Research Statement: Kate Lockwood, PhD Candidate
Qualitative Reasoning Group, Northwestern University

Current Research

I am interested understanding how to build systems that can learn from multi-modal presentations of information such as text and diagrams. I am especially interested in approaching this problem from a cognitive modeling perspective. By understanding how people use these sources of information I hope to build systems that can learn from the same sources of information, that can support human learners, and that can evaluate human learning. I am interested in how the regularities and conventions of instructional media can be exploited to aid in learning. I am also interested in exploring the roles of different types of reasoning, such as analogy in multi-media learning and in spatial reasoning.

My work so far has focused on multi-modal learning from text and diagram pairs like those found in textbooks. In multi-modal knowledge capture experiments, I have taken examples from textbooks. Each example is translated into controlled English and the diagrams are sketched. As part of this system I have worked with both a natural language understanding system and a sketch understanding system. I have also modeled the learning and use of spatial language categories. Each of these areas of research is described in detail below.

Multi-modal Knowledge Capture. Humans learn a variety of concepts through a combination of text and diagrams. One example of this are science textbooks, where diagrams are often used to demonstrate the make-up of a physical system while the accompanying text describes a process that occurs in the depicted system. In order for intelligent systems to be able to learn by reading, they need to be able to combine information from text and diagrams into a coherent representation suitable for reasoning. I am building a system that takes as input simplified English text and sketched diagrams which are taken from a physics textbook (*Basic Machines and How They Work*). As output it creates a coherent representation that is suitable for storing in a large-scale knowledge base. The system's knowledge is tested using problems from the back of the book – the same questions used to test human comprehension.

One problem I have tackled while addressing the multi-modal knowledge capture problem is depiction conventions. Across domains, there are semi-standard ways that diagrams are drawn. These conventions, such as call-outs or cut-aways, often have implications for how we reason about the information presented. Humans are easily able to recognize these conventions and use them to interpret diagrams, even for less clearly defined conventions. For example, a single wavy line might be used to indicate the level of liquid in a cylinder. A human user is easily able to recognize that the line indicates that the cylinder below the line is full of the liquid. Situations like this often confuse recognition-based diagram understanding systems. My approach to diagram understanding is based on sketched diagrams where users have labeled the objects in the diagram (see [5]).

In this system the wavy line might be labeled "water". Then the system can use common sense knowledge about water, taken from an underlying knowledge base, to make the same inferences that a person would when examining the same diagram. This requires combining basic knowledge about the entity being represented (water) and spatial information from the sketch (what other objects in the sketch might constrain the extent of the water). Humans rarely encounter a diagram in a vacuum; knowing what entities are in a sketch (either through a caption, accompanying text, or a direct written or verbal label) allows humans to understand the information being communicated. This is the behavior that I am modeling. My approach contrasts with that of recognition-based diagram understanding systems which require objects in a sketch to match against a library of predefined templates. There are several problems with this approach: it requires you to know *a priori* the objects you will encounter, it allows each template to match to only one interpretation, and it constrains your system to a given domain.

I am taking a domain-independent approach that exploits labeled sketches to allow for much greater flexibility in what can be depicted and also allows for much deeper reasoning. My approach to diagram understanding is to place as few constraints as possible on how diagrams are drawn. The goal for my system is for it to automatically pick out important entities and to be able to distinguish conceptually important entities from superfluous detail or sketch variation (like shading).

Modeling Spatial Language. Text describing physical systems tends to include a lot of spatial prepositions. These are words like *in, on, above, below*, etc that describe the spatial relationship between pairs or sets of objects. The assignment of a spatial relationship to a visual scene is a complex decision involving information about the spatial configuration of objects and functional information about the objects themselves. There is a large body of research from the psychology and linguistics communities about how people learn and use spatial prepositions, making spatial language an ideal candidate for cognitive modeling experiments. I have done several experiments modeling the use of spatial prepositions.

In [1][6] I built cognitive models that could learn spatial preposition categories from sketched examples of the relationships. In [2] I showed how to use knowledge about how spatial language is used to assign prepositions to novel scenes. In [6] I was able to learn different sets of prepositions (Dutch and English) using the same set of sketches. In this work, I rely on common-sense information about objects in sketches from an underlying knowledge base (the Cyc KB). I combine that information with qualitative spatial information from the ink in the sketches to learn the spatial language categories. Learning is done via progressive abstraction, which is driven by analogy and similarity.

I am currently interested in expanding this work to look at how a system can use partial information about spatial prepositions to strengthen spatial language categories or to learn new ones. For example, if a system has a basic understanding of some simple prepositions involving two objects and comes across an example of *between*, can the system build a rudimentary understanding of *between*? Can that understanding be refined through further exposure to new examples of *between*? And can the understanding of *between* that is built be used in new situations?

Sketch Understanding. Much of my work involves sketched input. To these ends I have been very involved in the CogSketch open-domain sketch understanding system [4]. CogSketch was used to sketch all of the inputs used in my experiments. In addition to being a tool for cognitive modeling, CogSketch is also intended to be used as an educational tool. Teachers can use CogSketch to create worksheet problems and supply the system with a correct solution. Based on comparing the teacher solution to the student problem, CogSketch can provide the student with tutoring advice.

Future Research Plans

Multi-modal Knowledge Capture. I plan to expand my work in multi-modal knowledge capture in several directions. First, I am interested in increasing the autonomy of my system by adding the ability to discover and learn novel terms. In textbooks students often learn a lot of new vocabulary, especially in highly technical domains. I plan to explore ways for the system to recognize that a term is novel and then develop an understanding for that term. I am also interested in how the structure of instructional materials can be exploited to aid in learning and tutoring systems. Many instructional materials have the same general sections (e.g. diagrams, examples) and the same discourse structure. I am interested in examining how this structure and specific cues in the text can give clues to how different modalities should be integrated.

Sketch Understanding. I plan to continue to work on sketch understanding, including building on the depiction convention system using a wider variety of diagram sources. I am also interested in exploring how diagrams help students learn and how an intelligent system can use student-drawn diagrams to diagnose misconceptions. I also plan to use human-computer interaction studies to understand how to make sketching technologies more natural and easy to use. Additionally, I would like to study how people reason about motion using static sketches and how series of sketches can be used to depict multi-step processes and building systems that can reason about the changes between sketches in a series. All of these projects will involve a combination of sketching studies with human users, and computational modeling of results.

Language Understanding. I also plan to build on some of my work in language understanding. I am interested in looking at the structure of language, in particular, how a system can exploit the structure of instructional language to help bootstrap learning of concepts. I also plan to continue my work on spatial language. In addition to continuing my work on spatial prepositions, I hope to work on how an understanding of spatial language can help intelligent systems to carry on discussions with human users about spatial objects such as diagrams, pictures and maps.

Computer Science Education. As I am interested in both instructional media and computer science, a natural progression is to examine methods of teaching computer science. As enrollments in computer science departments decline, it is more important than ever to develop ways to make computer science courses applicable to a wider variety of students and to encourage students with an interest in computer science to engage with the material. I am particularly interested in examining ways to make success in introductory computer science courses more accessible to non-traditional populations.

Establishing a Research Program

The projects that I have been involved with are highly interdisciplinary, involving collaboration with colleagues from psychology. My work also has overlap with linguistics, learning sciences, and communications. In establishing my own research program I plan to actively seek out colleagues from across the university as well as researchers from other schools to collaborate with.

Within each of the areas that I hope to pursue, there are a variety of smaller, discrete chunks of research that are ideally suited to student involvement. I think that it is very important to get students involved in research as early as possible and that early exposure to interesting research problems can help motivate students to study computer science. Within my research goals, I have many projects suitable for senior theses, independent studies, and summer research opportunities that would appeal to a variety of students.

Selected References

[1] Lockwood, K., Forbus, K., Halstead, D. & Usher, J. (2006). Automatic Categorization of Spatial Prepositions. Proceedings of the 28th Annual Conference of the Cognitive Science Society. Vancouver, Canada.
[2] Lockwood, K., Forbus, K., & Usher, J. (2005). SpaceCase: A model of spatial preposition use. Proceedings of the 27th Annual Conference of the Cognitive Science Society. Stressa, Italy.
[3] Lovett, A., Lockwood, K., Dehghani, M., & Forbus, K. (2007). Modeling human-like rates of learning via analogical generalization. *Proceedings of Analogies: Integrating Multiple Cognitive Abilities*. Nashville, Tennessee.
[4] Forbus, K., Lovett, A., Lockwood, K., Wetzel, J., Matuk, C., Jee, B., and Usher, J. (2008). CogSketch. *Proceedings of AAAI 2008*.
[5] Lockwood, K., Lovett, A., Forbus, K., Dehghani, M., and Usher, J. (2008). A Theory of Depiction for Sketches of Physical Systems. *Proceedings of QR 2008*.
[6] Lockwood, K., Lovett, A., and Forbus, K. (2008). Automatic Classification of Containment and Support Spatial Relations in English and Dutch. In *Proceedings of Spatial Cognition*.