Tools for performing organic reaction mechanisms over the web

John H. Penn, Chemistry Department, West Virginia University, Morgantown, WV 26506-6045 USA

Christoph Steinbeck, Cologne University Bioinformatics Center, Research Group for Molecular Informatics, Institut für Biochemie, Zülpicher Str. 47, Köln D-50674, GERMANY

Ada Casares, Department of Chemistry, Richard Stockton College, Pomona, NJ 08240 USA

AUTHOR EMAIL ADDRESS John.Penn@mail.wvu.edu

ABSTRACT

Web-based instructional methods have shown constant advancements in recent years. In the world of organic chemistry, on-line structure drawing has become possible through a variety of applets and 3-D visualization techniques are beginning to become highly commonplace. The next mountain to be climbed is that of drawing organic reaction mechanisms, and then to have the computer evaluate its correctness. This contribution will focus on the progress towards that goal and the various techniques that might be used to help students draw and, more importantly, to understand reaction mechanisms.

Introduction

This conference is unique, in that it bridges the very formal written manuscript and the standard presentation of a conference. I hope that you will allow me to use the informality of a standard conference to just write as if I were speaking to you now in an auditorium. I'll open my comments with a couple of lines that I use in almost every talk that I give.

"Practice Makes Perfect": An Introduction to WE_LEARN

Our overall strategy for teaching is based upon the premise that ï¿½Practice Makes Perfectï¿½. It has always been astounding to me that our society (and in the context of this conference, the entire world) expects the top athletes to practice every day. This is particularly so prior to the big championship game of any sport. So, for this conference, let's consider the World Cup of Soccer, currently being played in Germany. It would be absolutely crazy for anyone to think about any team from any country to not have practiced prior to, and during, the event. The same idea is true in the area of music, where professional and amateur musicians must devote lots of practice time prior to the event. Yet, at the same time, the US American public does not put the same emphasis upon education. In fact, at one presentation of my philosophy of repetitions to help the learning process (and I must add that I had not yet started using the phrase ï¿½Practice Makes Perfectï¿½), a high school teacher said that the volleyball player wouldn't have time to do the repetitions in chemistry because they had volleyball practice. We probably won't be able to change the overall attitude of society through this conference, but the same attitude seems to exist among the academic chemistry world.

Let's now try to break apart what happens when one practices. One attempts something and receives immediate feedback about the correctness of the attempt. In music, one plays some series of notes and provides themselves immediate feedback about the correctness, based upon what one personally hears. The immediate feedback is self-provided or usually after some time, a music teacher will provide some feedback about the quality of the performance.

The concept of homework arose from the attempts to get students to learn, in a process where the teacher could provide some feedback. There have been many different models of how to have the homework administered. Trying to make a long story short, they are very time intensive. Let's look at what would typically happen in the large American University. Homework is assigned. Students submit the homework for grading. There is a delay in time while the homework is graded. The homework must be distributed back to the students. In principle, it would be great to do this again, but in most systems, this only occurs when the students are given an exam. Obviously, the demands of teaching large lecture section classes makes this impractical for a professor.

Here's where the advantages of CAI and information technology come to advantage and beg to be used. Computers do not tire from handing out questions, grading answers that are submitted. In fact, computers excel at boringly repetitive processes. They do not sleep (although occasionally, they need rebooted or a maintenance period), so they are available 24/7 (i.e., around the clock). If properly programmed, they keep track of what they've done, so there is a ï¿½paperï¿½ trail available to know what has transpired.

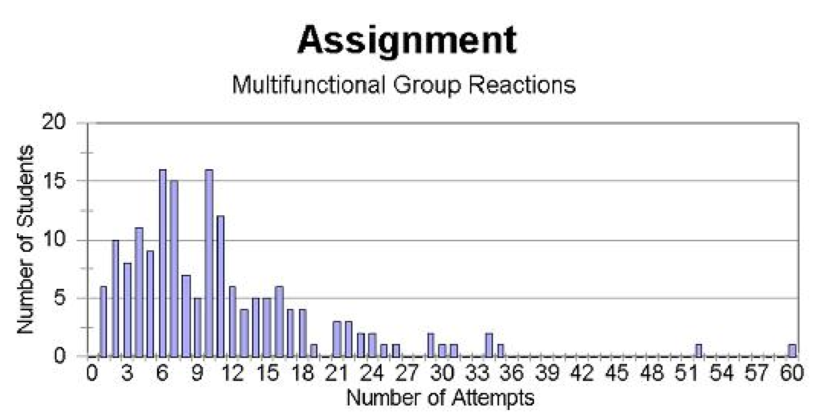
What is WE_LEARN?

With all of these considerations in hand, we developed the ï¿½WE_LEARNï¿½ system for organic chemistry (Penn, et al, 2000). The idea is that there are many examples of a similar concept placed in the computer and the computer will randomly select a question to pose to the student. If the students have truly mastered a concept, then they should be able to answer the question correctly every time. The students submit an assignment for grading and the computer can grade it immediately (or at least the length of time for electrons to travel to the server computer and back). If they get the question wrong, then the computer will provide an explanation of why the answer is incorrect. Hopefully, this is as close to the model of athletic practice or music practice as we can get.

OK, let's demonstrate what we have just said. Point your browser tohttp://www.welearnhorizon.com/q/perception.dll and log into the system with the account name CONFCHEM and password CONFCHEM. Try assignment CONFCHEM01 by clicking on the ï¿½Startï¿½ button. Each question in this assignment is what we would consider to be a separate concept. Therefore, the reaction of an alkene with HI, HBr, and HCl, respectively, are three different, but related concepts. Each question has approximately 30 questions in our database from which one question is chosen randomly. The important issue is to complete the assignment (i.e., click ï¿½Submitï¿½ at the bottom of the page) and then try the assignment again. You will be given a different set of questions.

Why Use WE_LEARN?

The idea that each attempt leads to a different set of questions makes several new and unique possibilities for the instructor. This leads to what is the major strength of our system. Homework is now administered in a completely different way than in the past. ADVANTAGE 1): In the traditional method of administering homework, where all students have the same assignment, one student routinely does the work, while the other students copy. Therefore, there is no benefit to the students who copy. Since all students receive a different assignment, the likelihood of student copying is minimized. It takes a special relationship between two students to get one student to double their own workload to complete the assignment. ADVANTAGE 2): The instantaneous grading and feedback on wrong answers can be used to teach students and encourage mastery of the subject. We believe that this is a true mastery-based learning system. To promote mastery-based learning, we give students a bonus based upon their maximum score for an assignment, regardless of how many times that the students try an assignment. An example of the number of attempts/student is shown in the figure below, where the number of attempts/student is shown on the X-axis, while the frequency of that happening is shown on the Y-axis. Clearly, most students are attempting an assignment anywhere from 3-7 times, therefore, working ca. 30-70 problems, instead of the normal 10 problems that they might have worked under the traditional homework regimine.

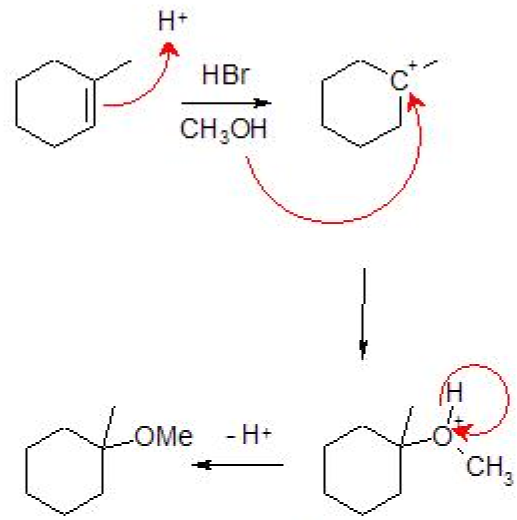


Drawing structures on-line

In preparation for the discussion to come, it is important to be sure of the capabilities of computers on-line. In the beginning of the development of this database and others, organic chemistry talks about concepts in terms of graphical structures. At the onset of development of the WE_LEARN program, it was impossible to draw structures. However, about three years ago, several groups realized that Cheminformatics provided a pathway to draw organic chemical structures on-line and be evaluated for correctness. The following list is not meant to be an all-inclusive list, but is meant to acknowledge that we were not the only group to learn how to evaluate student-drawn structures. The list includes our group, the OWL group led by Bill Vining (one of the presenters in this conference), Webassign, andACEorganic. The trick is to convert the graphical image into a canonical form, such as a SMILES string. The ASCII string generated as the SMILES string can now be compared to a string stored within the database that represents the compound that is the correct answer. This strategy has been used successfully in our classrooms for three years with the major issue being to get the students to practice drawing the structures.

OK, let's demonstrate what we have just said. Point your browser tohttp://www.welearnhorizon.com/q/perception.dll and log into the system with the account name CONFCHEM and password CONFCHEM. Try assignment CONFCHEM02 by clicking on the ï¿½Startï¿½ button.

The Problem: Teaching Organic Reaction Mechanisms



Mechanism for the Addition of Alcohols to Alkenes

With this introduction, we can now turn to the issue of trying to teach students about organic reaction mechanisms. Organic chemists use reaction mechanisms to rationalize a host of data and to understand stereochemistry, reaction rates, etc. The current methodology was stated by Ingold in 1969 and has remained relatively unchanged since that time. ï¿½Curved arrowsï¿½ indicate the flow of electrons, and potential energy intermediates along the reaction coordinate help to understand the chemical reaction. Most teachers believe that students can best predict chemical reactivity using reaction mechanisms or rationalizations. Let's illustrate this with an acid addition to an alkene, which will be the source of most of the examples for this paper. I know that all of you know this, but let me make sure that everyone understands. What we are trying to teach the students is that the methyl group is an electron-donating group, that the acid has a choice of which carbon atom to attach to, and that the subsequent carbocation adds a source of electrons to satisfy its electron deficiency.

This now brings us to the problem for this paper. How can one evaluate student reaction mechanisms via on-line methodology? The development of representational methods for organic reaction mechanisms could lead to a CAI (or a ï¿½Practice Makes Perfectï¿½) solution if the solution can be reduced to something that the computer can compare to a solution already in the computer database (or memory). There are currently a variety of methods for representing chemical reactions via computer formats. These have been summarized recently by Holliday, Murray-Rust, and Rzepa. Since they have already done such a good job of summarization, I'll just take liberally from their synopsis. Cambridge Soft uses the CDX file format. Products and reactants are identified by their precise positioning in a graphical arrangement. MDL ISIS/Draw uses the RXN file format. Daylight uses the SMILES format for molecules and has extended that format to the SMIRKS format for reactions. Last, but not least, is the CMLReact approach proposed by Rzepa and his group in the publication just cited. The problem is that none of these formats include both ï¿½curved arrowsï¿½ and structural representations that can be easily compared by standard computer techniques.

Since there is no method available with current technology, let's back up from this task and look at a host of intermediate solutions to the problem of teaching students how to draw and how to understand organic reaction mechanisms.

Method 1 ï¿½Discoveryï¿½ or ï¿½Inquiryï¿½ Method for Learning Reaction Mechanisms

When we first began the development of WE_LEARN, there was much excitement in the chemical education community about Inquiry-based learning or Discovery-based learning. Hopefully, this will give this conference lots to discuss, because I never have truly understood what the Inquiry-based system was all about. My understanding was that you could present a multitude of data to the students and they would develop their own theories of how things work. Hopefully, with a little guidance, the students would develop the same theories that the scientists of yore developed. And, of course, the students would have fun while doing it.

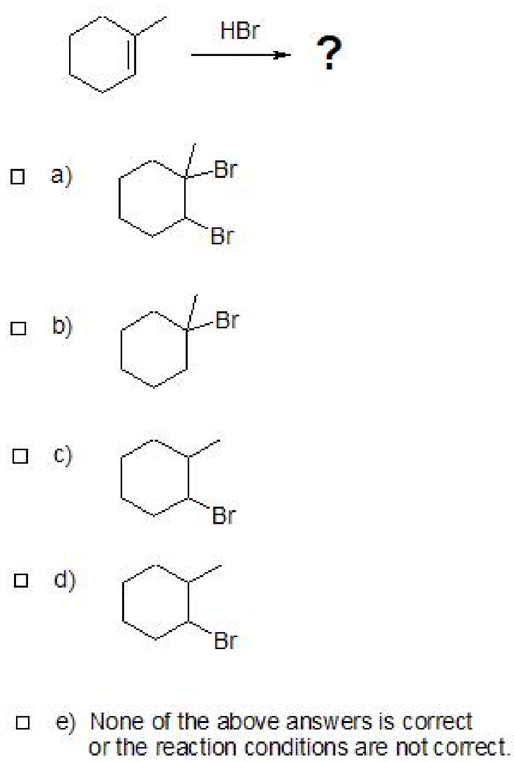
Discussion of the Method

In this regard, we thought that if this learning theory is correct, then we should be able to give the students enough different examples of how reactions work. By carefully keeping our database of questions in balance, we should be able to provide students with an opportunity to generate their own rationalizations of chemical behavior which will allow them to get to the correct predictive answer. For example, I have been using the addition reactions of alkenes as my examples for this conference. Within the dataset of questions, we carefully balance the number of groups that are electron-donating (*e.g.* , 1-methylcyclohexene), the number of symmetrical examples (*e.g.*, cyclohexene), and the number of groups that are electron-withdrawing (*e.g.* , 1-nitrocyclohexene).

Demonstration of the Method

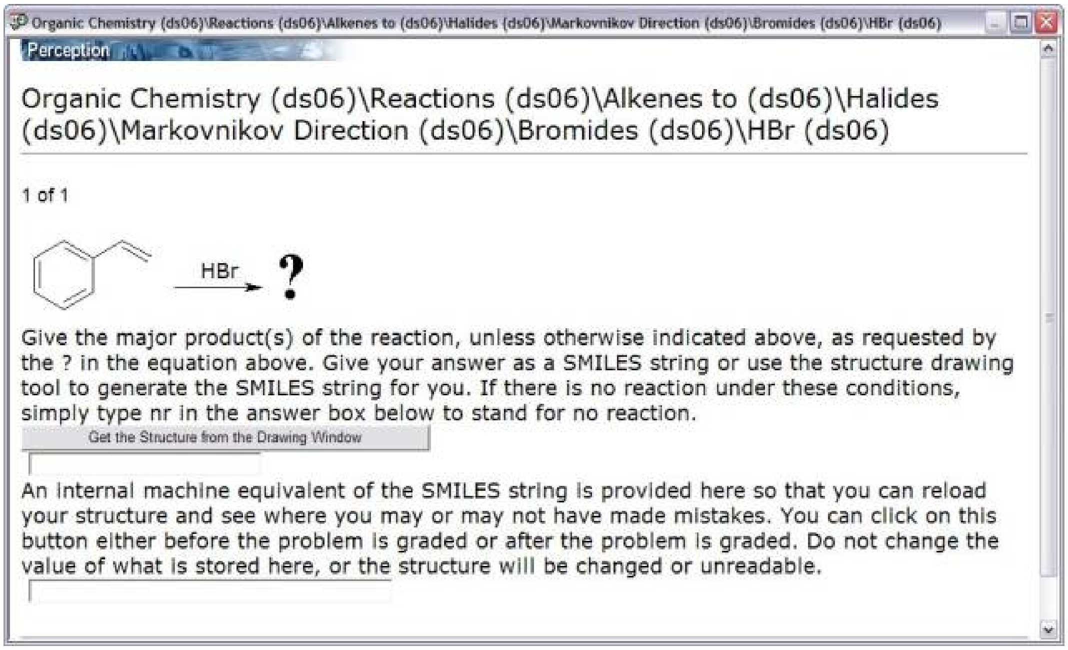
Before discussing the data, let's demo the system to see what happens and what the student sees. OK, let's demonstrate what we have just said. Open up another window on your computer and point your browser to http://www.welearnhorizon.com/q/perception.dll and log into the system with the account name CONFCHEM and password CONFCHEM. Try assignment CONFCHEM01 or CONFCHEM02 by clicking on the ï¿½Startï¿½ button. Think about the assignment now as providing examples of the different types of reactivity that one might envision in all of organic chemistry. Make sure to click on the submit button at the bottom of the page to observe the computer grading and explanations for wrong answers.

For those of you that are just printing out the paper, here's a sample of what you would see if you went to that assignment. In the typical Multiple-Select Question, the answer choices are always randomized, so that the students must identify the correct structure and not learn that answer b is always correct for this question.



Typical Multiple Select Question

Almost every multiple-select question has a corresponding draw-the-structure of the product question. Here is a sample of what that will look like and how it will act. One only needs to click on the button which says ï¿½Get the Structure from the Drawing Windowï¿½. This will activate a Java-window which will allow the structure to be drawn. This window is the now infamous JME window, generously donated to me by Peter Ertl of Novartis. After drawing the structure in the window, another button says submit the structure back to the main window and a SMILES string is placed into the answer box automatically.



Discussion of the data collected for this method

Despite in-class insistence that students should work through problems by thinking about the mechanisms of the reactions, students did, indeed, develop their own rationalizations of how these reactions work. In most cases, their solutions were nothing close to the correct mechanism and most often the students never develop the skills or recognition that the nitro-group is an electron-withdrawing group and will change the ï¿½Markovnikovï¿½-position of addition on the unsymmetrical alkene.

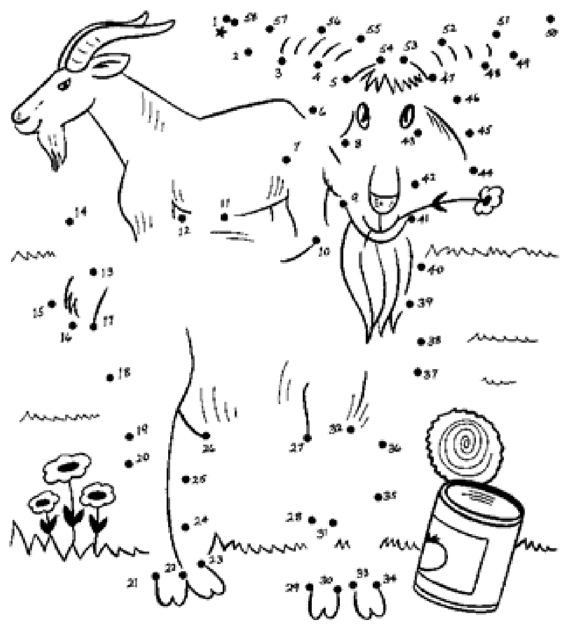
We have quantitative data on this method for learning reaction mechanisms. We asked students to draw standard organic reaction mechanisms on an exam. They were told in advance that they would have to draw a reaction mechanism. The major difference between this material is the fact that they had no ï¿½practiceï¿½. The mechanisms were evaluated using a rubric in which the correct product gave ca. 30% of the total score. For the ca. 200 students in my class, the sad fact is that the final average for this portion of the exam was only 30%. Of course, the top students were able to learn the mechanisms and draw them correctly. When we combine these students with the students who were not even able to get the product correctly, the overall average score was 30%. Similar results have been obtained at Richard Stockton College using similar methodology. Clearly, we are failing as teachers, when we use this methodology to teach reaction mechanisms.

Again, I hope that the believers in ï¿½Inquiry-based learningï¿½ will jump in here and tell me what I am doing wrong. This crusty old professor has become a firm anti-Inquiry-based learning professor.

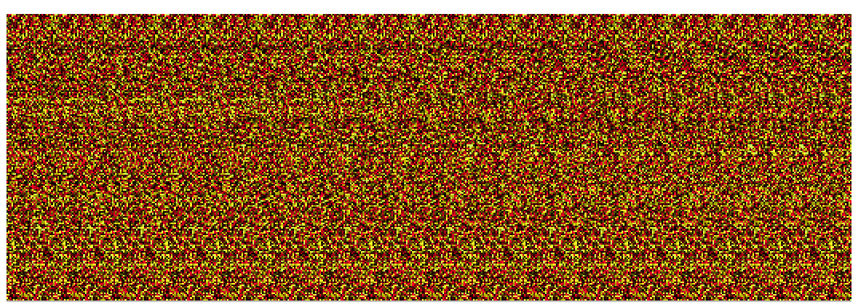
Method 2 ï¿½Arrange them in the correct orderï¿½

With the overwhelming failure of the method above, we have been considering alternative methods to try to train the students in thinking about the consequences of the individual steps of an organic reaction mechanism. At the moment, there are not ways to provide canonical versions of the ï¿½curved arrowsï¿½ of Ingold. This limits our possible methods for teaching mechanisms in a ï¿½Practice Makes Perfectï¿½ paradigm that would be suitable for CAI. However, we have organized several ï¿½connect-the-dotsï¿½ types of exercises as methods to try to get students to think their way through the mechanisms of reactions.

Ultimately, we will be asking the question of whether our "connect-the-dots" puzzles are too simplistic and provide no information to the student, as is our first ï¿½connect the dotsï¿½ puzzle that is shown here to the right.

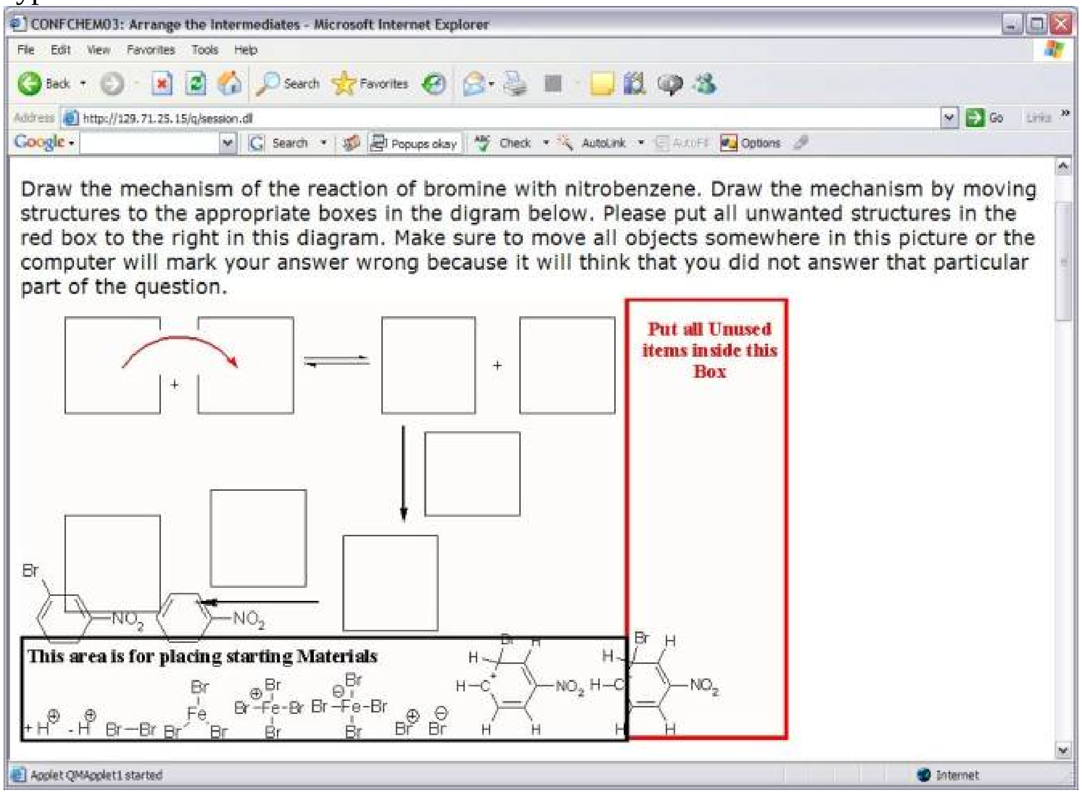


Alternatively, are the connect-the-dots puzzles too complex to be easily soluble (Note Example 2), and therefore bring their own problems with interpretation?.



Solve this seemingly random set of colored dots and see what this puzzle says

As our next attempt to teach organic reaction mechanisms to students in a way that they will be able to follow the flow of electrons in a chemical reaction, we have made ï¿½road mapsï¿½ showing all chemical intermediates in a chemical reaction between the starting materials and products. All significant potential energy intermediates that are normally hypothesized



in a reaction are given, along with some distracting species, are shown to the students, together with the reaction arrows to indicate the flow along the reaction coordinate. As the structures are given to the students, it is up to the students to place the structures correctly on the roadmap.

This is illustrated by the bromination of nitrobenzene. One can see the roadmap in the background, giving boxes which must be the location of some answer. What is less obvious from this picture is that the structures in the area marked ï¿½starting materialsï¿½ are movable on the computer screen. Therefore, one can move the structure to the appropriate place. There are additional distracting answers to make sure that the students understand concepts such as charge, etc. I really urge you to try the on-line demonstration in the next paragraph.

OK, let's demonstrate what we have just said. Point your browser tohttp://www.welearnhorizon.com/q/perception.dll and log into the system with the account name CONFCHEM and password CONFCHEM. Try assignment CONFCHEM03 by clicking on the ï¿½Startï¿½ button. Although this assignment does not cover the mechanism of the addition reactions to alkenes, hopefully, this will give an idea of what I mean for this methodology.

This method has the advantage that all of reaction pathway can be shown to the students, including the ï¿½curly arrowsï¿½ of electron movement. This method has the disadvantage that all intermediates are shown, and the students only need to come to a recognition level about what is happening in the reaction. A secondary disadvantage is a software quirk in that the answer choices cannot be randomized and remain in the order in which you placed them into the picture.

The major disadvantage is the time commitment to produce these questions, given the question manager from Questionmark, is huge.

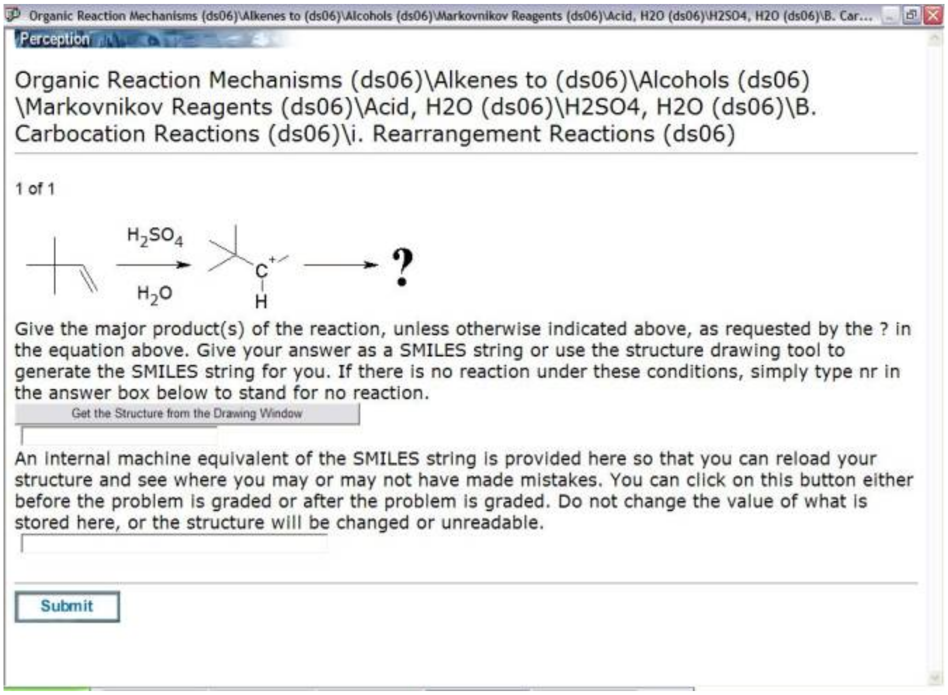
Unfortunately, I do not have quantitative information concerning the effectiveness of this method of teaching reaction mechanisms. Given the time commitment to generate a single example, we were unable to generate the 30-40 examples of each reaction type to properly illustrate this methodology. Seemingly, students were able to memorize the few examples that we were able to provide of this reaction type. As evidence of this phenomenon, students who completed this exercise on the computer routinely had higher scores than those students who had to write a mechanism in the traditional method on paper, even if it was the same reaction that they had seen in the computer exercises.

Method 3 ï¿½Fill-in-the-Blankï¿½ Method

A conceptually similar method for teaching reaction mechanisms is to draw a series of intermediates along a reaction coordinate and then to ask the students what comes next. This derives from our abilities to have students draw structures and then to have them evaluated by the computer.

OK, let's demonstrate what we have just said. Point your browser tohttp://www.welearnhorizon.com/q/perception.dll and log into the system with the account name CONFCHEM and password CONFCHEM. Try assignment CONFCHEM04 by clicking on the ï¿½Startï¿½ button. This assignment goes back to our addition reactions to alkenes and shows the strategy of the approach. This assignment contains 9 different examples of either halogen addition to alkenes or acid addition to alkenes. By allowing for random selection, each mechanism is shown at different stages in the reaction. Any help that the students get from other examples within the assignment is purely intentional, in that they are now realizing that the reactions of H 3 PO 4 and H 2 SO 4 are conceptually similar. In principle, students should be able to understand the ï¿½roadmapï¿½ of structures involved in the reaction mechanism after they have performed this assignment a few times. The only thing missing from this approach would be the ï¿½curved arrowsï¿½ showing the electron flow. Hopefully, the students would be able to make this connection by simply ï¿½connecting-the-dotsï¿½ at each reaction step.

This is shown in the following picture, where obviously, the next intermediate is that arising from rearrangement of the carbocation. Notice that the students must draw a structure and get no help from a list of answer choices.



We do have limited quantitative data on this approach. Given that the experiment has only been performed one time, and the selection of reaction intermediates was more limited (i.e., I rushed to finish off a number of examples for the purposes of this on-line conference), let's consider this to be preliminary data that may point to some trends. Further data collection may refine these trends and show more clearly the effectiveness of this approach.

The experiment is now similar to that described above. Students have access to this type of assignment prior to the in-class examination. During the examination, they will have to write a mechanism in the traditional fashion, using paper and pencil. The students were made aware of the fact that a mechanism or two would have to be drawn on the exam more than one week prior to the administration of the exam. The students were also made aware of the fact that the practice homework to assist them in studying the reaction mechanism was available on-line. A rubric was used to grade the mechanism, with approximately 1/3 of the value of the question being to achieve the final product. Based upon one exam, the mechanism score increased from the 30% being typical of the mechanism scores observed above to somewhere between 4.5 and 5. This indicates some improvement in student comprehension of reaction mechanisms, but there is still room for significant improvement.

Method 4 ï¿½Draw the Mechanismï¿½

So, at the same time, let's really tackle the big question. Back when I was a post-doc with Prof. Josef Michl at the University of Utah, there was a group slogan that was founded in Prof. Michl's superb scientific wisdom that ï¿½no experiment should not be performed, just because it is impossibleï¿½. Of course, this led to a diminished number of papers, particularly for the principal author of this paper, since the exceedingly new and exciting results were often found to be the results of an incomplete understanding of the instrumentation being used to measure the particular scientific phenomenon in question.

Application of Prof. Michl's philosophy to the problem of drawing mechanisms on-line suggests that there must, indeed, be a way to have the students draw reaction mechanisms, complete with curly arrows, and then to have the computer evaluate these mechanisms. Needless to say, it was my original hope to have the first preliminary working model ready for this on-line conference. However, the duties from teaching in this past semester, combined with other professional obligations, have kept the end goal tantalizingly close, yet still out of reach. Maybe before the end of the conference or at the BCCE in Purdue.

Methodology

Our full methodology appears in another manuscript. Without going into full detail, I'll just say that the SMIRKS format fits our needs best, in that a canonical representation of a reaction is given in an ASCII code that can be compared to something that is stored in a computer database. The missing component is the ï¿½curved arrowsï¿½ to show electron flow. If one realizes that the curved arrows can be represented in a canonical form that can be separated from the ASCII string, then one can develop an ï¿½Extended SMIRKS stringï¿½ which can be used for comparison of what someone else has drawn. As I said, I will provide examples of this strategy at the BCCE meeting.

With such a tool in hand, there is no question that we will be able to represent simple reactions (i.e., one or two steps) and to be able to compare these to strings that are already stored in the database. The long-term challenge will be to evaluate multi-step reactions where partial credit will be a must in the overall grading scheme.

Summary and Conclusions

ï¿½Practice Makes Perfectï¿½ drawing of organic reaction mechanisms is possible and will soon be available. It will be interesting to see if the enhanced capabilities of the students to draw these reactions will be accompanied by an increased understanding of the underlying chemical phenomena. Obviously, that will be the subject of a paper to be written in the near future.

ACKNOWLEDGEMENT

I thank the many students who have come through my class and helped me to work out the problems in the WE_LEARN system. I thank Horizon Learning Solutions for providing support for this work.

ACEorganic, http://www.prenhall.com/aceorganic/

A similar web-site concerning Discovery-based learning can be found athttp://ws.cc.stonybrook.edu/Reinventioncenter/resinquiry.html

Holliday, G.L.; Murray-Rust, P.; Rzepa, H.S. ï¿½Chemical Markup, XML, and the World Wide Web. 6. CMLReact, an XML Vocabulary for Chemical Reactions, *Journal of Chemical Information and Modeling*, **2006,** *46*, 145-157.

Horizon Learning Solutions, www.we-learn-horizon.com.

Ingold, C.K. "*Structure and Mechanism in Organic Chemistry*, 2nd Edition, Cornell University Press, Ithaca, NY, **1969**.

An interesting web-site concerning Inquiry-based learning can be found athttp://www.thirteen.org/edonline/concept2class/inquiry/index.html.

Professor Michl is currently in residence at the University of Colorado, Boulder.

OWL, http://owl1.thomsonlearning.com/aboutOwl.html

Penn, J.H.; Nedeff, V.M; and Gozdzik, G. "Organic Chemistry and the Internet: A New Approach to Homework and Testing Using the WE_LEARN System", *Journal of Chemical Education*, **2000**, *77 (2)*, 227-31 .

Penn, J.H.; Steinbeck, C. *Journal of Chemical Information and Modeling*, submitted.

Webassign, www.webassign.com