

# ChemPad: A Pedagogical Tool For Exploring Handwritten Organic Molecules

Dana Tenneson

Brown University Providence, RI 02912

## Abstract

Molecules are inherently three-dimensional objects that are represented by chemists on paper and classroom blackboards by a system of two-dimensional notations. A major hurdle for aspiring chemists in college courses is to be able to visualize the actual 3D structure from the dimensionally-deficient 2D drawings. ChemPad, a new Tablet PC application with a pedagogical focus was designed and implemented to generate 3D molecular structures from hand drawn digital ink. This paper addresses the ChemPad system, the algorithms involved in interpreting 3D molecular structure from 2D ink, and the pedagogical design of the software.

## 1 Introduction

Organic Chemistry is a hard course. Early in their college studies, students who wish to pursue a career in medicine must have a solid understanding of organic chemistry and, perhaps more importantly to the students, a good grade in Organic Chemistry. The difficulty of the material prevents many students from continuing with this career path. One of the major difficulties for students in Organic Chemistry is understanding the three-dimensional nature of molecules. Students usually have no background in three-dimensional visualization and have great difficulty converting between the two-dimensional drawings used in text books and on classroom blackboards to represent molecules and their three-dimensional structures. Without this understanding, to survive the course, students must memorize a large vocabulary of molecules and rules to fake an understanding of the three-dimensional structures. Although this is possible for some, good students and good chemists tend to learn to visualize the molecules in three dimensions and apply a much simpler set of rules to these visualizations.

ChemPad, a Tablet PC application shown in Figure 1, assists students learning to visualize molecules from standard two-dimensional drawings. ChemPad has been designed with pedagogy as its focus and introductory Organic Chemistry students as its target audience so that it is much more of a teaching and learning tool than general molecule modelling software. In many ways, it acts as a digital substitute for physical ball-and-stick modelling kits. Students can build the same structures they would with the modelling kit and see the same 3D features of the molecules. However, students using ChemPad can construct these models much quicker than with a physical model and receive interactive feedback about

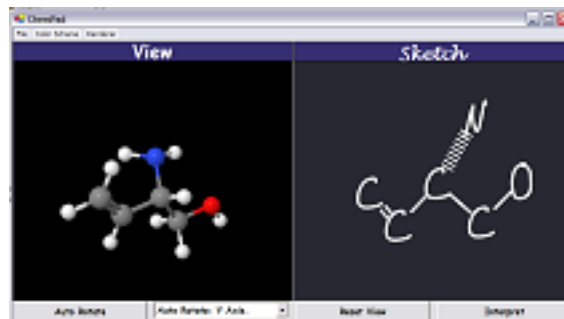


Figure 1 The ChemPad application.

their 3D intuitions which would otherwise be unavailable without a TA or a professor available.

## 2 Using ChemPad

Using ChemPad consists of drawing a molecule in the Sketch panel and observing the three-dimensional structure in the View panel. Instead of selecting atoms and bonds from toolbars and menus the way a chemist does in professional modelling software, students use the Tablet PC stylus to draw a molecule in digital ink on the tablet that looks very similar to the same drawing they would make on an exam or on paper when talking to another chemist about the molecule. These drawings consist of single-stroke symbols representing atoms, bonds, and special pedagogical functions. For example, a student drawing ethanol would draw two C's and an O to represent the two carbons and one oxygen in the molecule. The atoms are then connected with covalent bonds by drawing straight lines between letters. As the atoms and bonds are being drawn, the View panel displays the balls and sticks of the atoms and bonds being recognized and the Sketch panel is incrementally prettifying the user's input. By tapping the "Interpret" button, the student indicates that the drawing is complete and ChemPad presents the user with a 3D scene showing the ethanol molecule with implicit hydrogens attached and the molecule oriented in the configuration approximated by the user's drawing.

A more interesting example of a student exploring a 3D structure in ChemPad is the geometric difference between isobutyl alcohol and acetone (Figure 2). Isobutyl alcohol consists of three carbons and an oxygen each connected directly to a fourth carbon. In this configuration, the central carbon assumes a tetrahedral structure with its connected atoms. By removing an outer carbon and moving the empty bond to the oxygen, acetone is created. In acetone, the central carbon assumes a trigonal planar structure<sup>1</sup>. A student using ChemPad could explore this difference in the following set of steps. First, the student draws the molecule in much the same way that ethanol was drawn, except with two additional carbons. Upon interpreting this molecule, ChemPad will remind the student that

<sup>1</sup>A trigonal planar structure has the three neighbors share a plane with the central atom and the angles between neighbors are 120°

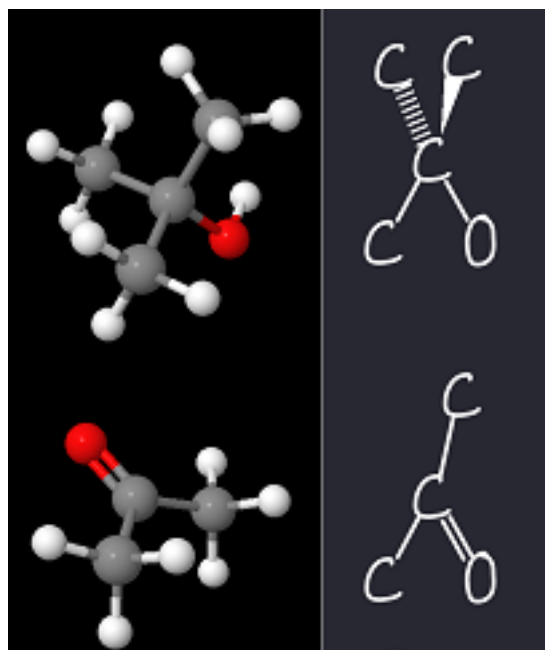


Figure 2 Isobutyl Alcohol and Acetone

this structure is extra-planar and that the student must specify the bond directions with standard wedge and dash notation. The student then scratches out two of the bonds to erase them and replaces them with single-stroke variations of the solid triangle and dashed lines used by chemists to specify three-dimensional structure. Interpreting will now display the 3D structure the student has drawn of isobutyl alcohol. Erasing a carbon and its bond and drawing a new bond to the oxygen will create acetone. Now, the student will be informed that the drawing is incorrectly using wedge and dash notation because the central carbon is now planar. Replacing the wedge and dash with normal bond strokes will allow the interpreting to continue.

## 2.1 Stereochemistry

Understanding that molecules are three-dimensional is a first step towards solving the difficult problem of naming chiral molecules. Chirality is a property of molecules where a carbon atom is bonded to four distinctly different structures. Because there are exactly two different rotationally equivalent ways to attach four structures to a tetrahedron, two different molecules exist with the same chemical formula and connectivity. These mirror image molecules are called enantiomers and can have important differences in their chemical properties. Chemists therefore label one version of the molecule as R (Rectus) and one version as S (Sinister). For a student to determine which version of a molecule is displayed on paper, the student must first order the four branches of the molecule by the Cahn-Ingold-Prelog (CIP) rule set. Next, the student must visualize the lowest priority branch facing away from the student's point of view and observe whether the other three branches are ordered in clockwise or counter-clockwise fashion. Without the ability to visualize the molecule, a student must perform a complicated set of tricks and rules to convert a molecule drawing to an orientation that is 'equivalent' to one the student knows to be R or S.

Using ChemPad, a student can be taken through the steps of the visualization and gain a deeper understanding of the structure. By drawing an R or S next to a stereocenter (a carbon with four differ-

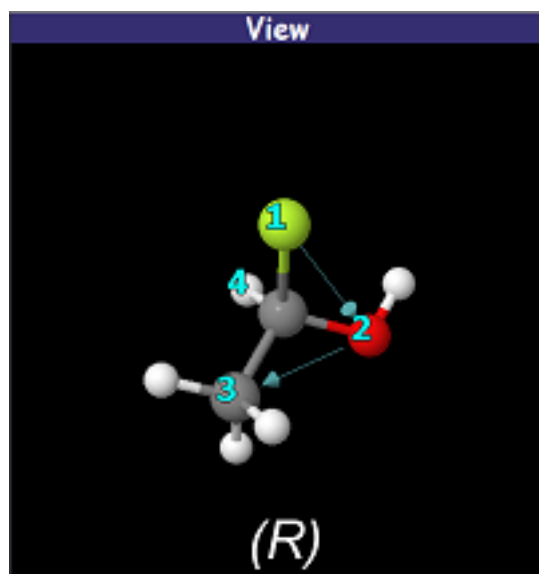


Figure 3 The stereochemistry visualization

ent branches), the student will receive feedback in the View panel showing the CIP priorities drawn over each branch of the molecule and arrows connecting the 1, 2, and 3 branches. When the user rotates the lowest priority branch towards the back, a dialog is displayed giving the user one more chance to reevaluate their naming and then the proper chirality is displayed as shown in Figure 3. In addition to enantiomers, ChemPad also contains a visualization system for a student to explore the naming of cis and trans (Z/E) stereoisomers, another form of stereochemistry based on the rigidity of double-bonds.

## 3 Recognizing Chemistry Handwriting

Figure 4 gives an overview of how handwritten data is processed into 3D graphics by ChemPad. Before any visualization can occur, ChemPad must generate a 3D molecular structure from the user's digital ink drawing. The first step of ChemPad's process is the recognition of the components (tokens) of the user's drawn molecule. The single-stroke tokens input by the user are processed by a version of the fluid inking [8] library that has been adapted to recognize chemistry specific tokens. Fluid inking uses the Windows XP: Tablet Edition recognizer to identify letters and a grammar of stroke components to recognize other symbols. The recognized atoms and bonds are stored in a list data structure with methods to access the data as a graph<sup>2</sup>. This allows the renderer, a .NET port of the open source molecule viewer Jmol<sup>3</sup>, to quickly get a list of 3D locations and the chemistry methods to process the data with the ease of a graph structure. To make these methods faster, some information about how the bonds connect to atoms is cached after it is computed.

Once ChemPad recognizes inked gestures as Atoms or Bonds, a data object of either the Atom or Bond type is created for the ink. In addition to the ink data, pixel coordinates of the centers of the atoms and the end points of the bonds are stored and unique identifiers are generated. The Atom or Bond is then stored in the Molecule data structure which contains a list of each type of data in the molecule

<sup>2</sup>These methods include checking for atoms adjacent to atoms, bonds adjacent to atoms, and atoms adjacent to bonds.

<sup>3</sup><http://jmol.sourceforge.net/>

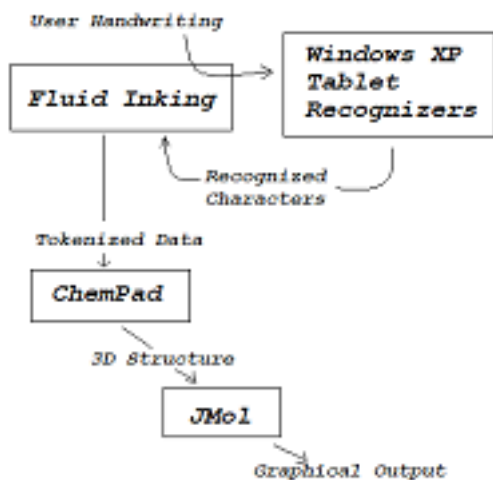


Figure 4 The flow of data through ChemPad.

because this format is easy to pass on to the Jmol visualizer. Additionally, the Molecule structure contains lists of temporary bonds and temporary atoms. These lists contain implicit molecular structure not represented by drawn ink. In a neutrally charged molecule, each atom bonds to a specific number of other atoms based on its atomic number. Drawings of organic chemistry molecules usually do not contain all the bonds and atoms present in the molecule because a large number of hydrogen atoms is common and hydrogen has only one bond available. Therefore, hydrogens can be implied in the drawings and ChemPad must add these atoms into the molecule to be processed. Implicit hydrogens and bonds are kept in these separate lists from the explicit atoms so that they can be removed from the molecule easily when the user changes the drawing.

When a user taps “Interpret” in ChemPad, the software needs to map the drawn image to an appropriate 3D model of a molecule. Atoms in molecules exert forces upon each other constantly causing individual atoms to take various positions in 3D space. These combinations of 3D positions, or conformations, change frequently with conformations of relatively low energy being much more likely than higher energy conformations. Due to the many degrees of freedom (DOF) available to a molecule such as rotations around bonds (torsional DOF), bond lengths, and bond angles, a great deal of conformations are possible. Therefore, it’s not surprising that calculating appropriate conformations is not a new problem for computers. [2] and [9] show examples of how to explore molecules for low energy conformations using pseudo-exhaustive and stochastic search algorithms, respectively. These solutions, while near optimal for computationally determining molecular properties, are based upon atom connection tables and are therefore oblivious as to any conformational intent represented in a drawing. While professional chemists may want the perfectly accurate conformation of a molecule and be willing to wait for it, a student chemist should be given a quick, reasonable conformation which accurately reflects conformation cues drawn into the molecule. Besides the speed benefit, this allows the student to quickly recognize obvious conformation mistakes made in the drawing.

To achieve a faster algorithm that incorporates the user’s drawing of the molecule, we first note that bond lengths and angles tend to be determined largely by localized information and possible values are discretized. We can then think of building our 3D molecule

### 3.1 Mapping From 2D to 3D

Algorithm 3.1: INTERPRETMOLECULE(*molecule*)

```

procedure CONNECT(connecteeAtom, connectorAtom)
{
  connecteeAtom.Template ←
    StructureTemplate(connecteeAtom)
  connectionVector ← OptimalConnection(connectorAtom,
    connecteeAtom)
  rotation ← RotateToMatch(−connectionVector,
    connecteeAtom.Template.PrimaryConnection)
  connecteeAtom.Template.Rotate(rotation)
  connecteeAtom.3DLocation ←
    connectorAtom.3DLocation + (connectionVector*
    BondLength(connectorAtom, connecteeAtom))
}

procedure OPTIMALCONNECTION(connector, connectee)
{
  if IsImplicitHydrogen(connectee)
  then {
    connectorAtom.Template.
    RotateAroundParentAxis(
    DegreeForEclipsedConformation())
  }
  else if { BondBetween(connector, connectee).Type ==
    Planar and connector.3DLocation == 0
  }
  {
    connectionVector ←
      connector.Template.PlanarConnection();
    {
      TurnDirection(connector.Parent.
      3DLocation, connector.3DLocation,
      connector.3DLocation + connectionVector)
    }
    if {
      != TurnDirection(connector.parent.
      2DLocation, connector.2DLocation,
      connectee.2DLocation)
    }
    then {
      connector.Template.
      RotateAroundParentAxis(180)
      connection ← connector.Template.
      PlanarConnector
    }
    else comment: Do similar thing for other bonds
  }
  return (connectionVector)
}

procedure ROTATETOMATCH(goalVector, startVector)
{
  axis ← Cross(goalVector, startVector)
  angle ← Dot(goal, start)
  return (axis, angle)
}

main
{
  comment: Preprocessing
  if ErrorsExist(molecule)
  then fail
  AddImplicitHydrogens(molecule)
  comment: Start the molecule

  startAtom ← MostStereochemistryCuesAtom(molecule)
  startAtom.3DLocation ← (0, 0, 0)
  neighborAtom ← AtomAcrossPlanarBond(molecule,
    startAtom)
  penSlope ← startAtom.InkLocation
  penSlope.Normalize()
  penSlope.Multiply(BondLength(startAtom, neighborAtom))
  neighborAtom.3DLocation ← (penSlope.X, penSlope.Y, 0)
  Connect(startAtom, neighborAtom)
  Connect(neighborAtom, startAtom)
  comment: Grow outward from the center

  repeat
  for each unconnectedAtom ∈ AtomInPriorityOrder
  do Connect(unconnectedAtom, ConnectedNeighbor(a))
  until AllConnected(molecule)
}
  
```

Figure 5 System of Mapping a 2D drawing to a 3D Molecule.

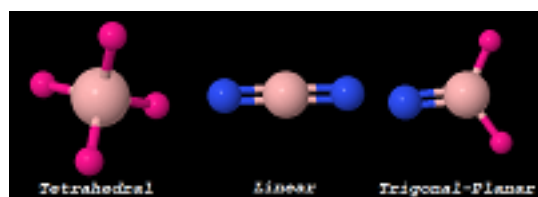


Figure 6 Examples of common atom templates.

with the metaphor of building a molecule using a plastic ball-and-stick molecule kit. We know exactly which “pieces” to use based on the connection table of the molecule. This reduces the difficulty of finding a conformation to one of finding the torsional angles, a not uncommon technique in Computational Geometry solutions to conformation finding [9]. We additionally note that chemists (student and professional alike) take care with drawing their molecules to try to create a pseudo-perspective drawing of the 3D structure. So, we can treat the drawing as a “photo” of the actual 3D molecule sitting in its proper conformation. This combination of a skeleton of parts with many DOF and a 2D “photo” of the 3D structure, leads to the conformation problem requiring more of a Computer Vision solution in the same category as human position tracking than a Computational Geometry solution. This choice of solution paradigm intuitively makes sense because of ChemPad’s goal of helping students *visualize* the molecules that are drawn. Techniques such as [5] and [10] solve this kind of problem for human motion and pose and it would not be difficult to generate the model-based images these algorithms need from our chemical models. However, human positioning algorithms are concerned with dealing with a slightly different class of obstacles and aids in the photographs. Obstacles in photographic deficiencies of humans are obscured body parts and fuzzy images rather than excluded hydrogens and completely wrong projected locations. A major aide in keyframed tracking is a very close model pose that is considered correct which is not to be found in a molecule interpretation system. On the chemistry side, molecule drawings contain bond cues for depth that aren’t found in human photos and an algorithm for conformation determination should exploit this information. Hence, a different system is needed.

ChemPad’s algorithm for interpreting a molecule (shown in Figure 5) is deterministic and rule-based. It starts with a couple pre-processing steps before 3D positioning is attempted. The first is a quick error-checking step to make sure that the molecule is connected and that no atom violates its valency restrictions. The second is to generate a list of implied hydrogens and the bonds that connect them to the explicit atoms. Finally, ChemPad attempts to fit the molecule to the tokenized drawing data so that a rough match is found. Organic chemistry molecules contain carbon chains at their core and ChemPad starts the process by identifying the atom (normally carbon) with the most different connections because that atom has the most spatial information. That atom becomes the origin of the scene and its connection to two of its neighbors defines the XY plane of the 3D scene. A three-dimensional template, as shown in Figure 6, is assigned to each atom based on its atomic number, bond orders, and neighboring bond orders. These templates are a list of vectors representing the relative directions where other atoms can be bonded. Distances between atoms are assigned based on the atomic number of the atoms and the order of the bond. Using these constraints and the XY plane constraint, the first two atoms can be positioned in 3D such that the slope of their connecting bond in the XY plane is the same as the slope of the drawn bond in the Sketch panel. From here, additional atoms are attached to the existing 3D structure. When an atom is attached, the ‘parent’ atom’s template is queried to find the most appropriate connection point.

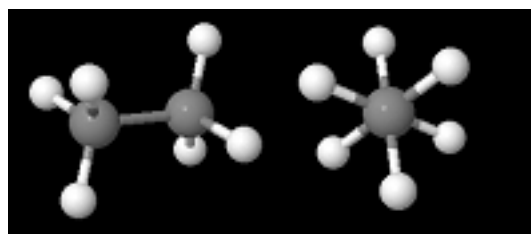


Figure 7 Two views of adjacent tetrahedral carbons.

Which connection point wins depends on the orientation of the template, the bond notation used (line, wedge, or dash) and the angle at which the bond was drawn. In some cases, the template must be rotated around the axis formed with its parent atom to achieve the best fit. Because implicit hydrogens contain the least structural information, they they require additional heuristics for coaxing the templates into reasonable chemistry conformations and are added last so that they cannot override decisions made by explicit atoms. One important heuristic is that adjacent tetrahedral atoms, such as those in Figure 7, when viewed down the axis of their bond, tend to have their immediate neighbors evenly spaced at  $60^\circ$  angles because this conformation requires the least energy to maintain.

The simplistic solution above does a very good job at its goal of orienting molecules typically found in introductory Organic Chemistry<sup>4</sup> into a low energy conformation without performing computationally intense functions. Computational complexity is polynomial instead of exponential or stochastically reduced exponential and runs in real time. Although computationally imperfect, these conformations are educationally *better* since they adhere to the rules of thumb chemists use to visualize and describe molecules. Shortcomings in the conformations become most apparent when there is a lack of handwriting cues that can be drawn upon to determine the intended conformation. For example, atoms at a distance at least 2 from the main chain of the molecule are less likely to be in a low energy conformation due to the loss of standard bond line indicating a planar bond. Similarly, the use of non-standard drawing notations<sup>5</sup> reduce the handwriting cues further and are therefore rejected by ChemPad as being improper.

#### 4 Determining CIP Ordering

The CIP ordering system orders constituent groups from highest to lowest when determining stereochemistry naming. Chemists describe CIP in vague terms such as those in Figure 9 which are not well spelled out from an algorithmic standpoint.

A human being performing this operation in Organic Chemistry tends to use a little intuition to solve the problem. Given four different constituent groupings, try to pairwise compare the branches to find out which priorities to assign. Often two of the branches are different non-carbons (a hydrogen and an oxygen for example) which can quickly be eyeballed for absolute priority, by only looking at the first atomic weight. When two of the branches are carbon branches, the human eye can usually note the depth of the first difference in the chain and then deduce which is the higher priority.

<sup>4</sup>Molecules with less than 30 atoms.

<sup>5</sup>“Chemist hacks” such as a main-chain carbon with two wedges and two dashes which are accepted as technically correct, but frowned upon for being improper.

**Algorithm 4.1:** SORTEDPRIORITIES(*branchesList*)

```

procedure CIPSCORE(constituentBranch)
  if IsAnAtom(constituentBranch)
  then {
    string ← FormatToThreeDigits(constituentBranch.
      AtomicNumber)
    return (string)
  }
  neighborValuesList = newList()
  subscoresList = newList()
  for each childBranch ∈ constituentBranch. Children
  do {
    neighborValuesList.Add(
      FormatToThreeDigits(constituentBranch.
        Root. AtomicNumber))
    subscoresList.Add(CIPScore(childBranch))
  }
  else {
    neighborValuesList.SortAlphabetically()
    subscoresList.SortAlphabetically()
    score = newString()
    for each string ∈ neighborValuesString
    do score+ = string
    for each string ∈ subscoresString
    do score+ = string
    return (score)
  }

main
  scoresList = newList()
  for each branch ∈ branchesList
  do scoresList.Add(CIPScore(branch))
  scoresList.SortAlphabetically()
  return (scoresList)

```

**Figure 8** Assigning CIP priorities using string comparison

1. Higher atomic number takes precedence over lower.
2. When two atoms directly attached to the stereogenic center are identical, compare the atoms attached to these two on the basis of their atomic numbers. Work outward from the point of attachment and evaluate substituent atoms one by one. Precedence is determined at the first point of difference.
3. The difference is determined by the substituent of the highest atomic number and is not additive if there is more than one substituent.
4. Where there is a double or triple bond, both atoms are considered to be duplicated or triplicated.

**Figure 9** A chemist's description of the CIP rules. Taken from [1].

O-H 008001001

C-H 006001001001001001001

C-C-H 006006001001006001001001001001001001001

**Figure 10** Examples of constituent branches and the generated CIP score strings

For the computer to determine the priorities of the branches, it's a bit more difficult. When there is a difference one or two atoms away from a stereocenter, it is not difficult to quickly determine the priority orderings, however as the constituent branches, the algorithm needs to branch down the higher priority subbranch first, which the computer doesn't intuitively know how to do. Although the CIP rules have been around since the mid 60's, the importance of absolute stereochemistry determination has only been emphasized for the last decade [3] and CIP algorithms are sparse. [6] uses a global partitioning system to prioritize every atom in the molecule independent of the queried stereocenter. This solution allows for quick computation of the chirality of multiple stereocenters in a molecule, but is not intuitive from the perspective of the CIP rules and therefore hard to use as a basis for an algorithm animation (a requested, but as of yet unimplemented important feature). A straightforward implementation if the CIP rules calls for recursion, but a recursive order in which to traverse the molecule is not apparent.

ChemPad uses a novel application of string comparison in a recursive algorithm to compute an individual stereocenter's constituent orders on the same order of computation complexity as [6], but which also notes the path traversed to the point of first difference for an algorithm animation. The algorithm is detailed in Figure 8. Each constituent is assigned a numerical value that is generated recursively and stored as a string. The goal is to create strings of digits that when sorted into alphabetical order will correctly prioritize the constituent groups. Here, the recursive base case is that of a terminal atom which produces its atomic value as a 3-digit number. Oxygen is 008, for example. If the atom is non terminal, the recursive case is to generate a string with three parts. The first part is its atomic number. The second part is the ordered atomic numbers of its neighbors. The third part is the ordered recursive strings for its neighbors. Ordering here occurs by using standard string comparison. Because we are using 3-digit numbers and the periodic table contains no elements with a 4+-digit atomic number, alphabetic ordering will sort the numbers into numeric ordering.<sup>6</sup> Furthermore, alphabetic ordering will give us the correct priority order because alphabetic ordering looks for the first point of difference scanning left to right and the algorithm is placing numbers in the order they are to be considered for CIP. Consider the branches in Figure 10 and their numerical values. Here the Oxygen gets the highest priority, the 2-carbon chain the next highest priority, and the methyl group the lowest priority. The differences come up quickly in the string comparison. Note that using numerical comparison here would generate incorrect answers because the longer branches would get higher priority. In fact, the format of the strings may diverge wildly after the point of first difference with numbers no longer meaning the same thing in parallel.<sup>7</sup>

In the more complicated example of Figure 11, the branching carbons make it difficult to know where the algorithm should pro-

<sup>6</sup>An extension to the system to account for differences in isotopes is a straightforward addition of digits to each atom's value.

<sup>7</sup>For example, the third number in the O-H branch refers to the local atomic number of a hydrogen. The third number in the C-H branch refers to the atomic number of a hydrogen adjacent to the local carbon.

$C-C-O-H$  00600600600100600800100100800100100100...

$C-C-O-Br$  00600600600100600800100100803503500100...

**Figure 11** Examples of branching constituent branches and the generated CIP score strings

ceed. However, by doing sorting of the substrings as recursive generation is occurring, the longer branches are placed first in each string and the Bromine vs. Hydrogen difference comes first in the strings to be compared.

Once we have a 3D structure, determining the Z/E or R/S stereochemistry of molecules becomes possible. Given a CIP ordering of constituents, Z/E stereochemistry is determined by calculating the turn directions (left or right) of the angles formed by the central carbons and their higher priority constituents. If the directions are the same, the molecule is trans (E) and if they are different, the molecule is cis (Z). Similarly, for R/S stereochemistry, the turn direction of the angle formed by the projections of the high priority constituents onto the XY plane when the low priority constituent is on the Z axis determines the chirality. A left turn from 1 to 2 to 3 is S and a right turn is R.

## 5 Pedagogical Methodology

Once the software can determine the 3D structure and chirality of software, helping teach students do the same can be addressed. For a student with a 3D model of a molecule, the task of determining chirality becomes the simpler one of assigning a CIP ordering and remembering that R is clockwise. Students who cannot accomplish this task on paper but can using ChemPad are working within Vygotsky's Zone of Proximal Development (ZPD) [11]. The visualization assistance ChemPad offers *scaffolds* the student's learning of the material. By receiving assistance to reach correct answers, students develop the skills to reach correct answers without assistance.

Several Organic Chemistry course web pages provide students with the opportunity to manipulate pre-built 3D molecules with browser plugins such as Chime<sup>8</sup>. These tools allow exploration of similar visualizations to what is provided by the 3D component ChemPad. However, they lack the capability for students to make modifications to the molecule and experiment with it. The value of giving students the tools to raise and answer their own questions instead of following a linear curriculum, such as that found on a web page, has been espoused in Duckworth's concept of "Having Wonderful Ideas" [4]. A typical user with a basic knowledge of chemistry can learn the fundamentals of ChemPad in only five minutes. From there, the student has a great deal of freedom within ChemPad to experiment with generating different molecules, exploring the differences between them, and generating their own "wonderful ideas" for understanding chemistry.

## 6 Use in the Classroom

ChemPad is currently being used as part of Brown University's Organic Chemistry course. Professor Matthew Zimmt is using the program in class to demonstrate the process of determining chirality and lab sections are being held twice a week where students can use the software to explore and work through example problems in their own time. Feedback from students and Professor Zimmt has been

- "It's cool how you can draw a molecule in 2D and have it represented in 3D."
- "I like being able to flip the molecules to see it in the right orientation."
- "Having a 3-D representation of the molecule for learning purposes is extremely helpful...much better than on paper."
- "I like that it shows the C.I.P. numbering so that you can see if it is clockwise or counter-clockwise very easily."
- "It helped me understand rotation better. It also helped me visualize which atoms were stereocenters."

**Figure 12** Student feedback regarding ChemPad

generally positive and has exposed to the research team the depth of the problems these students have visualizing three-dimensional structures. Figure 12 shows quotes from typical user responses on a survey imbedded in the software. Additionally, a study is currently underway to use quiz scores to determine the efficacy of ChemPad. Students who did poorly on quiz questions that specifically require three-dimensional visualization skills were sent specific invitations to attend the lab sections. 3D questions on later quizzes will be identified so that two comparisons of performance on this class of questions can be analyzed. This first is performance before and after using ChemPad. The second is performance of students who received invitations and attended the labs to those who received the invitations but chose other methods of studying.

## 7 Future Work

There are many directions in which ChemPad can continue to grow. Interpreting molecules from entire drawings that are drawn exactly the way they are in chemistry instead of in our single-stroke language would be a challenging and important step towards having a general chemistry interface from digital ink. Adding additional visualization tools for students and molecules accepted as valid for professionals would expand the audience who could use ChemPad on a regular basis. Professor Zimmt and others outside of Brown have expressed that integrating ChemPad with the math sandbox in MathPad<sup>2</sup> [7] would make a tool useful for even more tasks. With this functionality, ChemPad could become an integral part of the entire first year of chemistry studies.

## 8 Acknowledgements

ChemPad would not have been possible without the profound efforts of Professor Matthew Zimmt who provided not only a class of students, but endless hours of software testing and chemistry tutoring for the author. Deploying a successfully-debugged version of ChemPad with high quality 2D and 3D interfaces for the Organic Chemistry classroom was made possible by the extra programming efforts of Sascha Becker, Loring Holden, and Robert Zeleznik of the Brown University Graphics Group under the direction of Professor Andries van Dam.

The also author would like to thank Atlantic Philanthropies for their grant to support ChemPad development and Hewlett-Packard for their donation of Tablet PC's to create the ChemPad lab.

<sup>8</sup><http://www.mdl.com/products/framework/chime/>

## 9 References

### References

- [1] Stereochemistry (substituted butanes, cyclopentanes, and cyclohexanes). <http://orgchem.colorado.edu/courses/ModelExStereo.pdf>.
- [2] S. Kahn A. Smellie and S. Teig. Analysis of conformational coverage. 1. validation and estimation of coverage. *Journal of Chemical Information and Computer Sciences*, 35:285–294, 1995.
- [3] T. Cieplak and J.L. Wisniewski. A new effective algorithm for the unambiguous identification of the stereochemical characteristics of compounds during their registration in databases. *Molecules*, 6:915–926, 2001.
- [4] Eleanor Duckworth. “*The Having of Wonderful Ideas*” & *Other Essays on Teaching & Learning*. Teachers College Press, 1987.
- [5] I. Reid J. Deutscher, A. Blake. Articulated body motion capture by annealed particle filtering. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1144–1149, 2000.
- [6] P. Labute. An efficient algorithm for the determination of topological chirality. *Journal of the Chemical Computing Group*.
- [7] J. LaViola and R. Zeleznik. Mathpad<sup>2</sup>: A system for the creation and exploration of mathematical sketches. In *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2004)*, volume 23, pages 432–440, August 2004.
- [8] Modeless Combination Of. Fluid inking: Using punctuation to allow.
- [9] L. Kavraki et. al. P. Finn, D. Halperin. Geometric manipulation of flexible ligands. In *Proceedings of the First ACM Workshop on Applied Computational Geometry*, 1996.
- [10] Cristian Sminchisescu and Alexandru Telea. Human pose estimation from silhouettes. a consistent approach using distance level sets. In *WSCG International Conference on Computer Graphics, Visualization and Computer Vision*, 2002.
- [11] L. S. Vygotsky. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, 1980.