Recognition of Hand Drawn Chemical Diagrams

Tom Ouyang OUYANG@CSAIL.MIT.EDU
Randall Davis DAVIS@CSAIL.MIT.EDU
MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge MA, 02139 USA

1. Introduction

Chemists regularly use diagrams to capture and communicate ideas about chemical compounds. By laying out a molecule's spatial structure, they can convey information about its chemical properties and inter-molecule interactions much more clearly than by just giving a chemical formula. While there are a number of software systems today for specifying these chemical structures to a computer, none of them provide the ease of use, naturalness, and speed of drawing on paper. If machines are to interact with us in a way that feels natural, we will have to bridge this gap between how people naturally express ideas and how computers interpret them.

In this paper we describe an approach to building natural and robust sketch understanding systems that is inspired by how people interpret these sketches. When we examine at a sketch, we bring to bear a wide range of information: we look at the low level geometry (e.g., individual lines and arcs), the high level structure (e.g., the relationships between objects), and use our understanding of the semantics of the domain (e.g., what constitutes a valid configuration of atoms in a molecular compound). While some work has explored the first two of these, relatively little effort has been devoted to using knowledge about the domain to better understand and interpret a sketch. To explore this last idea, we have developed a sketch interpretation system that embodies the physical properties governing how components behave and interact. While our implementation focuses on hand-drawn chemical diagrams, such as the one in Figure 1 we believe that the approach presented here could be applied to other domains as well.

Interpreting sketches of molecular diagrams presents several challenges. As seen in Figure 2 A, handwriting can be very messy, and it is often difficult to decompose these symbols into low level primitives such as lines and arcs commonly used by other sketch recognition systems. There are also inherent ambiguities in interpreting messy diagrams. For instance, without the aid of context, the two vertical lines in the "H" in Figure 2 B can also reasonably be interpreted as either as a double bond or as a part of a hashed bond. The circled symbol in Figure 2 C can be

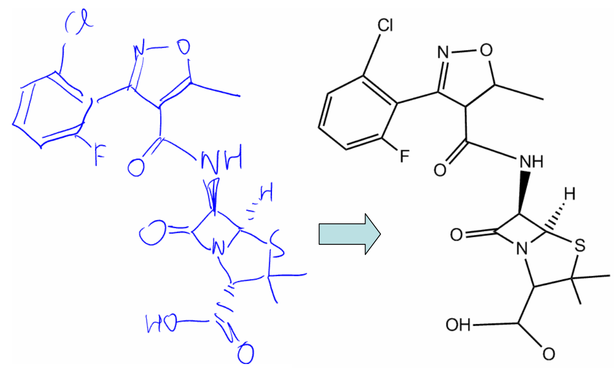

Figure 1. Example drawing of a molecular diagram and its recognized interpretation

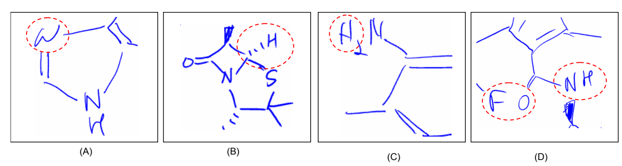

Figure 2. Some examples of recognition difficulties that appear chemical diagrams

interpreted either as an "H" or an "N", and it may be difficult to determine which without using knowledge about chemistry. A third challenge is deciphering the intended structure of the drawing. It is, for example, not necessarily clear in Figure 2 D that the "F" and the "O" should not be connected while the "N" and the "H" should be. Our system attempts to resolve these kinds of ambiguities by using context and chemistry knowledge.

2. System Overview

The sketch interpretation process involves three stages: segmentation and symbol recognition, structure interpretation, and domain verification. We will describe each in turn in the following sections. Once the sketch has been interpreted, the resulting structure can be passed to a variety of different programs such as ChemDraw (which can draw the structure neatly) or SciFinder (which can retrieve information about compounds such as patents or publications).

2.1 Segmentation and Recognition

In order to interpret objects drawn using multiple strokes, we first need to determine which strokes belong to the same symbol. The space of possible groupings that we consider consists of any combination of strokes, up to a certain size, that were drawn sequentially in time. While this approach does require that the user finish one symbol before starting the next, we found in our observations that this not a significant limitation. One possible reason for this is that the symbols used in molecular diagrams are small and spatially contained.

Our system uses a statistical classifier that attempts to interpret each potential group as either an element, a hash bond, or a wedge bond. This classifier considers general geometric features such as size and stroke density as well as text recognition features which are described below. Optimal values for these features are learned during training. The highest scoring groups, along with their recognized symbol labels, are stored for further interpretation. In addition, the system remembers all candidate segmentations generated by this process in case any part of the sketch needs to be re-interpreted. Since we consider groups of consecutive strokes only up to a certain size limit, the total space requirement scales linearly with the number of strokes. Any strokes not classified as one of these three types of symbols are considered solid bonds by default.

In order to correctly interpret text symbols, we incorporate additional features from an independent handwriting recognizer into the classifier. These include the interpreted string returned by the recognizer, the length of the string, and the confidence measure for the match. For this implementation we use the handwriting recognizer provided with the Microsoft Tablet SDK, but the system is designed so that it can easily switch to other handwriting recognizers.

2.2 Structure Interpretation

Once all of the symbols have been extracted from the sketch, the next task is to determine how the different components combine to form the complete structure. Our system identifies connected components by looking for instances where the end of a bond and an element symbol are in close proximity and the direction of the bond is collinear with the element. Each bond is then attached to its nearest element based on these two distance metrics. We have also extended this connection framework to handle two special chemistry notations: implicit elements (carbons and hydrogens that are typically not drawn when they are part of a carbon chain) and benzene rings, denoted by a circle inside a six-carbon ring.

2.3 Domain Verification

Once an initial interpretation has been generated, the next step is to verify that the structure is chemically sound. One indication of a problem is that an element has too many or too few electrons to complete its valence shell. For example, hydrogen require one covalent bond to complete its valence while nitrogen would need 3 electrons to do the same. An inconsistency could mean a misinterpreted symbol (e.g., mistaking an "H" for an "N"), an incorrect bond connection, or a segmentation error. If such an inconsistency does arise, the system attempts to correct the problem by addressing each of its possible causes. Since we store alternatives hypotheses for each stage of the interpretation process, this task is equivalent to searching the space of relevant alternatives. For an improper valence, the system would consider the element in question as well as any bonds that are connected to it.

3. Related Work

Gennari et al. (Gennari et al., 2005) present an approach that uses domain knowledge to help interpret hand-drawn circuit diagrams. Our tasks differ in that chemistry diagrams involve a great deal of interspersed drawing and text and require more complex forms of domain knowledge. Others (Alvarado et al., 2002; Shilman et al., 2002) have used Bayesian networks and hierarchical methods for sketch interpretation, which differ from our work in the types of sketches considered. There have also been efforts to recognize chemical sketches and diagrams. Tenneson and Becker (Tenneson & Becker, 2005) developed a sketch-based system that helps students visualize the three dimensional structure of an organic molecule. Unlike our system, it requires all symbols to be drawn using a single stroke and does not try to use domain knowledge to correct errors in its interpretation. Casey et al. (Casey et al., 1993) presented a system for recognizing chemical graphics from scanned images. However, their work focused on typewritten chemical diagrams instead of freehand drawings.

4. Current Status and Future Work

We have tested an early implementation on a number of sketches collected from chemists and obtained proof-of-concept results. The system was able to recognize most of the diagrams correctly and recovered from many initial errors by using its knowledge of chemistry. We are currently conducting a user study to collect a larger corpus of example chemistry diagrams. This will allow us to examine variations in drawing styles and provide a richer source of training and testing data for our system.

5. Acknowledgements

This work was supported in part by the MIT Oxygen Project, by the MIT/Microsoft iCampus Project,by Intel, Inc., and by Pfizer, Inc.

References

Alvarado, C., Oltmans, M., & Davis, R. (2002). A framework for multi-domain sketch recognition. *AAAI Spring Symposium on Sketch Understanding*, 1–8.

Casey, R., Boyer, S., Healey, P., Miller, A., Oudot, B., & Zilles, K. (1993). Optical recognition of chemical graphics. *Document Analysis and Recognition*, 627–631.

Gennari, L., Kara, L. B., & Stahovich, T. F. (2005). Combining geometry and domain knowledge to interpret hand-drawn diagrams. *Computers and Graphics*, *29*, 547–562.

Shilman, M., Pasula, H., Russell, S., & Newton, R. (2002). Statistical visual language models for ink parsing. *AAAI Spring Symposium on Sketch Understanding*.

Tenneson, D., & Becker, S. (2005). Chempad: Generating 3d molecules from 2d sketches. *http://graphics.cs.brown.edu/research/chempad*.