketching assignments can be more practical to use broadly in science, technology, engineering, and mathematics education.*

Sketching is a natural way for people to think through spatial ideas, and to communicate about these ideas with others. This makes it attractive for science, technology, engineering, and mathematics (STEM) education, and there is indeed evidence that it can both help students learn and be used to help assess student knowledge (for example, Ainsworth et al. 2011; Jee et al. 2014). In geoscience, for example, sketching is heavily used by practitioners. Paradoxically, a survey of geoscience instructors indicates that, while most of them believe that sketching is useful for students as well, they use sketching assignments less than they desire (Garnier et al. 2017). One reason is that grading sketches is very time-consuming, compared with short-answer question forms. Providing support for semiautomatic grading of sketches can potentially increase their use in STEM education. The same mechanisms, if used correctly, can also provide on-the-spot feedback to students, anytime and anywhere, similar to the functionality provided by cognitive tutors (Koedinger et al. 1997) for nonspatial subjects.

Sketch worksheets (Yin et al. 2010; Forbus et al. 2017) are our approach to providing these capabilities for education. As we will describe, Sketch Worksheets use ideas from cognitive science and artificial intelligence (AI) to visually analyze student sketches, and provide advice based on comparisons with instructor sketches via analogy. Importantly, domain experts and instructors author the sketch worksheets, as opposed to AI experts or software developers needing to be involved. They have now been developed and deployed in two very different kinds of classes, including the use of sketch

worksheets developed by one university in classes at another. This article summarizes these experiences in deployment. We start by outlining the underlying technology, including how sketch worksheets are authored, to set the stage. Next we describe their deployment in geoscience, including the process of developing them for an introductory geoscience class, and lessons learned from using them in classes. Then we describe how they have been used in a knowledge representation and reasoning course, a topic far from the disciplines that motivated their original development. We close with conclusions and future work.

## Sketch Worksheets: The Basics

Sketch worksheets are built on CogSketch, an open-domain sketch understanding system (Forbus et al. 2011). This section provides just enough information about CogSketch to understand the rest of this article. The reader who wants more details should consult Forbus et al. (2017).

Most AI efforts on sketch understanding focus on recognition, and this has led to successful deployed educational software systems for structural mechanics (Lee et al. 2008; Valentine et al. 2012), aspects of analog electronics (de Silva et al. 2007), and chemistry (Cooper et al. 2009). Our open-domain approach is fundamentally different. Recognition makes sense in domains like handwriting recognition or circuit diagrams, where the mapping from abstract entities to shapes is one-to-one, and the particular spatial properties of the layout of the visual symbols constituting a sketch are irrelevant. By contrast, for most STEM domains, the mapping between abstract entities and shapes is many-to-many, and the specific spatial properties of what is drawn are often crucial to understanding. For example, three concentric circles might depict the layers of the earth, planetary orbits, or the cross section of a heat exchanger. The context of a specific exercise might enable discarding most of these interpretations, but even within an interpretation, an educational system should not presume that the student gets it right: The student might have the order of the layers mixed up, for instance. Hence it is important that students label what they draw, for the instructor (and the software) to accurately assess their knowledge. Thus, CogSketch requires students to draw visual objects, called *glyphs*, and label them with their intended meaning. The labels are drawn from an underlying knowledge base,<sup>1</sup> with textual renderings chosen by the instructor so that they can customize it.

The visual language of CogSketch provides three kinds of glyphs. *Entity glyphs* depict specific objects, concrete or abstract (for example, an orbit). *Relation glyphs* depict binary relationships between the entities depicted by glyphs (for example, *owns* in a knowledge graph). *Annotation glyphs* provide a way of specifying nonvisual properties in the sketch (for example, the temperature of an object). These glyphs

are compositional; that is, relation glyphs can apply to describing relations between and among dependency structures, and annotation glyphs can apply to relation glyphs (for example, to indicate rate of flow between carbon reservoirs).

CogSketch automatically computes a variety of visual and spatial relationships. Visual relationships are in the plane of the image, whereas spatial relationships are with respect to the viewpoint of the sketch (for example, *side view*, *top-down*, or *map*). It automatically computes positional relations (for example, *above*, *leftOf*), qualitative topological relationships (RCC8; Cohn et al. 1997), axis information, relative sizes, whether a shape is open or closed, and for arrows, what the head and tail are near. CogSketch is capable of computing a wider variety of relationships as well, on demand. Moreover, it is also capable of decomposing ink into edges and junctions, and combining glyphs into groups via gestalt principles, such as similarity and connectivity. All of these visual representations were motivated by psychological data concerning human vision where available. The goal is to ensure that CogSketch sees a sketch in ways similar to the ways that people do. This has led to CogSketch models of multiple visual problem-solving tasks, including mental rotation and paper folding (Lovett and Forbus 2013), geometric analogies (Lovett and Forbus 2012), an oddity task (Lovett and Forbus 2011), and Ravens' Progressive Matrices (Lovett and Forbus 2017). In all of these models, the simulations are consistent with human data; for example, what is hard for CogSketch is hard for people, and vice versa. Our argument is that human-like sketch understanding must include the perception of these visual relationships.

Sketch worksheets are based on the insight that many important domain ideas can be mapped onto visual and spatial relationships. A student asked to draw the layers of the Earth, for example, should include the crust, mantle, outer core, and inner core, in exactly that order, going inward. Other exercises provide photographs, maps, or diagrams that need to be annotated to illustrate important ideas, such as where faults and marker beds are in a photo of geological strata. Being able to automatically analyze a student's sketch, and compare it to an instructor's solution, provides the grist for computing how they are similar and what the differences are. As long as the important domain concepts in an exercise can be tied to visual relationships, this sketch-and-compare model can be used to give constructive feedback. In CogSketch, a sketch consists of subsketches, each providing related information about the subject of the sketch. In sketch worksheets, there is one or more solution subsketch, where the instructor depicts a correct solution to the problems posed by that worksheet. (Instructor-authored sketches depicting common misconceptions are also supported, but were not used in the deployments described here.) Another subsketch is used by students in doing their work,

and unless the worksheet is unlocked, all solution subsketches are hidden from the student.

Authoring sketch worksheets is designed to be done via domain experts, to support dissemination. This means that they must not rely on domain-specific reasoning, nor can they go beyond the normal visual capabilities of CogSketch. The overall workflow for sketch worksheets is illustrated in figure 1. The author, a domain expert or instructor, sets out the problem to be solved by the student and draws a sketch that represents a good solution to it. CogSketch analyzes their sketch, and the instructor marks a subset of the facts it generates as important, assigning points for getting them correct, and feedback to be provided otherwise. When students tackle the worksheet, they draw their sketch, which CogSketch analyzes and compares to the teacher's sketch via a computational model of human analogy, the structure-mapping engine (SME), as will be explained. The differences SME finds are used to help provide feedback to the student, so they can improve their sketch. In addition to getting the final product, detailed data about the process of sketching is also provided to the instructor for assessment purposes. CogSketch includes a variety of ways to visualize such data, and export tools for doing further analysis with other software.

The authoring process starts by instructors selecting what concepts need to be used (by browsing the knowledge base) and drawing their solution sketch. Selecting concepts involves spelunking through the knowledge base, using string completion and a hyper-text browser. The concepts selected by the instructor include everything they need to conceptually label their solution sketch, plus appropriate distractors. Types for entities are drawn from the collections (that is, what the Cyc ontology (Matuszek et al. 2006) uses to represent concepts) in the NextKB knowledge base,<sup>2</sup> as are the vocabulary of binary relationships and annotations. There is a definite learning curve to knowledge-base spelunking,<sup>3</sup> but authors are also free to add their own — as long as they use the same collections and relations consistently, the analogical matching process will work as intended. Importantly, instructors can provide strings for the names of types and relations, as well as documentation, which controls what students will see (figure 2). This enables them to author materials that are domain- and age-appropriate, as well as using languages other than English.

As the instructor draws their sketch during the authoring process, CogSketch incrementally and automatically performs a visual analysis of the sketch. If instructors desire more types of visual relationships to be computed, they can be specified via a menu. English-language templates in the knowledge base provide readable text for each relationship, and instructors can select a fact to see the entities it relates highlighted in the sketch, to help make sure that they have selected the right relationship. Particular facts can be marked as important facts for tutoring. For example, in a worksheet on the layers of the

Earth, the mantle must be inside the crust. Tutoring facts can have an associated piece of feedback to be provided when that fact isn't true in a student's sketch, and a point value that is used in grading. The feedback is provided via a text string, which enables instructors to craft a message that ought to be helpful. Empowering authors to write the feedback, like concept and relation names, enables them to make the message domain- and age-specific, and use whatever language is appropriate for their student population.

An important kind of tutoring fact that CogSketch provides is quantitative ink constraint. In quantitative ink constraints, the student's sketch is supposed to accurately align with the instructor's sketch for that glyph, as when annotating a photograph or diagram. The instructor specifies an error tolerance, and when comparing the student and instructor sketches, CogSketch will use a numerical comparison between the instructor's ink and the student's ink, providing feedback based on how they mismatch.

Instructors also provide text posing the problem(s) to the student, and optionally provide a background image as part of the problem (for example, a photograph or diagram to be annotated). They can also provide multiple-choice questions to be asked of students before and after they complete a worksheet, as additional assessments and opportunities for reflection. CogSketch also includes a gradebook, which does batch processing of a directory of sketches (for example, as downloaded from a course management system when assignments are turned in). It uses a web interface to enable instructors to browse each student's work, including its full history, and view the automatically assigned points, based on the instructor-provided rubrics.

Feedback to students is generated on demand, to avoid interrupting them while thinking. When feedback is requested, a multistep process ensues: (1) CogSketch tests to see if any glyphs do not yet have a conceptual label. Without conceptual labels, it is not clear what should match to what, and so it will point those out, and not continue until they are fixed. (2) Glyphs that are mentioned in facts that the instructor marked as important are considered to be *required glyphs*. If any required glyphs are missing, the tutor mentions which glyphs are missing, so that the student can add them. If all glyphs have conceptual labels and all required glyphs are present, then CogSketch prepares to compare the instructor and student sketches. (3) It runs any optional visual computations needed to derive tutoring facts, if there are any. (4) It compares the instructor and student sketches using the SME, a cognitive model of analogical matching (Forbus et al. 2016). Mismatches are detected by analyzing SME's candidate inferences, and any associated advice is retrieved and presented to the student, hyperlinked with the glyphs involved to help them make sense of it.

Two sources of additional guidance provided by the instructor are also used during tutoring. For

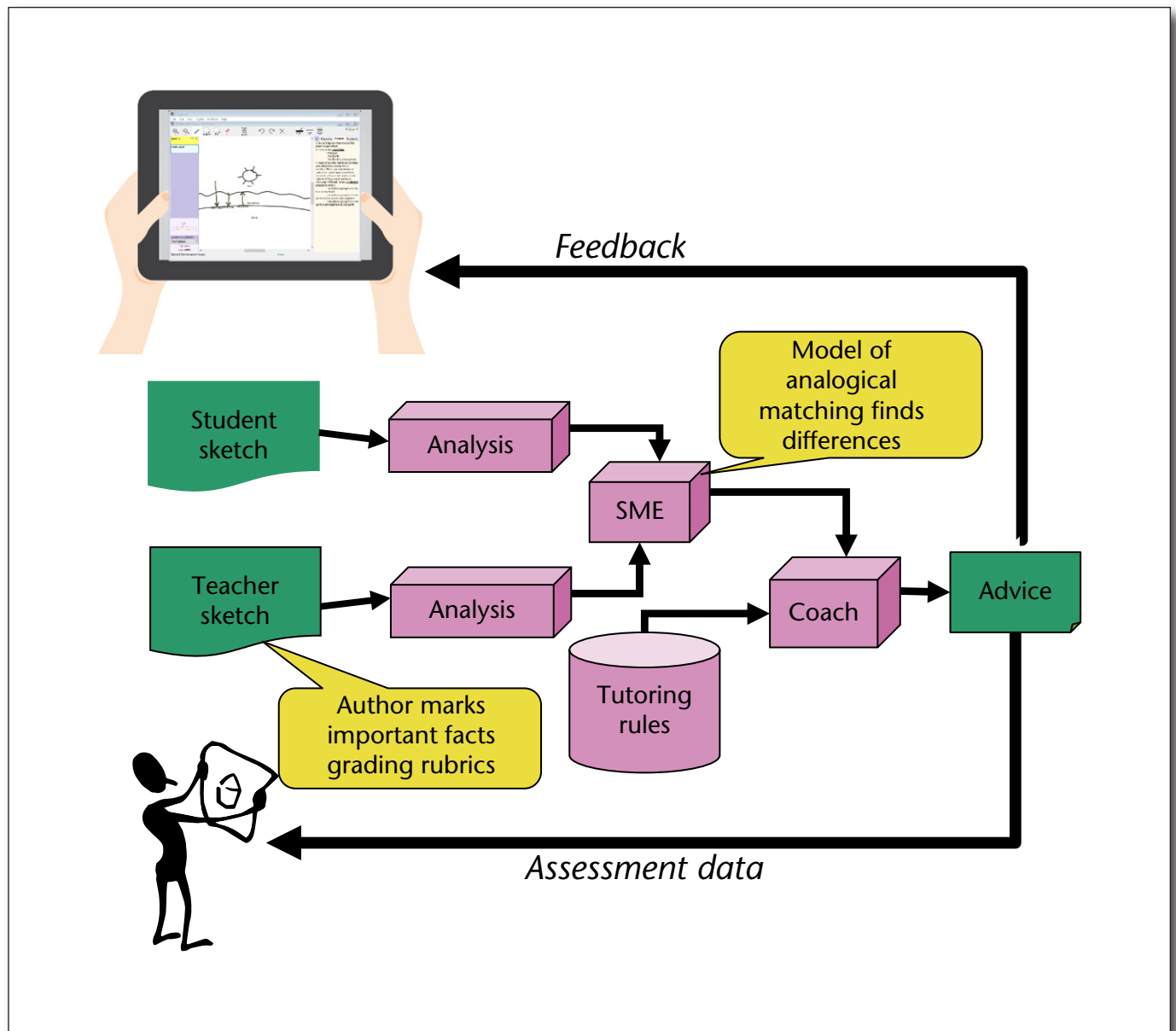


Figure 1. The Sketch Worksheet Authoring Process.

the first, recall that annotation glyphs can supply additional information about existing glyphs; for example, that the entity represented by another glyph is 2.1-m long. Candidate inferences about quantitative annotation values receive special handling. The instructor can provide a range of allowable values, as well as different advice depending on whether the student's value is too low or too high. The second additional guidance is associated with quantitative ink constraints. Instructors specify a tolerance region around the outline of the solution glyph's location that is used to determine what will be acceptable. Moreover, different advice can be given depending on whether the student's glyph is above, below, left of, or right of the correct location.

## Sketch Worksheets in Geoscience

Geoscience is one of the most spatially intense STEM disciplines, hence it is a natural discipline for sketching. Here we discuss how sketch worksheets were developed for an introductory course at the University of Wisconsin-Madison, and deployed in an introductory course at Northwestern University.

### Developing Sketch Worksheets for Geoscience

Geoscience sketch worksheets were developed through a collaboration among geoscientists, cognitive scientists, and CogSketch developers. Collaboration

Collections Allowed in Sketch:	All Collections:
Fault-FootWall (foot wall)	Fault
<b>Fault-HangingWall (hanging wall)</b>	Fault-ActiveRegion
Fault-Topographical (geological fault)	Fault-FootWall (foot wall)
StoneObject-Natural (stone)	Fault-Geological (fault)
	Fault-HangingWall (hanging wall)
	Fault-Topographical (geological fault)
	Fault-Topographical_DipSlip (dip-slip f

←  
X

This side of a fault is above the plane of the fault-line.

**Human-readable name for the term Fault-HangingWall:**

hanging wall

Figure 2. Authors Select Concepts to Use From the Knowledge Base, and Can Edit How They Are Presented to Students.

was necessary to ensure that each worksheet: (1) used results of cognitive-science based research to support the learning of spatially complex topics through sketching; (2) properly used the spatial, feedback, and grading capabilities of CogSketch, which results in an activity that is beyond a typical paper sketching worksheet; and (3) required primary authorship by a domain expert to ensure that the worksheet was rigorously correct and relevant to geoscience instruction.

Twenty-six sketch worksheets were developed for introductory geoscience courses (typically, physical geology). In developing sketch worksheet content, geoscientists identified specific concepts commonly taught in introductory geoscience courses that are difficult for students to grasp. In addition, cognitive scientists helped identify four spatial skills that are necessary for success in the geosciences: disembedding, reasoning about dynamic processes (figure 3), penetrative thinking (figure 4), and scaling (see Garnier et al. 2017). Geoscience concepts and spatial skills are important to develop at the introductory level to aid student understanding and possible success in future STEM courses.

Because most geoscience concepts involve the use of at least one spatial thinking skill, it was important to incorporate cognitive science-based research into sketching activities to support simultaneous learning of geoscience concepts and spatial

skills development. Once a geoscience concept was chosen, sketch activities were created based on activities that would support development of the spatial skill, often taken from techniques or activities in educational research of geoscience, or other STEM disciplines.

Sketch activities were also tailored to take advantage of CogSketch's interactive and spatial capabilities. Each worksheet incorporated tasks that allow students to move and rotate objects, draw or annotate on top of photos and diagrams, free sketch, and draw arrows to show motion or relationships. Worksheets included a background image and/or objects to manipulate, directions to complete various spatial tasks, and multiple choice questions to answer when the worksheet is complete. All worksheets are currently accessible on the Science Education Resource Center at Carleton College website.<sup>4</sup>

Creating the solution sketch and feedback was an important part of worksheet development. The goal of solution sketches and feedback was to lead students to the correct answer through a trial-and-error process. Preliminary testing of each worksheet with student volunteers showed that poor solutions/feedback leads to student frustration if correct answers were identified as incorrect, if feedback was not helpful, and/or if students could not fix an incorrect sketch. These errors may initially be seen as errors



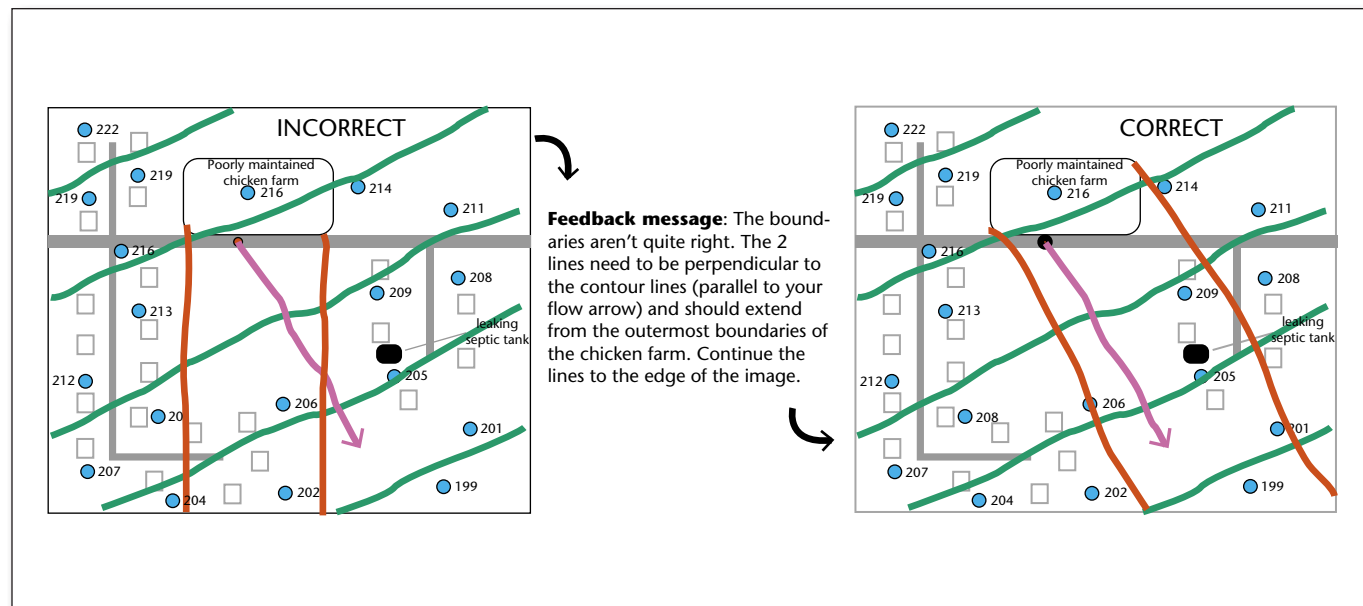


Figure 3. Example of Student Work on a Groundwater Contamination Worksheet.

Contour lines (green) and flow paths (purple and red) are difficult concepts for students, so drawing them and the flow paths are useful. On the left is a common student mistake; in the middle is the feedback from the worksheet; and on the right is the student's corrected sketch.

with the SME or tutor, where, in fact, it is a reflection of errors in the authoring process. For example, if a student correctly traces a fault but part of the line goes outside of the ink tolerance set by the worksheet author, the student will receive feedback for an incorrect sketch. To correct this situation, a worksheet author would increase the ink tolerance to include acceptable human error when drawing a line. Worksheet authors must thoroughly test and continually update solution sketches as a wider range of errors are observed. This process can be time-consuming but results in improved solution/feedback advice for each worksheet.

All worksheets were authored in CogSketch by a geoscientist, using the authoring environment. The first few worksheets each took 1–1.5 weeks to fully develop, from idea to completed worksheet. Development time greatly decreased with each worksheet, to the point where it took ~3 days per worksheet. About three-quarters of development time was spent creating the material (that is, worksheet idea, images, text) and the remainder was spent authoring the worksheet in CogSketch, as well as testing the worksheet and correcting problems.

Initial testing of 16 developed sketch worksheets was conducted in an introductory geoscience course at the University of Wisconsin-Madison, spring 2014. By analyzing completed student worksheets, we learned that: (1) the worksheet tutor allowed students to use different strategies to complete worksheets; (2) the tutor identified and helped correct common mistakes; and (3) almost no instructor time was needed to grade and provide feedback on sketch worksheets. Quantitatively, it took, on average, five hours per week

to grade 196 paper worksheets (~1.5 minutes per worksheet, which includes hand-writing feedback messages similar to those that students automatically receive from the CogSketch tutor) versus only 7.5 minutes per week to grade 66 sketch worksheets (~0.11 minutes per worksheet), an order-of-magnitude difference. Therefore, we saw that sketching opportunities could increase in courses and save time for instructors, while still providing helpful, effective feedback for students.

Through the development process, we also learned that the CogSketch program and sketch worksheets are adaptable and able to change with continual use. Instructors have the capabilities to make changes to sketch worksheets to better serve their courses. This well positions the program and sketch worksheets for continued growth and usage as new tasks and worksheets develop in courses.

### Deploying Sketch Worksheets in Geoscience at Northwestern

Sketch worksheets were used alongside paper exercises in physical geology laboratory sections (Northwestern's Earth 201) during three 10-week quarters. Course enrollments ranged from 20 to 41 students. Students completed worksheets on topographic maps, geologic time, geologic structures, earthquakes, floods and flood recurrence, and glacial movement. This section briefly describes the pedagogical approach to using worksheets in the laboratory, summarizes qualitative observations from the classroom, and presents quantitative estimates of grading times.

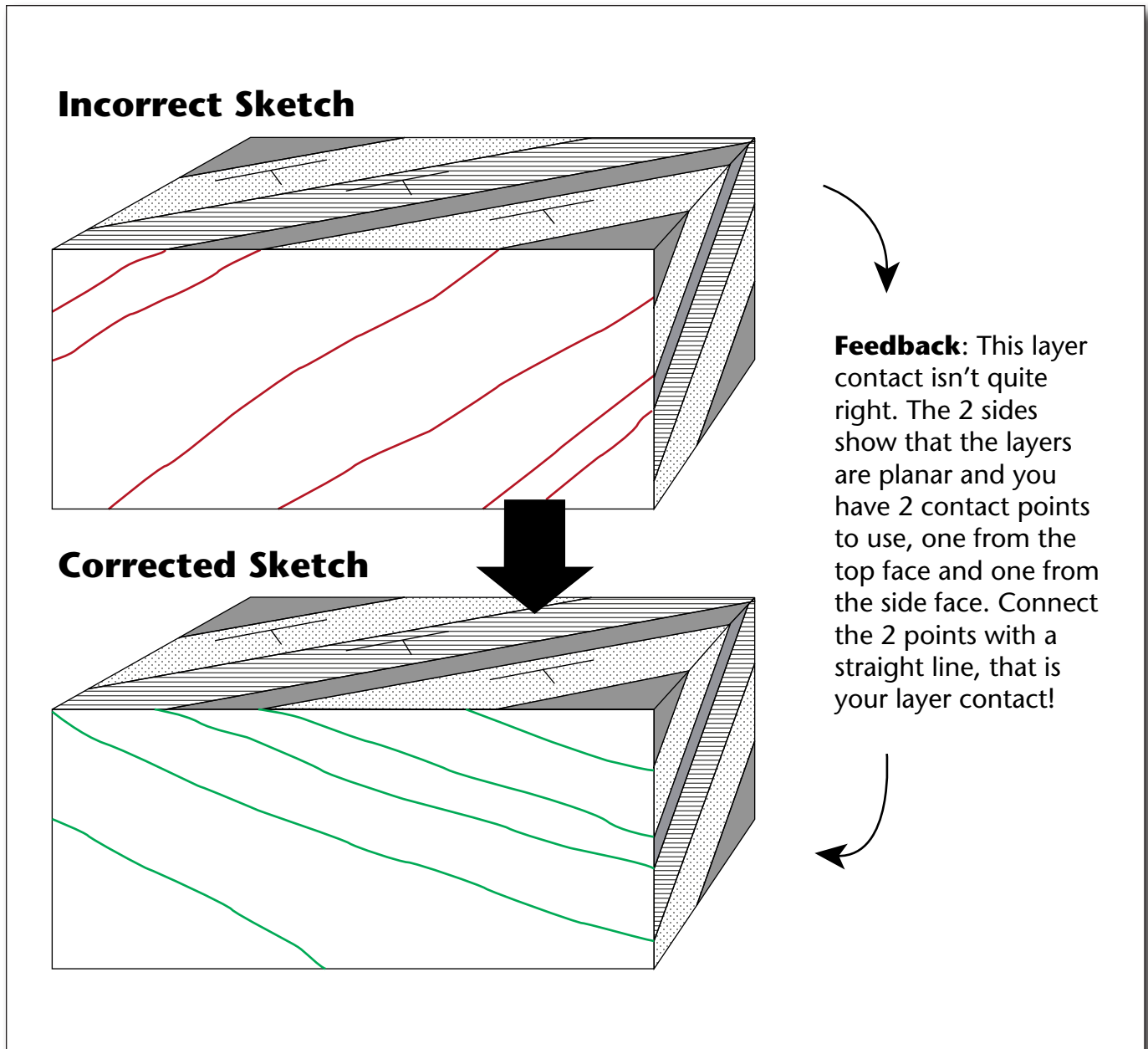


Figure 4. Learning How to Model the Unseen World.

This kind of penetrative thinking is particularly difficult for geoscience students.

Worksheets were selected from the assignments developed by Garnier et al. (2017; and see previous subsection, Developing Sketch Worksheets for Geoscience), based on alignment with class learning objectives and estimated time needed to incorporate the exercises into the laboratory session. Nearly 500 worksheets were completed, submitted, and graded. One or two Sketch worksheets were typically completed by students in class on tablet personal computers (PCs) after they finished their paper exercises and were submitted as image files that were uploaded to

the campus learning management system. Students used the graphical feedback meter that is part of the program to evaluate their progress and to determine if their sketches met the minimum requirements to receive credit for their work. Sketches were graded using a pass/fail scheme. For a passing grade, the following criteria needed to be met: All components of the worksheet were attempted and at least 70 percent of the sketch was correct, as judged by the sketch worksheet tutor. Initial grading of the sketches was done in the course learning management system.

Using tablet PCs with sketch worksheets had impacts on both social dynamics and efficiency. In social dynamics, both potential benefits and drawbacks were observed. The drawbacks include decreased content-related questions, a shift toward technical support of tablets or worksheets, and a decrease in student–student interaction. However, an observed benefit was increased group focus when students switched to using tablets. Next we explain the significance of student–student interactions in a laboratory environment and our concept of group focus, and evaluate these concepts with respect to achieving course learning objectives.

Student–student interactions are commonly engineered into most STEM laboratory environments by pairing or grouping students to create opportunities for interaction. These interactions are intended to reduce the average time needed to complete laboratory work, reduce student demands on the instructor, and to bring students along more rapidly. As such, laboratory activities that promote interaction are pedagogically more sound and are broadly considered advantageous in achieving learning objectives. The fact that introducing sketch worksheets into the laboratory environment reduced the frequency of student interactions might be regarded by some as a drawback, because students are completing the laboratory problems on their own. However, several observations suggest that reduced interaction in the context of a digital tutor is present, but this may not be a problem. First, the results of Garnier et al. (2014) indicate that students using sketch worksheets typically had similar or improved performance relative to students who completed paper worksheets. Second, student demands on the instructor did not increase — instead, the opposite was observed. Students continued to interact with their laboratory partners while using tablet PCs, and if they were unable to succeed on a worksheet, they typically consulted their laboratory partner first before going to the instructor. In other words, some student–student interaction was replaced by student–computer interaction. The one additional cost is that some in-class time was used to manage problems with the tablet PCs, a factor that should shrink as the technology continues to improve.

Introducing sketch worksheets also improved the focus of the group. Within the context of this deployment, group focus is defined as the apparent collective increase in student engagement with course content. Observations that support this interpretation included fewer and quieter student conversations and individual student attention directed toward tablets. The benefit here would likely be increased classroom productivity. However, similar observations have been made in other laboratory environments where sketch worksheets and tablets were not used. The transition from paper to digital exercises appeared to be more abrupt in this deployment and suggests that there may be a causal

relationship between student behavior and the transition to a different content delivery medium. There are three plausible explanations for increased focus. First, using digital devices often pulls students into an activity: Focusing on screens is a social behavior that seem ubiquitous and even habitual for today's students. Second, it could be more difficult to share work or results on tablet PCs. Although it can be more difficult for the instructor to see tablet screens over the shoulders of students, laboratory partners sitting next to each other would swap or share tablets. Third, and what we suspect is most likely, the increased group focus was due to students being able to access immediate and personalized feedback from the sketch worksheet tutor, and they sought this feedback out first. Additional research is needed to explore this.

Sketch worksheets on tablet PCs influenced teaching efficiency in the areas of laboratory preparation, in-class use, and grading. In terms of laboratory preparation, tablet PCs needed to be maintained and access provided to assignments via the course management system. Occasional bugs when the software tutor did not recognize changes made by students were observed. These were easily fixed by saving and then restarting the worksheet, and selecting *update* or by quickly redrawing the sketch in a new worksheet. The small screen size (~10.5 inches) sometimes made it more challenging to read feedback messages and to find errors in sketches, especially in exercises that involved scaling (for example, requiring the user to zoom in and out) and the positioning of many small glyphs. Benefits occurred in the areas of in-class use and grading. There were lower barriers to starting worksheets. Few if any students had questions about how to get started on a worksheet. It was also easier to read labels on sketches and there was a reduction in paper assignments to track and hand back in class.

None of the drawbacks presented were significant enough to deter further use of sketch worksheets in future laboratory sections and plans for new deployments are being made. Tasks such as tablet maintenance will need to be absorbed into the teaching process and the issue of diagnosing problems with sketches may be solved in several ways: using larger tablets or laptops, using a classroom management system to observe student work, or by designing worksheets to optimize the length of feedback and minimize the use of small glyphs.

The gradebook software was not originally used in the Northwestern classes in part because paper worksheets were also still being used during laboratory sessions, which had to be graded by hand anyway, and the pass/fail rubric on images was fairly efficient to grade. We used the Northwestern data to perform an additional evaluation of the efficacy of the software gradebook, using 20 batches of sketches composed of 10–19 exercises, sampling across the exercise types used in class. Batches of exercises were regraded using the tutor and its



Rubric	Paper + Tutor-Based		Tutor
	Pass/Fail ( $n = 7$ )	Tallied ( $n = 7$ )	Tallied ( $n = 6$ )
Time (minutes)/sketch	0.15(0.025)	0.40(0.090)	0.42(0.094)

Table 1. Estimates of Average Grading Times in the Northwestern Geoscience Deployment.

Note:  $n$  = number of batches;  $2\sigma$  values shown in parentheses.

built-in quantitative rubric for which points were deducted and tallied, and in a paper format using the same rubric. Grading times varied both by exercise and grading method. Average results from these rubrics and methods are presented in table 1. Single sketches were graded in less than a minute and batch times ranged from 2 to 10 minutes. On average, the pass/fail method was the quickest per sketch. Paper-based and machine grading using tallied points produced similar grading times per sketch.

These times are substantially less than those found in the University of Wisconsin–Madison deployment, for two reasons. First, all students at Northwestern were using sketch worksheets, with the requirement that their score was at least 70 percent on the tutor before they turned in their work to be graded by hand. So, students had already benefited from the feedback the sketch worksheet tutor provides, thereby reducing the number of mistakes in what was turned in and simplifying the manual grading process. Second, the grading rubric at Northwestern, as previously noted, was pass/fail, whereas for the paper worksheets at University of Wisconsin–Madison, not only did grades have to be assigned but the instructor also had to provide feedback, because that was their only source of help. This is further evidence that sketch worksheets can help with grading efficiency. Another potential advantage is that, unlike people grading paper worksheets, the software never suffers from fatigue.

## Sketch Worksheets for Knowledge Representation and Reasoning

Arguably, computer science is one of the least spatial disciplines in STEM. Spatial models are often used in introductory courses (for example, contour models for depicting variable scope; box-and-pointer notation for describing data structures; process diagrams for describing the flow of computation). However, textual media dominates the everyday work of computer scientists, at least in terms of what they share with each other most often. Nevertheless, we believe that spatial models provide an important part of the tacit knowledge that computer scientists use. In knowledge representation, the rise of knowledge

graphs provides a new opportunity for using spatial learning to support students. CogSketch's visual language can be used to express knowledge graphs by using entity glyphs to denote concepts, and relation glyphs to express the links between them. Moreover, unlike existing concept map tools, arbitrary ink can be used to depict nodes, hence providing additional scaffolding and mnemonics. However, the same issues of grading efficiency arise in knowledge representation classes, even more so with the current flood of computer science enrollments. Consequently, we developed and deployed sketch worksheets in CS 371: Knowledge Representation and Reasoning, taught by Forbus at Northwestern University. This section summarizes what worksheets were developed and our experiences in deploying them. We describe each in turn.

## Using Sketch Worksheets for Knowledge Representation

Sketch worksheets were used in four out of the five homework assignments in the course. (The fifth assignment was an update on students' progress on their term projects, which absorbed their energies for the second-half of the 10-week course.) The first assignment asked students to express relationships between concepts (Dog, Cat, Animal, Plant, Carnivore, Organism) via Venn diagrams, using containment and disjointness to represent what (in Cyc) would be *genls* and *disjointWith* relationships, thereby connecting these new ideas with prior learned models (see figure 5). All but two students received a perfect score, which is not surprising given that this was essentially a warm-up exercise. Even so, feedback was needed, because students often forgot about the existence of carnivorous plants. Students used the feedback four times on average during this assignment.

The second assignment gave them practice in representing everyday situations. It consisted of two sketch worksheets. The first provided a drawing in the background, of a person standing on a floor, with a ceiling overhead that had a light connected to it. Students were expected to add entities representing what was depicted, choosing the most appropriate concept to represent each one of them, and to draw a specific subset of the relationships that held between them. The set of concepts and relationships

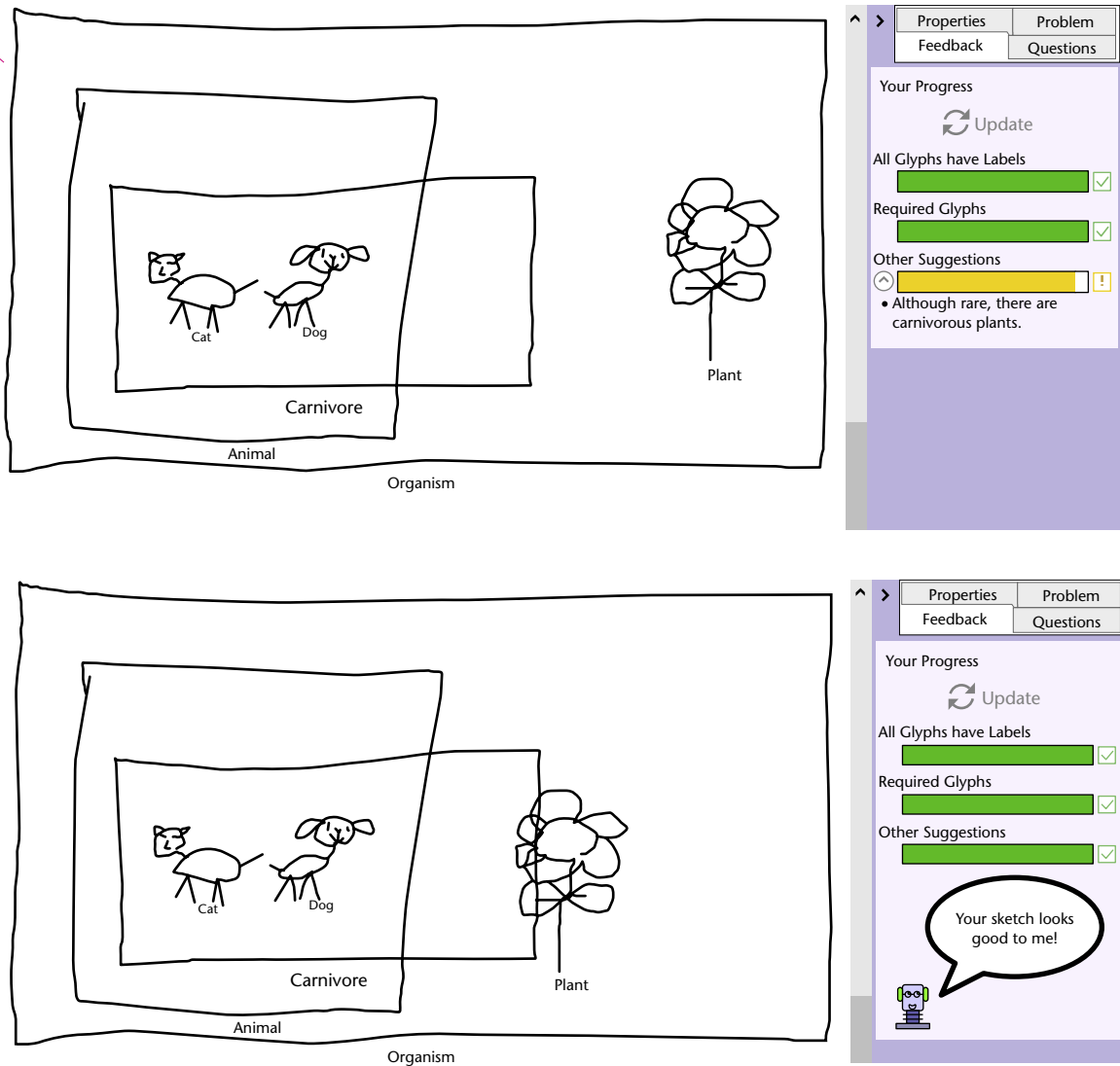


Figure 5. Student Misconception and Corrected Version.

was chosen from the knowledge base to have both the correct concept and a reasonable set of tempting distractors. Similarly, a second worksheet in that assignment asked them to draw landmarks on Northwestern's North Campus "the way you might draw them to explain their layout to a visiting friend" and then add relation arrows to indicate the spatial relationships between adjacent objects (five in all). The only wrinkle with this worksheet was that new blank relationships had to be added and renamed to look like the actual Cyc relationships that should be used. Otherwise, CogSketch automatically inferred

the correct geospatial relationships and declared the sketch to be finished before the student drew anything! 57 out of 58 students received perfect scores, accessing feedback six times and nine times on average, for the two worksheets respectively.

The third assignment required them to fill out a worksheet on a mythical soap opera ("The Eternal Turmoil"), whose contents had been informally specified via a student-driven discussion in class. Soap operas, as noted by Brachman and Levesque (2004), provide marvelous scope for practice with representing events, relationships, and causality. This story

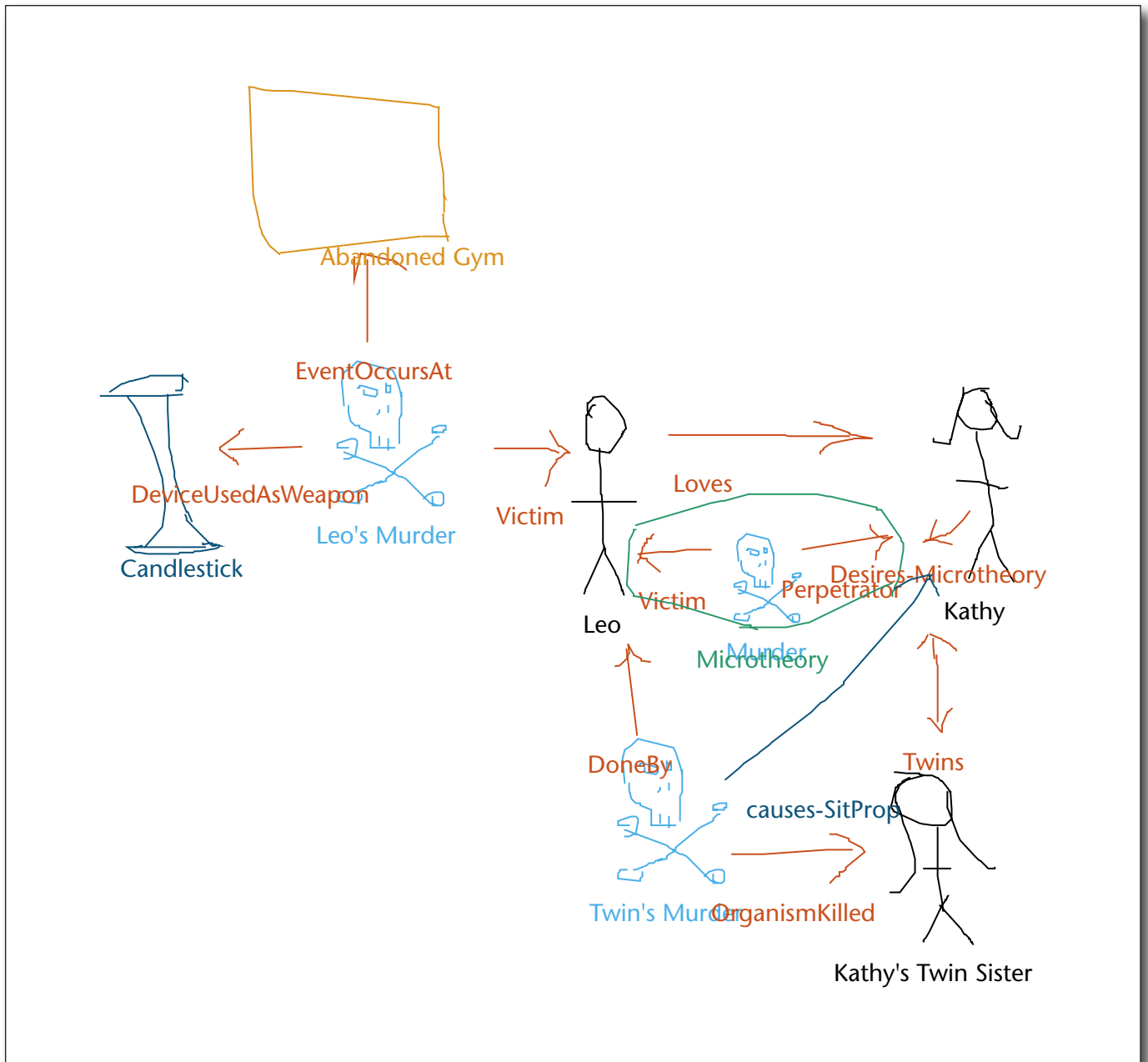


Figure 6. Encoding a Murder Mystery.

Note the glyphs inside a region, indicating a microtheory corresponding to one of the participant's desires, rather than the world itself.

included an event ("Leo is murdered in an abandoned gym with a candlestick"), which is depicted by an entity for the event itself, entities for the roles in the event, and role relations connecting them. A subtlety, supported by Cyc's use of microtheories — which provide Cyc's notion of local context — is handling desires. Here, as shown in figure 6, "Kathy wants to murder Leo because Leo killed her twin sister." This desired murder must be distinct from the actual murder that occurred, which is done in the formalism via a separate microtheory, linked to

the person who wants it via the Desires-Microtheory relationship. Microtheories can be depicted via entity glyphs, with glyphs inside them depicting the facts specific to that microtheory. Seventy-eight percent of the students received perfect scores, but even so, 22 percent turned in worksheets with one or more problems, typically concerning a missing event or relationship. The additional complexity of this assignment can be seen from the use of the feedback system 40 times on average by each student during its completion.

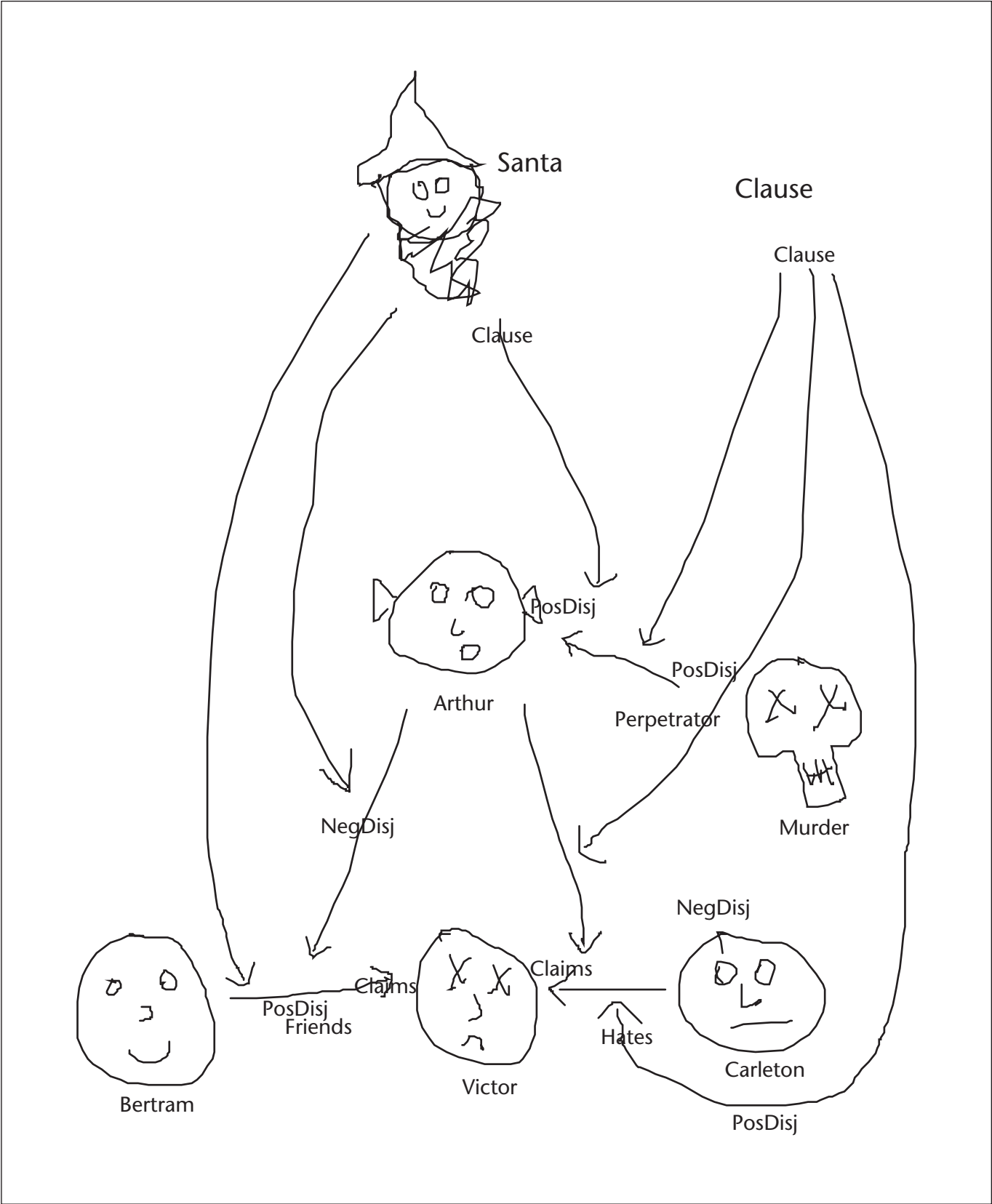


Figure 7. Student Sketch.

This student sketch is an example of a correct rendering of a murder mystery in clause form, from a textbook problem.

The final sketch worksheet had students encode a different murder mystery from a homework assignment in the textbook,<sup>5</sup> to give them practice in producing disjunctive normal form clauses from natural language specifications (figure 7). Each clause was depicted by an entity glyph that was connected to its terms by relationships arrows (*posDisj* indicates that a proposition is a positive disjunct in the clause, and *negDisj* indicates that a proposition is a negative disjunct). This worksheet went beyond the limit of the built-in analogical matching support in sketch worksheets, with the ambiguity in the multiple *posDisj* relations often leading to errors in mapping (and hence feedback and grading) in pilot testing,<sup>6</sup> so we turned feedback off (warning students of this, because by now they expected it) and grading the worksheets by eyeballing them using the gradebook. Students did indeed do worse, with only 38 percent of them achieving perfect scores without on-the-spot feedback.

## Conclusions and Future Work

Our experience with deploying sketch worksheets indicates that the technology has reached the point of achieving most of its goals. Specifically, they can be used by students in more than one discipline (geoscience, AI), and they can be authored by domain experts and instructors (as indicated by the geoscience experience). Grading efficiency is enhanced, as is the ability for an instructor to gain a deeper understanding by browsing through the history of a student's work on a sketch, something that is simply unavailable with pencil and paper sketches (barring video analysis and drawing with multiple color pens, two laboratory practices that are completely impractical for classroom-scale use and impossible for homework assignments). The one remaining goal to be demonstrated is showing that using sketch worksheets actually improves student learning, compared with both nonsketching exercises and sketching on pencil and paper. There is already evidence that sketching can provide gains over verbal self-explanation in understanding texts (Scheiter et al. 2017), and the sketching experience for students is sufficiently fluent that we would expect it to hold for sketch worksheets as well. But such experiments remain to be done, ideally with randomized controlled trials across balanced classrooms. Removing the bottleneck of grading burden should facilitate those experiments being done in the future. However, we note that the geoscience worksheets focus on implementing research-based techniques that have previously been shown to improve learning. In addition, the fact that CogSketch and sketch worksheets can be used in an actual course and greatly reduce grading time for instructors is a major accomplishment that not all educational tools can claim. Finally, many of the lessons concerning feedback in cognitive tutors may be applicable to sketch worksheets, but again, this is a subject for future experimentation. CogSketch's visual analysis

capabilities provide the prospect of using sketches as a medium for educational data mining, and using analogical generalization to help instructors identify common patterns of misconceptions (Chang and Forbus 2014). We hope that these deployments are just the next step of helping spread sketching more broadly through STEM education.

## Acknowledgments

This work was supported by National Science Foundation-funded Spatial Intelligence and Learning Center Award SBE-1041707.

## Notes

1. The contents of the knowledge base are derived from NextKB, which integrates material from OpenCyc, FrameNet, VerbNet, and WordNet, along with a large lexicon and support for qualitative reasoning, spatial reasoning, and analogical reasoning and learning. NextKB is available under a Creative Commons Attribution license, from [www.qrg.northwestern.edu/nextkb/index.html](http://www.qrg.northwestern.edu/nextkb/index.html).
2. <http://www.qrg.northwestern.edu/nextkb/index.html>.
3. The current version of NextKB has 82,438 collections and 20,327 binary predicates. There are 38 types of annotations, supporting the indication of various quantity values, directions of motion and rotation, and other physical properties.
4. [serc.carleton.edu](http://serc.carleton.edu).
5. Problem 3, Chapter 4 in *Knowledge Representation and Reasoning* by Brachman and Levesque (2004).
6. Because the material was abstract, the quantitative grounding techniques introduced previously (Chang and Forbus 2012) were inapplicable.

## References

- Ainsworth, S.; Prain, V.; and Tytler, R. 2011. Drawing to Learn in Science. *Science* 333(6046): 1096–7. doi.org/10.1126/science.1204153
- Brachman, R., and Levesque, H. 2004. *Knowledge Representation and Reasoning*. San Francisco, CA: Morgan-Kaufmann.
- Chang, M., and Forbus, K. 2012. Using Quantitative Information to Improve Analogical Matching Between Sketches. *Proceedings of the 24th Annual Conference on Innovative Applications of Artificial Intelligence*. Palo Alto, CA: Association for the Advancement of Artificial Intelligence.
- Chang, M. D., and Forbus, K. D. 2014. Using Analogy to Cluster Hand-Drawn Sketches for Sketch-Based Educational Software. *AI Magazine* 35(1): 76–84. doi.org/10.1609/aimag.v35i1.2505
- Cohn, A.; Bennett, B.; Gooday, J.; and Gotts, N. 1997. Qualitative Spatial Representation and Reasoning with the Region Connection Calculus. *GeoInformatica* 1(3): 275–316. doi.org/10.1023/A:1009712514511
- Cooper, M.; Grovea, N.; Pargas, R.; Bryfczynski, S.; and Gatlin, T. 2009. OrganicPad: An Interactive Freehand Drawing Application for Drawing Lewis Structures and the Development of Skills in Organic Chemistry. *Chemistry Education Research and Practice* 10(4): 296–301. doi.org/10.1039/B920835F
- de Silva, R.; Bischel, T. D.; Lee, W.; Peterson, E. J.; Calfee, R. C.; and Stahovich, T. 2007. Kirchhoff's Pen: A Pen-Based



## AAAI Members — Get Ready to Vote!

*The link to the electronic version of the annual AAAI Ballot will be emailed to all regular individual AAAI members soon. The online voting system is expected to close on June 26. The ballot will be available via the online system only. If you have not provided AAAI with an up-to-date email address, please do so immediately by writing to membership20@aaai.org.*

Circuit Analysis Tutor. In *Proceedings of the 4th Eurographics Workshop on Sketch-Based Interfaces and Modeling*. doi.org/10.1145/1384429.1384447

Forbus, K.; Chang, M.; McLure, M.; and Usher, M. 2017. The Cognitive Science of Sketch Worksheets. *Topics in Cognitive Science* 9(4): 921–42. doi.org/10.1111/tops.12262

Forbus, K.; Ferguson, R.; Lovett, A.; and Gentner, D. 2016. Extending SME to Handle Large-Scale Cognitive Modeling. *Cognitive Science* 41(5): 1152–1201. doi.org/10.1111/cogs.12377.

Forbus, K.; Usher, J.; Lovett, A.; Lockwood, K.; and Wetzel, J. 2011. CogSketch: Sketch Understanding for Cognitive Science Research and for Education. *Topics in Cognitive Science* 3(4): 648–66. doi.org/10.1111/j.1756-8765.2011.01149.x.

Garnier, B.; Chang, M.; Ormand, C.; Matlen, B.; Tikoff, B.; and Shipley, T. 2017. Promoting Sketching in Introductory Geoscience Courses: CogSketch Geoscience Worksheets. *Topics in Cognitive Science* 9(4): 943–69. doi.org/10.1111/tops.12291.

Jee, B.; Gentner, D.; Uttal, D.; Sageman, B.; Forbus, K.; Manduca, C.; Ormand, C.; Shipley, T.; and Tikoff, B. 2014. Drawing on Experience: How Domain Knowledge Is Reflected in Sketches of Scientific Structures and Processes. *Research in Science Education* 44(6): 859–83. doi.org/10.1007/s11165-014-9405-2.

Koedinger, K. R.; Anderson, J. R.; Hadley, W. H.; and Mark, M. A. 1997. Intelligent Tutoring Goes to School in the Big City. *International Journal of Artificial Intelligence in Education* 8: 30–43 Berlin: Springer.

Lee, W.; de Silva, R.; Peterson, E.; Calfee, R.; and Stahovich, T. 2008. Newton's Pen — A Pen-Based Tutoring System for Statics. *Computers & Graphics* 32(5): 511–24. doi.org/10.1016/j.cag.2008.05.009

Lovett, A., and Forbus, K. 2011. Cultural Commonalities and Differences in Spatial Problem Solving: A Computational Analysis. *Cognition* 121(2): 281–7. doi.org/10.1016/j.cognition.2011.06.012

Lovett, A., and Forbus, K. 2012. Modeling Multiple Strategies for Solving Geometric Analogy Problems. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Sapporo, Japan.

Lovett, A., and Forbus, K. 2013. Modeling Spatial Ability in Mental Rotation and Paper-Folding. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society*. Berlin, Germany, 930–935.

Lovett, A., and Forbus, K. 2017. Modeling Visual Problem Solving as Analogical Reasoning. *Psychological Review* 124(1): 60. doi.org/10.1037/rev0000039.

Matuszek, C.; Cabral, J.; Witbrock, M.; and DeOliveira, J. 2006. An Introduction to the Syntax and Content of Cyc. In *2006 AAAI Spring Symposium on Formalizing and Compiling Background Knowledge and its Applications to Knowledge Representation and Question-Answering*. Palo Alto, CA: AAAI Press.

Scheiter, K.; Schleinschok, K.; and Ainsworth, S. 2017. Why Sketching May Aid Learning From Science Texts: Contrasting Sketching with Written Explanations. In *Topics in Cognitive Science*. New York: Wiley. doi.org/10.1111/tops.12261.

Valentine, S.; Vides, F.; Lucchese, G.; Turner, D.; Kim, H.; Li, W.; Linsey, J.; and Hammond, T. 2012. Mechanix: A Sketch-Based Tutoring System for Statics Courses. In *Proceedings of the 24th Innovative Applications of Artificial Intelligence Conference (IAAI 12)*. Toronto, Canada.

Yin, P.; Forbus, K.; Usher, J.; Sageman, B.; and Jee, B. 2010. Sketch Worksheets: A Sketch-Based Educational Software System. In *Proceedings of the 22nd Annual Conference on Innovative Applications of Artificial Intelligence*. July 11–15, 2010, Atlanta, GA.

**Kenneth D. Forbus** is the Walter P. Murphy Professor of Computer Science and Professor of Education at Northwestern University. He received his degrees from the Massachusetts Institute of Technology (PhD in 1984). His research interests include qualitative reasoning, analogical reasoning and learning, spatial reasoning, sketch understanding, natural language understanding, cognitive architecture, reasoning system design, intelligent educational software, and the use of AI in interactive entertainment. He is a Fellow of the Association for the Advancement of Artificial Intelligence, the Cognitive Science Society, and the Association for Computing Machinery. He is the inaugural recipient of the Herbert A. Simon Prize, a recipient of the Humboldt Research Award, and has served as Chair of the Cognitive Science Society.

**Bridget Garnier** is a PhD candidate in the Department of Geosciences at the University of Wisconsin–Madison.

**Basil Tikoff** is a geologist and professor at the University of Wisconsin–Madison. In addition to disciplinary studies in the geological sciences, he has worked with cognitive scientists and geoscience educators on spatial learning within the natural sciences.

**Wayne Marko** is an adjunct earth science instructor at several community colleges in the Chicago area. His research interests are in the areas of igneous petrology, magma emplacement, and tectonic environments.

**Madeline Usher** is a computer programmer and researcher at Northwestern University. Her research interests include spatial and qualitative reasoning. She is the chief programmer for CogSketch, an open-domain sketch understanding and spatial reasoning system.

**Matthew McLure** completed his PhD in computer science from Northwestern University in March 2019. His research focuses on machine learning via analogical matching and structured spatial representations.