COMPUTER SOFTWARE REVIEWS

Kekulé for Windows: The Complete Structure Input System†,‡,⊥

Anne Rogers§

Chemical Registry, The Dow Chemical Company, Midland, Michigan 48667

Received July 8, 1994

INSTALLATION AND SETUP

Installation of the Kekulé program is straightforward, using (mostly) Windows (a product of Microsoft Corp.) operating system standard processes. Instead of the usual status bar to tell you how far along the installation process is, you get a slowly filling circle in a benzene ring. Otherwise, it looked like all standard Windows interface installation procedures are followed. We have two copies of the Kekulé program and wanted to install it for two concurrent users on our PC local area network. Like many of the better vendors today, PSI did not have a problem licensing the software this way. With a little help from Tech Support (available and helpful on Christmas Eve day) on a few minor network configuration issues, the software was quite easy to install and very well behaved on the LAN. A nice installation feature is the ability to locate the software easily in a drive/directory of your choosing, something not all structure drawing packages let you do. The installation program also gives you the option to add the Kekulé icons to an existing Windows desktop group instead of automatically creating an additional group.

The Kekulé program was ready-to-run after installation. The only modification needed for complete functionality was to specify scanner details. A wide variety of defaults are customizable, including colors, selector tool (frame or lasso), fonts (atom, subscript, text), bond styles (type, width, length, angles), contents of the tool bar, group formulas on the drop down menu, and the contents of the template menu. Export and save file defaults can also be modified. If there is a default setting you do not like, chance are you can change it. You can save changes in the default configuration file or have different configuration files for different situations. The program prompts you on exit to save any configuration changes you have made.

DRAWING

The program's drawing features have been dramatically improved with the 2.0 release. When I started reviewing version 1.11, my impression was that while the Kekulé system had adequate drawing features, its main attractions were its structure scanning capabilities and advanced "chemical intelligence". It is now a full-featured structure drawing package, with quick and intuitive drawing tools. A wide variety of easily accessible templates is provided, with the most commonly used ones available on the toolbar. Other nice drawing features include the following: bond and ring sprouting; a chain tool; a drop-down, customizable menu for group formulas; reaction notation, many positioning tools; an online periodic table; and lots of keyboard shortcuts. Functions for easily rescaling your structure are available, and the program watches for and points out node overlays during mirroring, moving, and rotation operations. Adding text and graphics (arrows, orbitals, curves, elipses, and rectangles) is an easy process, with lots of control over fonts, line widths, fill attributes, and colors.

The Kekulé system has more chemical intelligence than any other drawing program I have tried, which means it gives you a lot of help when you're creating structures. Valence checks are made on all nodes during drawing procedures. Common elements are available on the button bar and a good pop-up periodic table makes it easy to add "uncommon" elements to your drawing. Brackets and parentheses are recognized as part of the structure, and included in molecular formula (MF) calculations. The MF is calculated as you draw, and you can choose to have it displayed on a bar at the top of the screen or only on demand. If you are redrawing or scanning a structure, the MF provides a quick way to validate your structure. This program also produces the most complete MOLfile (a connection table format developed by MDL, Int'l.) of any drawing package except the ISIS Draw software (an MDL, Int'l. product). Bracket information does not export to the MOLfile, but then ISIS Draw program can not do that either.

CONVERSION OF SCANNED STRUCTURES

The most unique and impressive function of this program is its ability to convert images (pictures) of chemical structures to editable connection tables (chemically intelligent structures). These "intelligent" structures can then be used in word processing documents or presentations or moved to structure databases. Included in the purchase price is scanner software that works with any TWAIN (a Windows and Macintosh application program interface) compliant scanner. The scanner software is used to get a structure image into the clipboard. Once an image is in the clipboard, a "crop, copy, and exit" function moves the image from the clipboard to the Kekulé system, where the structure recognition process is started. The scanning software is accessible from a button on the toolbar, making scanning and conversion a fast, two-step process. With the Kekulé program and a hand-held scanner, scanning of structures was so easy I found myself testing the software's conversion abilities on structures from product advertisements, the journal literature, and monographs.

This program carries optical character recognition (OCR) several steps further. It not only recognizes characters but also recognizes when characters represent atoms or functional groups. Single, double, triple, and stereochemical bonds are recognized, as are rings. Considering the complexity of the graphical representations used to depict chemical structures, the Kekulé system does a remarkable job. Like OCR, its

† The *Kekulé* program is a product of PSI International, Inc., 810 Gleneagles Court #300, Baltimore, MD 21286. Tel: (410) 821-5980, Fax: (410) 296-0712.

‡ Version reviewed: 2.0 alpha 5.

§ The views expressed in this review are solely those of the author and not of The Dow Chemical Company.

⊥ Any trademarks used are those of their respective companies.

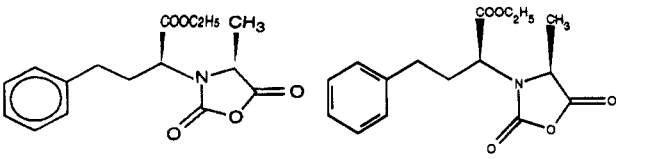

Figure 1.

capabilities are limited by the quality of the structure to be converted. Poor print quality and fuzzy or faint lines and character cause recognition problems. To help deal with the variability of printed structures, there are many parameters that can be adjusted. For example, if you are scanning structures that are printed with large, heavy type, you can adjust the character width and height if the program is not recognizing characters. You can also modify intercharacter spacing if there is too much distance between characters or characters and bonds for the program to normally interpret them as the same structure. There are many conversion parameters, and modification of them is not a trivial exercise. The manual is quite good, as is the technical support. If you are converting a large number of structures from the same source, optimizing the parameters is worth the effort, but it is not something you probably want to do for just one or two unusual structures.

The Kekulé program can recognize and translate benzene rings drawn with a circle in the middle, correctly translating them to alternating bonds. Stereochemical notation is also easily recognized, as are standard "shortcut" group formula notations, such as "COOH" (Figure 1). Group formulas can easily be "expanded" to show the detailed connections. If you some across group formula shortcuts that have not already been defined to the program, you add them.

The Kekulé system's structure recognition program does a credible job converting most printed structures. Even though PSI does not claim that hand-drawn structures can be processed, I tried it anyway and was surprised to find that the program was able to convert most of a clearly drawn structure. Complex ring structures, including fused rings and bridge structure are handled well. Some types of structures were not handled as well as others, such as structures containing radical notation (Figure 2) and π bonds. You can draw π bonds with the Kekulé system's drawing program, but the conversion program had difficulty recognizing them.

When the conversion program sees something it is not sure about, it shows you the area of the structure on the screen, and gives you its "best guess" interpretation. You can either accept the program's interpretation, fix the problem by entering the correct atom or functional group symbols, tell the program to "treat as bond", or discard. Because the conversion program is not going to be able to recognize everything, the ability to clean up structures during the conversion process is really nice.

MISCELLANEOUS FUNCTIONS

File open and save functions are greatly improved in version 2.0. The program follows Windows 3.1 standards, and it is now much easier to browse through different drives and directories. The Kekulé program can read many common structure storage formats, including ROSDAL, MOLfile, and ISIS (binary and ASCII). It can save in all these formats plus the SMILES, Windows Metafile, and WIMP graphics formats. Importing MOLfile format files from other structure drawing packages did not always work, sometimes causing a complete shutdown of the program. Some Macintosh operating system drawing programs are not able to export complete MOLfiles, and when the Kekulé program did not understand them it would abort back to the Windows environment. An error message would be preferable, but most other structure drawing programs just interpret what they can and do not tell you anything is missing, which is much more annoying than a program that just will not convert it at all.

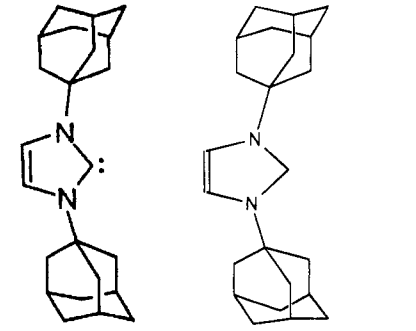

Figure 2.

The Kekulé system also contains a communications interface program, which can be configured to access third-party communications software. I did not try this, but this interface could give you the opportunity to integrate this structure entry program with a mainframe-based structure database. PSI also advertises the Kekulé system for the Macintosh environment, which I did not try.

SUMMARY

As an optical structure recognition program, Kekulé has no peer—I am aware of no other commercially available software packages with this functionality. In a production oriented structure data entry environment, Kekulé certainly will not replace your data entry personnel, but it can definitely give them some help. Someone skilled at drawing chemical structures can still draw simple structures faster than Kekulé can scan and interpret them. Where this tool can aid the work process is for the entry of complex printed structures. Even if the conversion program only gets 80% of a complex structure, editing an existing structure is faster than drawing it from scratch.

As a chemical structure drawing program, the Kekulé system is quite easy and intuitive, with many drawing functions plus templates, common atoms, and bonds available on button bars. Almost every drawing default available can be modified to your preferences. If you want to draw pretty pictures, the Kekulé program can do that. If you want to draw "smart" structures, the Kekulé package is one of the most intelligent around.