Methodology for Realizing Graphs in Stable Polycyclic Unsaturated Hydrocarbons

Aaron Germuth and Alex Aravind

Department of Computer Science
University of Northern British Columbia
E-mail: (germuth,csalex)@unbc.ca

June 24, 2013

Abstract

This report outlines the methodology used to find realistic graphs for classified Kekulé cells.

1 Introduction

Hesselink [1] has previously classified all possible Kekulé cells of rank <7. Additionally, he has discovered graphs for (nearly) every classification. However, most graphs do not correspond to realistic molecules. We wish to improve upon the classification, by providing more accurate graphs.

In order for a graph to be considered 'realistic' we have composed several restrictions:

1. Interval vertices are restricted to a degree of <4

2. Ports are restricted to a degree of <3

3. Cycles within the graph must be of size >4

4. All Graphs must be connected

5. No intersection of edges

With these restrictions we have found a graph for every classification of rank 4 and 5 using methods such as: exhaustive search, template molecules, and the editing of existing graphs.

2 Exhaustive Search

Hesselink (Section 4.2) [1] started with the observation that Kekulé cells are monotonic (preserve order) in the set of edges and therefore:

$E_0 \subseteq E \implies K(V, E_0) \subseteq K(V, E)$

Where V = set of vertices, E = set of edges, and K(V, E) = Kekulé cell of graph G = (V,E).

Every port assignment found in K(V, E_0) is also found in K(V, E). This implies the possibility that we can start with a smaller graph G_0, and slowly add edges to approximate K(V, E). If at any point $E_0 \nsubseteq E$, we have failed to find a graph and the process is repeated with more possible vertices. In order to get a terminating procedure, we fix a finite set V of vertices and a set of edges L, which can be added to the graph.

This method relies on starting with a smaller subset graph. This graph is obtained by including only the ports as vertices and allowing 'border channels' between the ports which have been proven to not effect the Kekulé Cell. From this small graph edges are slowly added.

Since we consider all sets E where:

$$E_0 \subseteq E \subseteq E_0 \cup L$$

the time complexity is already quite large. All possible internal vertex arrangements are added by considering each graph of the isomorphic class of matched graphs. When 8 vectices are used, this class is 10413 graphs long. Considering there are 2^{6*8} relations between a set of six vertices and a set of 8, and over 10000 graphs to try, it is unfeasible to attempt to search for any graphs with 8 or more internal vertices.[1]

Instead of simply terminating after finding the first graph, we can continue searching until the vertex limit is reached and retrieve all graphs which satisfy the cell. We then prune out graphs with unrealistic degrees on their vertices. This may still result graphs which are unrealistic (such as disjoint), but in most cases we can alleviate this by editing the graph (see Section 4).

3 Using Known Aromatic Hydrocarbons as Templates

The greatest disadvantage of last method is its limitation in the amount of vertices. Many polycyclic polyunsaturated hydrocarbons contain upwards of 10 atoms (pyrene, pyracylene, see Figure 1). Not only are molecules such as these stable, but they may be necessary. Adding more nodes to a graph tends to lower the degree of the nodes, and so larger graphs may be needed.

In this approach we take many commonly known conjugated hydrocarbons, and use them as a template. We consider all possible port locations (all nodes with degree <3) and classify the molecule with each combination. For example, Hesselink [1] looked at all 14 possible combinations of 5 ports on pyracylene and classified 5 different cells [1].

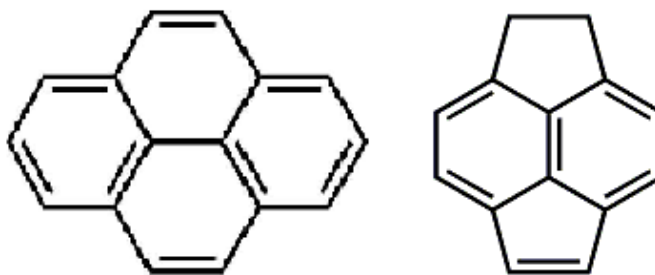[1] Only even sets of vertices are considered due to Lemma 2 of Hesselink Section 3.2 [1]

Figure 1: Two large polycylic aromatic hydrocarbons. From left to right, pyrene, and pyracylene.

In order to do this efficiently, we must consider symmetry. Consider the example given above, pyracylene with 5 ports. Since there are 8 nodes in pyracylene which can support a port due to the laws of carbon chemistry (<3 degree), there are 8 positions to put 5 ports. This gives $\binom{8}{5}$ or 46 possible combinations of port locations. However, this can be reduced to only 14 if we consider the symmetry groups of the nodes and trim out combinations which are the same.

The advantage to this approach is that one molecule may result in multiple classifications, and we know the molecule is realistic. However, we could never look at every one, and it is some what of a matter of luck in picking the right molecule.

4   Editing structures

There has been some literature on methods to change a graph without changing its cell. Hesselink et al [2] show a way to split a node of high degree into multiple nodes of lower degree. M.H Van der Veen [3] classifies a list of operations which can be used to edit structures. Of course, in order to edit a structure you have to have a graph to edit, so this method is used in addition to one of the above methods.

4.1   Joining Disconnected Graphs

Some graphs obtained from section 1 are disjoint, as connectedness is not a requirement. Molecules however, must be connected. If an atom is not directly bonded within the molecule, it is not part of that molecule. Luckily, we can use usually connect two disjoint graphs without affecting the cell. If we add internal vertices which must always conjugate 'within' themselves, they will not interfere with the cell. For example, see Figure 1.

In the right-hand graph of Figure 2, *y* is not a port and so must always contain a double bond. The only other node which is available to form a double bond with *y* is *x*, so in every resonance structure of this molecule x and y will contain a double bond with each other. Therefore *d* or *b* can never form a double bond with their new edge, and the Kekulé cell has not been changed. This now resembles a stable connected molecule.

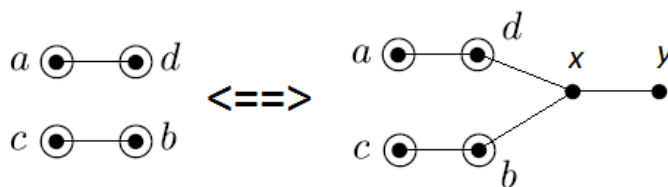This procedure resembles *Operation iii* and *Operation vi* of [3].

3

Figure 2: Transforming a disjoint graph into a connected one.

4.2 Reducing the degree on all ports

Sometimes the best graph you can find still has a high degree at one or more nodes. One common strategy is to 'extend' all the ports away by one bond, and leave a new internal vertex at their old position. I'm not sure yet whether this always holds true in all cases, but regardless is a common strategy to reduce the degree.

Hesselink et al. [2] show a technique to converge 3 nodes into one. They proved that the Kekulé remains constant while transforming the graph this way. If the process is done backwards, it can split a node of high degree into two nodes of lesser degree. It can even be done twice, by first merging and then re-splitting, in order to shuffle edges across nodes, similar to *Operation vii* of [3].

4.3 Removing Steric Hinderance

Another factor which causes results to not correspond to stable molecules is steric hinderance between carbons atoms. Carbon is tetrahedral and prefers to bond to 4 identities at approximately 109.5°. However, in triangular arrangements such as cyclopropene, the carbon bonds are forced to be 60°. In cyclobutadiene, the bonds are 90°. However, we can easily avoid such hinderance by again adding two internal vertices which are conjugated to everything (similiar to *Operation vi* of [3]). In the case of cyclopropene, we add two vertices to create cyclopentene. With cyclobutadiene, we get benzene (see Figure 3).
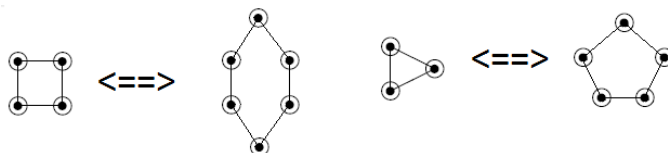

Figure 3: Removing steric hinderance by adding additional internal vertices. From left to right, cyclobutadiene, benzene, cyclopropane, cyclopentane

4.4 Adding or Removing Edges

After all of this there may be classifications which you haven't found a graph for. Based on the nature of Hesselink's classifications, as the Kekule classification 'number' increases (ex. K3 as

opposed to K7) the complexity of the cell increases. The number of resonance states and therefore port assignments increases until the last classification which contains every assignment of even length. We can use this fact to help us search for missing graphs. For example, assume you had a graph for K21 and were searching for K23. Since K23 has more port assignments than K21, we can try adding edges to our graph of K21 in order to find K23. This approach is by no means a guarantee, but has been useful in the past.

5   Current Results
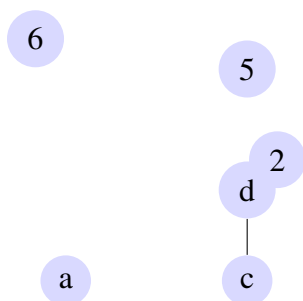
References

[1] W. H. Hesselink, Graph Theory for alternating hydrocarbons with attached ports. Indagationes Mathematicae, Elsevier, 24:115141, 2013.

[2] W.H. Hesselink, J.C. Hummelen, H.T. Jonkman, H.G. Reker, G.R. Renardel de Lavalette, M.H. van der Veen, Kekule Cells for Molecular Computation. Cornell University Library Online, 2013.

[3] M.H. van der Veen. π-Logic. PhD thesis, University of Groningen, May 2006.

[4] A.J. Heeger, S. Kivelson, J.R. Schrieffer, and W.-P Su. Solitons in conducting polymers, Rev. Mod. Phys.,60:781, 1988.