

# Extension of Structure-Oriented Lumping to Vacuum Residua

Stephen B. Jaffe\*

ExxonMobil Research & Engineering Company, 600 Billingsport Rd., Paulsboro, New Jersey 08066

Howard Freund and William N. Olmstead

ExxonMobil Research & Engineering Company, P.O. Box 900, Annandale, New Jersey 08801

The method of structure-oriented lumping (SOL) for describing the compositions, chemical reactions, and properties of complex hydrocarbon mixtures has been extended to molecules found in vacuum residua. The SOL approach was initially developed for gas oil and lighter fractions (i.e., those boiling at <1100 °F), where molecular compositions can be organized in terms of homologous series of single-core molecules (e.g., benzene and alkylated benzenes). In this work, residua molecules are represented as multicore molecules comprising linked assemblies of single-core species. The original SOL molecular description, using vectors of structural increments, is retained, along with the addition of information about the core linkages, and the extension meshes seamlessly with the conventional SOL notation. Using this formalism, properties of residua may be readily calculated and kinetic models of residua conversion processes developed.

## I. Introduction

Structure-oriented lumping (SOL), as proposed in 1992 by Quann and Jaffe,<sup>1</sup> represents individual hydrocarbon molecules as vectors of 22 structural increments. The SOL increments, shown in Figure 1, consist of three aromatic increments (**A6**, **A4**, **A2**), six naphthenic increments (**N6**, **N5**, **N4**, **N3**, **N2**, **N1**), a  $-\text{CH}_2-$  side-chain increment (**R**), and two isomeric descriptors: branches on side chains (**br**) and methyl groups on rings (**me**). A hydrogen increment (**IH**) is used to represent unsaturation, and a biphenyl bridging increment (**AA**) is also specified. Finally, two sulfur increments (**NS**, **RS**), three nitrogen (**AN**, **NN**, **RN**), and three oxygen (**NO**, **RO**, **KO**) are used to describe heteroatomic hydrocarbons. A complex mixture is then represented as a table of vectors (matrix): each row represents a molecule and each having a concentration in weight percent.

Molecules are built by selecting various numbers of these increments, as shown in Figure 2. SOL is consistent with a molecular class/homologous series organizational model of petroleum. It is convenient to characterize a molecule as comprising of a core structure plus a side-chain structure. Molecules boiling at temperatures less than vacuum residua (~1100 °F) readily lend themselves to this characterization. Members of a homologous series share the same core but differ in **R**, **br**, and **me**. Figure 3 shows several examples of SOL representations of homologous series, and Figure 4 shows 125 common core structures present in petroleum fractions boiling up through five condensed aromatic ring structures.

The SOL increment vector holds considerable molecular information useful for building kinetic rate expressions and calculating bulk physical properties. The stoichiometry associated with each increment is shown in Figure 5, and the matrix product of the SOL vector representation and the stoichiometry gives the carbon, hydrogen, sulfur, nitrogen, and oxygen content. The

molecular weight is similarly calculated. The aromatic content (aromatic carbons) used in various calculations (e.g., competitive adsorption) is obtained from **A6**, **A4**, and **A2**, and the molecular size for diffusion estimates can be made from the increments.

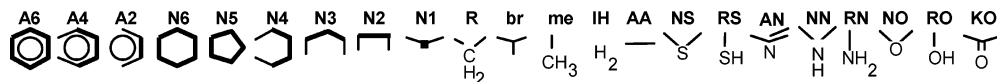
As discussed in Quann and Jaffe,<sup>1,2</sup> the increments reflect structural features that change during chemical reaction and provide a convenient framework to construct rule-based reaction networks of arbitrary size and complexity. SOL has been used successfully to develop detailed kinetic models for petroleum processes, where the feedstock molecules can be represented as a single-core structure and a side chain, which boil in the vacuum gas oil range or lighter (less than ~1100 °F). Molecules boiling in the residua range are more complex and require extensions to the traditional SOL approach. This paper discusses those extensions.

## II. Structure of Residua Molecules

Petroleum becomes far more difficult to analyze as molecular weight and boiling point each increase. It is no longer possible to analyze and quantify the individual compounds or isomer lumps, because the number and complexity of the molecular structures grows dramatically. Nevertheless, there has been much work on characterization techniques and attempts to assemble experimental information into molecular structures on entire vacuum residua as well as asphaltenes (the pentane- or heptane-insoluble fraction of vacuum residua). Altgelt and Boduszynski<sup>3</sup> have described the increasing complexity and many characterization techniques.

Individual molecules in vacuum residua cannot be measured; therefore, a combination of techniques are used and molecular structures are proposed that are consistent with the experimental data. Many measurements, such as elemental analysis (C, H, S, N, O, Ni, V), average molecular weight, and nuclear magnetic resonance (NMR), infrared (IR), and ultraviolet–visible (UV–Vis) spectroscopy give bulk properties. Separation techniques, such as short path distillation, high-performance liquid chromatography (HPLC), gas chromatography (GC), and solvent solubility give the amounts of the various fractions. By combining the two ap-

\* To whom correspondence should be addressed. Tel.: 856-224-3290. Fax: 856-224-3636. E-mail: stephen.b.jaffe@exxonmobil.com.



**Figure 1.** Depiction of the structure-oriented lumping (SOL) structural increments.

		A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	H	AA	NS	RS	AN	NN	RN	NO	RO	KO
2,3,5 Trimethyl hexane		0	0	0	0	0	0	0	0	0	9	3	0	1	0	0	0	0	0	0	0	0	0
Benzene		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Naphthalene		1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Phenanthrene		1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pyrene		1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Carbazol		2	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0
Benzofuran		1	0	0	0	0	0	1	0	0	0	0	0	-1	0	0	0	0	0	0	1	0	0
Fluorانون		2	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1
Adamantane		0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Cholesterol		0	0	0	1	0	2	1	0	0	10	2	2	-1	0	0	0	0	0	0	0	1	0
Water	H <sub>2</sub> O	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0

**Figure 2.** Examples of the SOL representations of hydrocarbons.

A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	H	AA	S	RS	AN	NN	RN	NO	RO	KO
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	3	0	1	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	4	0	1	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	5	0	1	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	6	1	1	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	7	1	1	0	0	0	0	0	0	0	0	0	0

A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	H	AA	S	RS	AN	NN	RN	NO	RO	KO
	2	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	2	0	1	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	3	0	2	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	4	0	2	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	5	0	2	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	6	1	2	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	7	1	2	1	0	0	0	0	0	0	0	0

A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	H	AA	S	RS	AN	NN	RN	NO	RO	KO
	2	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	2	0	1	1	1	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	3	0	2	1	1	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	4	0	2	1	1	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	5	0	2	1	1	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	6	1	2	1	1	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	7	1	2	1	1	0	0	0	0	0	0	0

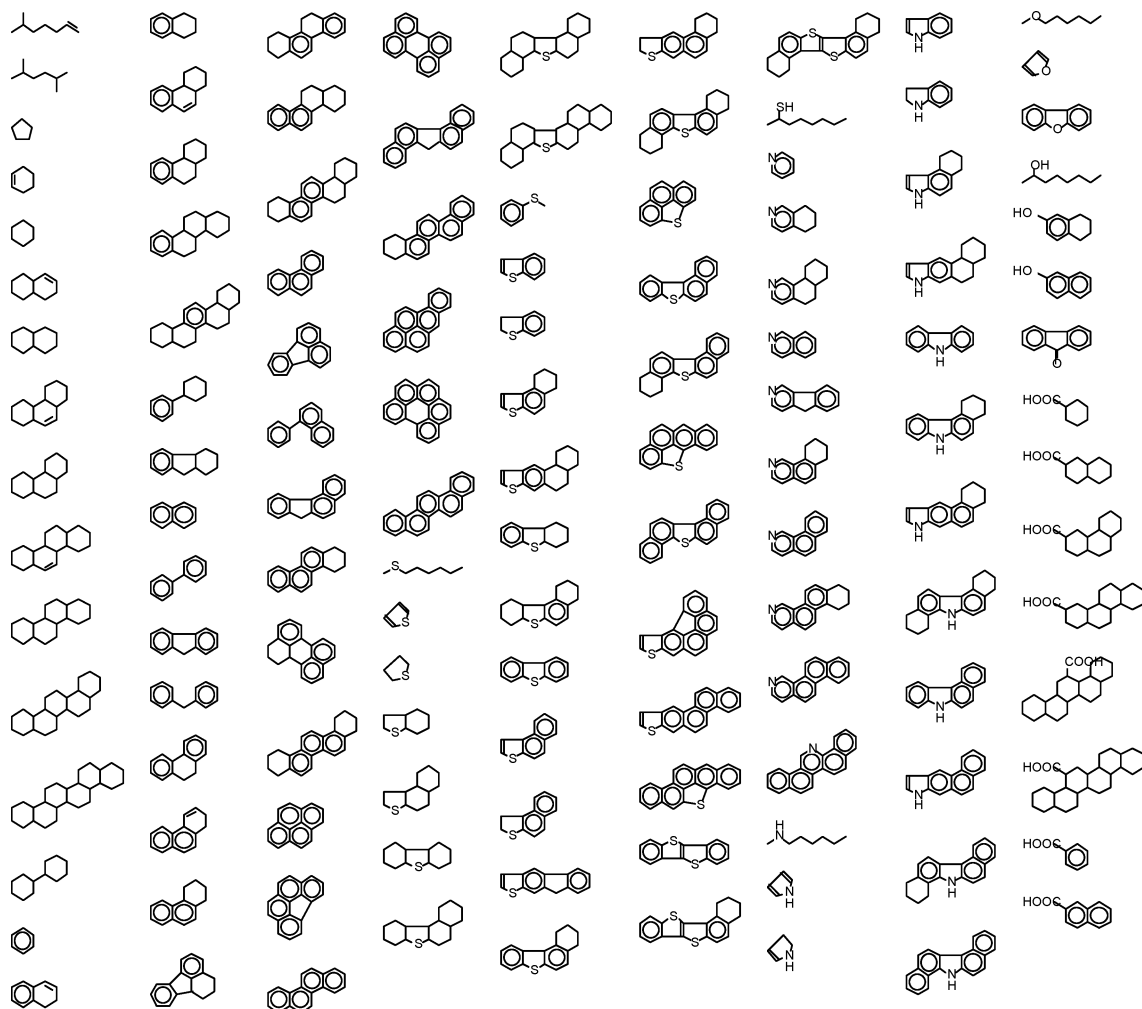
**Figure 3.** Sample homologous series using SOL notation.

proaches, as described in Altgelt and Boduszynski,<sup>3</sup> a more complete representation can be achieved.

Often, vacuum residua (including asphaltenes) are represented as average molecular structures, which is consistent with the characterization information. Although informative, the use of average structures is

inadequate for process modeling, because real conversion processes produce a wide variety of single-core molecules, which is inconsistent with a single starting molecule.

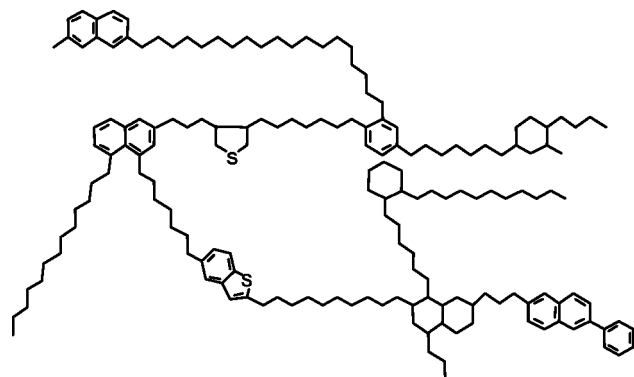
Ensembles of compounds have also been used to represent vacuum residua. Neurock et al.<sup>4</sup> have used a



**Figure 4.** Depiction of the known core structures in petroleum.

	A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	IH	AA	NS	RS	AN	NN	RN	NO	RO	KO
C	6	4	2	6	5	4	3	2	1	1	0	0	0	0	-1	0	-1	-1	0	-1	0	0
H	6	2	0	12	10	6	4	2	0	2	0	0	2	-2	-2	0	-1	-1	1	-2	0	-2
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1

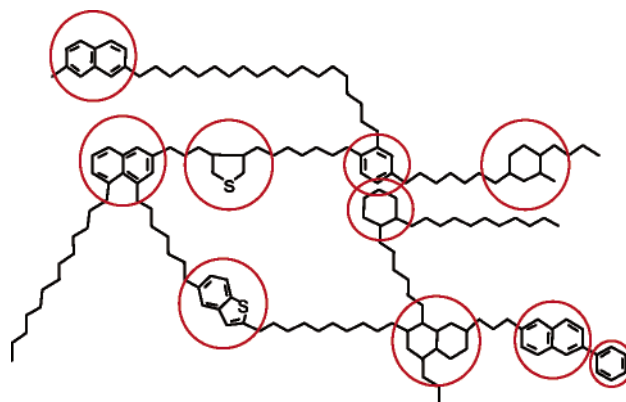
**Figure 5.** Matrix depicting the stoichiometry of SOL increments.



**Figure 6.** Depiction of a 10-core asphaltene molecule (from Strausz and Lown<sup>6</sup>).

stochastic (Monte Carlo) modeling approach to build ensembles that are consistent with characterization data. These ensembles can then be used as the starting material for kinetic and reactivity models.

Sheremata et al.<sup>5</sup> have extended this approach by incorporating more information and using a method of



**Figure 7.** Identified cores in an asphaltene molecule.

representation that facilitates the simulation of chemical reactions. They discuss two conflicting models for the molecular structures of asphaltenes. In the first, asphaltenes consist of large, highly condensed aromatic cores, usually more than seven rings per core. In the second, asphaltenes consist of smaller aromatic cores connected by covalent bonds or by aliphatic linkages, which can also contain heteroatoms. We prefer the second model because, when asphaltenes are selectively oxidized or pyrolyzed, the bulk of the products contain one to four aromatic and naphthenic rings with alkyl side chains.<sup>6-8</sup> These products could not arise from much-larger aromatic cores by any reasonable chemical reactions.

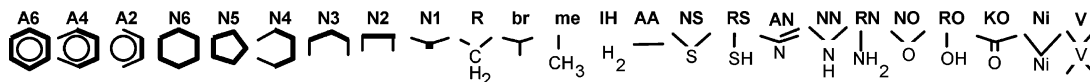
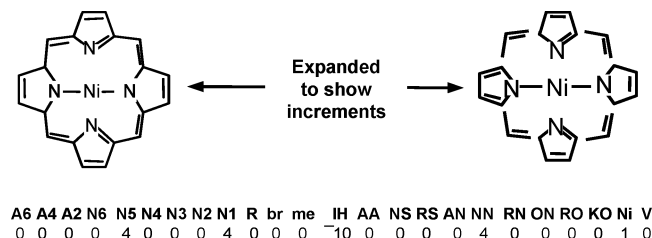


Figure 8. Depiction of the extended structural increments.

	A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	IH	AA	NS	RS	AN	NN	RN	NO	RO	KO	Ni	V
C	6	4	2	6	5	4	3	2	1	1	0	0	0	0	-1	0	-1	-1	0	-1	0	0	0	0
H	6	2	0	12	10	6	4	2	0	2	0	0	2	-2	-2	0	-1	-1	1	-2	0	-2	-2	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
Ni	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Figure 9. Matrix depicting the extended increment stoichiometry.

Chart 1



An example of an asphaltene molecular structure is shown in Figure 6. The much-larger portion of the residua that is not asphaltenes (the maltenes) can also be broken down to lower-molecular-weight, single-core products via pyrolysis or hydroconversion.

The approach we have adopted is to represent molecules boiling at temperatures lower than the boiling point of vacuum residua with the homologous series shown in Figure 4. For the vacuum residua, some of the molecules are continuations of these single-core homologous series with longer side chains, and others are multicore molecules composed of these same series. The multicore molecules, which are covalently linked, begin to appear as the boiling point and molecular weight increase. There is a smooth trend of increasing covalently linked cores all the way to the highest molecular weight asphaltenes. The single cores in Figure 6 have been identified in Figure 7.

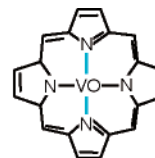
In this paper, we describe an approach to represent chemical structures in vacuum residua that is consistent with the analytical information available on the feeds and the products from chemical conversion.

### III. Extending the SOL Notation To Represent Residua

We intend to represent residua molecules as composite molecules built from the same cores found in the liquid fractions boiling at temperatures lower than vacuum residua. When these molecules react and decompose, they form single-core molecules that merge seamlessly with the kinetic models built for the lighter boiling fractions. The cores are linked by  $-\text{CH}_2-$  side-chain increments of specified length and bond strength; however, the exact location on the core will remain arbitrary. The notation used for the composite molecules will be similar to that for conventional SOL but will contain additional information.

**III.1. Nickel and Vanadium.** To represent porphyrins that contain nickel or vanadium, we must add two increments to the SOL structure vector. The revised list with a total of 24 increments is shown in Figure 8, and the stoichiometry is shown in Figure 9.

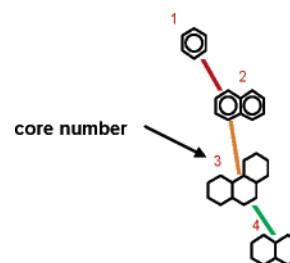
Chart 2



**III.1.a. Nickel.** The 23rd increment **Ni** is divalent and replaces two hydrogens in porphyrin molecules (see Chart 1).

**III.1.b. Vanadium.** The 24th increment **V** is tetravalent and, in the oxidized state, replaces two hydrogens, which are removed by the **KO** increment (see Chart 2).

**III.2. Multicore Molecules.** We represent multicore molecules as composites of single-core molecules with the appropriate connectivity information coded into the **AA** increment. Thus, the four-core molecule,



may be represented as four single-core molecules using conventional SOL, as presented in Chart 3.

Here, core 1 (benzene) is connected to core 2 (naphthalene), which is connected to core 3 (octahydrophenanthrene), which, in turn, is connected to core 4 (decalin). The red numbers in the drawing of the four-core molecule are the core numbers.

The coded **AA** increment requires additional discussion. We have organized connectivity information into a linkage table **link**, a bond strength table **type**, and the conventional biphenyl bridge increment. The **link** table for this molecule is

core 1:	2 0 0 0	(core 1 is linked to core 2)
core 2:	1 3 0 0	(core 2 is linked to code 1 and core 3)
core 3:	2 4 0 0	(core 3 is linked to code 2 and core 4)
core 4:	3 0 0 0	(core 4 is linked to core 3)

The linkage table has one row for each core and four columns, because we are allowing up to four cores to be linked to a single core.

We have distinguished three types of bond strengths. Bonds linking two aromatic cores are type 3 strength (red in the molecule drawing), which is the strongest. These are direct aromatic carbon–aromatic carbon bonds, in the 110–120 kcal/mol bond strength range. Bonds linking an aromatic core directly to a naphthenic core, or, more generally, to any nonaromatic carbon or heteroatom, are links of type 2 strength (amber in the molecule drawing), in the 90–100 kcal/mol bond strength range. Finally, we define bonds linking two naphthenic cores, or, more generally, bonds directly linking any



Chart 3

A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	IH	AA	NS	RS	AN	NN	RN	ON	RO	KO	Ni	V
1	0	0	0	0	0	0	0	0	0	0	0	0	172803750	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0	0	0	97204250	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	2	0	0	0	0	0	0	0	187202750	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	1	0	0	0	0	0	0	0	259201250	0	0	0	0	0	0	0	0	0	0

Chart 4

A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br
1	1	0	0	1	1	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
me		IH	AA					NS	RS	AN
0	0	0	0	0	0	0	0	0	0	0
0	0	0	1728	03750	9720	4250	1872	02750	2592	01250
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
NN	RN	ON	RO	KO	Ni	V				
0	0	0	0	0	0	0				
0	0	0	0	0	0	0				
0	0	0	0	0	0	0				
0	0	0	0	0	0	0				

Chart 5

A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	IH	AA	NS	RS	AN	NN	RN	ON	RO	KO	Ni	V
1	0	0	0	0	0	0	0	0	5	0	2	0	172802500	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	20006	0	0	0	97202750	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	2	0	0	0	6	1	3	0	187201500	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	1	0	0	0	30008	0	0	0	259201250	0	0	0	0	0	0	0	0	0	0

nonaromatic carbons or heteroatoms, to be type 1 strength (green), which is the weakest, with bond strengths of <82 kcal/mol. In this example, the **type** table is

core 1: 3 0 0 0 (core 1-2 linkage is type 3)  
 core 2: 3 2 0 0 (core 2-1 linkage is type 3 and the core 2-3 linkage is type 2)  
 core 3: 2 1 0 0 (core 3-2 linkage is type 2 and the core 3-4 linkage is type 1)  
 core 4: 1 0 0 0 (core 4-3 linkage is type 1)

Next, the conventional biphenyl bridge is

core 1: 0  
 core 2: 0  
 core 3: 0  
 core 4: 0

The complete **AA** information is

	linkage	type	AA
core 1:	2	000	3000 0
core 2:	1	300	3200 0
core 3:	2	400	2100 0
core 4:	3	000	1000 0

This matrix is then coded into a vector of three integers by changing to a mixed radix number system, using 24 24 24 24 5 5 5 10 as the base. This process is analogous to changing a measurement of 1 yard, 2 ft, and 3 in. to 63 in.

In APL, the notation is **A**⊖**B**, where **A** is 24 24 24 24 5 5 5 5 10 and **B** is

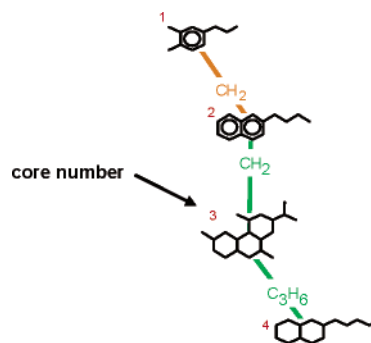
linkage	type	AA
2	000	3000 0
1	300	3200 0
2	400	2100 0
3	000	1000 0

172803750 97204250 187202750 259201250 ↔ **A**⊖**B**

(APL is a powerful vector-matrix oriented programming language with a concise notation and has been used here to describe SOL molecules and reaction rules.) Finally, the 4 × 24 matrix shown in Chart 3 can be converted to a vector of vectors. This is a vector of 24 elements, each of which has four elements (a so-called nested vector). (See Chart 4.) This multicore representation of a residua molecule has the same shape (a vector of 24 elements) as a conventional SOL representation and can be placed in the table of compositions.

The above multicore molecule has no side chains (**R** increment). When side chains are included, we must indicate how many carbons are associated with the

linkage. Consider the molecule



(the associated matrix is given in Chart 5). Here, each core has side chains: **R** = 5 6 6 8, while cores 2 and 4 have two and three  $-\text{CH}_2-$  imbedded in the linkages. One imbedded  $-\text{CH}_2-$  (**imbedded R**) on the second core is distributed across the two linkages. We have used the mixed radix number system approach with 10 000 as the base to code the **imbedded R** and the total side chain as a single integer. **Imbedded R** allows us to “weaken” the linkages between aromatic cores, changing type 3 to type 2 and type 2 to type 1.

In SOL, the precise location of the alkyl side chain(**R**), the branch (**br**), or the methyl group (**me**) is not specified. Likewise, in the extension to composite molecules, the precise location of the linkages is not specified, nor would they ever be known in a real mixture. Figure 10 shows several multicore molecules and their SOL increment notation. The decoded **AA** is also shown for clarity.

#### IV. Describing the Chemistry of the Residua

The major advantage of the original SOL approach is that petroleum reaction chemistry can be readily described and complex networks generated. SOL treats chemical reactions as transformations from reactant increments to product increments and encourages organization by class of reaction. The convenient framework of SOL “reactant selection rules” and “product generation rules” has been preserved with the extension to residua. Even though multicore molecules are represented as nested vectors, the single-core rules can still be applied, with some specific modifications. Yet, because of the multitude of pathways available with multicore molecules, care must be exercised to keep the number of species tractable. This is done by combining elemen-

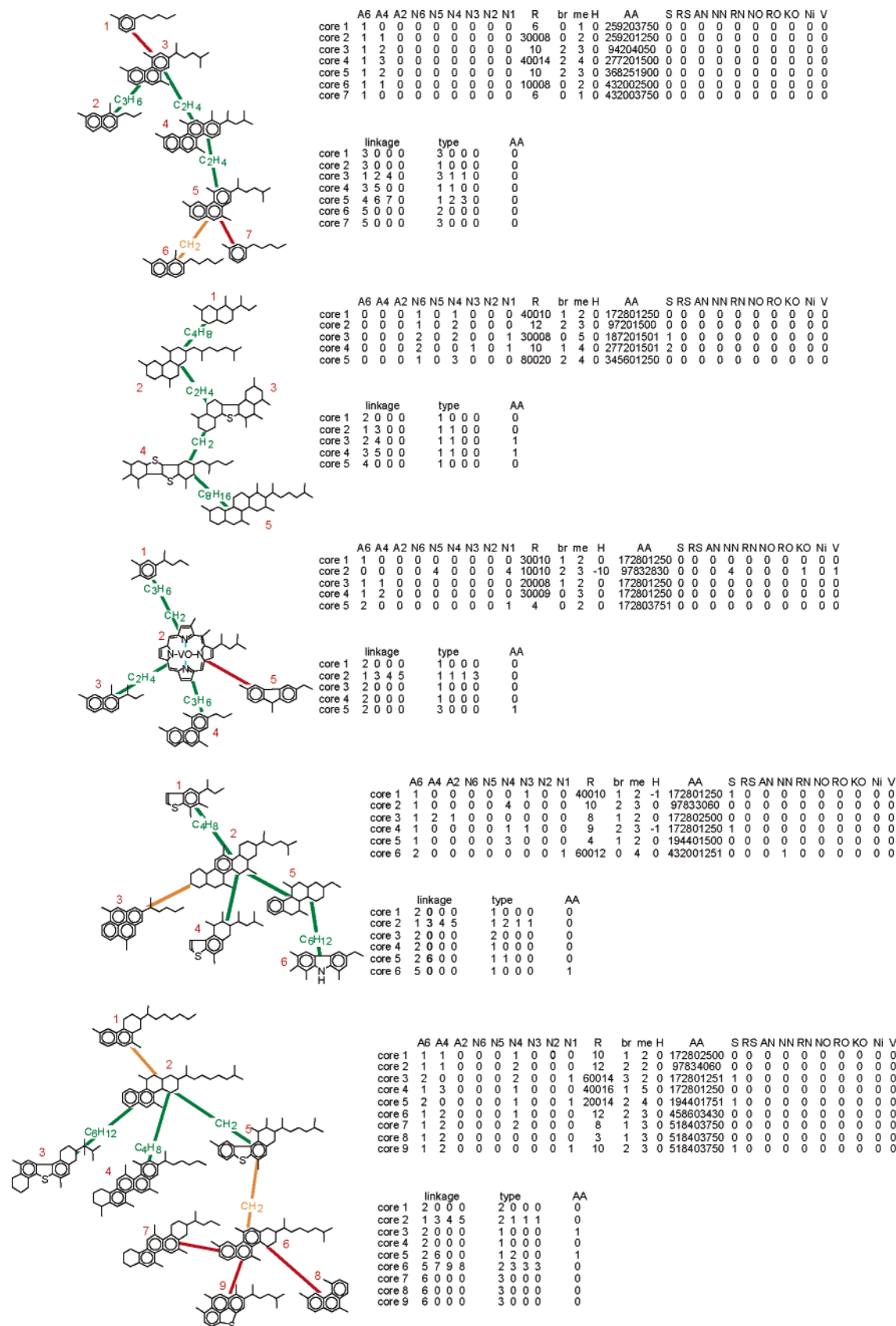
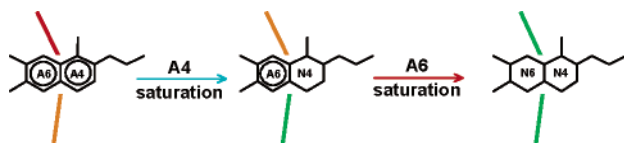


Figure 10. Multicore molecules and their expanded SOL representations.

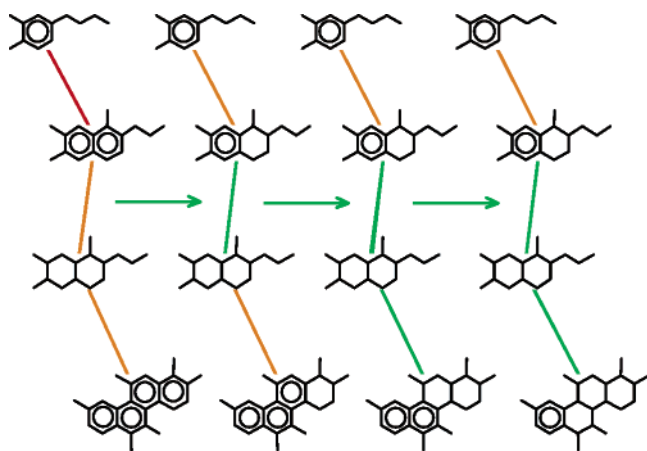
tary reaction steps into a single step. The reactions shown in the following figures often imply specific transformations; however, given the isomeric uncertainty of the structures, these are known only in a general way. We have chosen to focus on hydroprocess-

ing chemistry to illustrate the chemistry for multicore residua molecules. The hydrogen involved in the pathways is not shown explicitly.

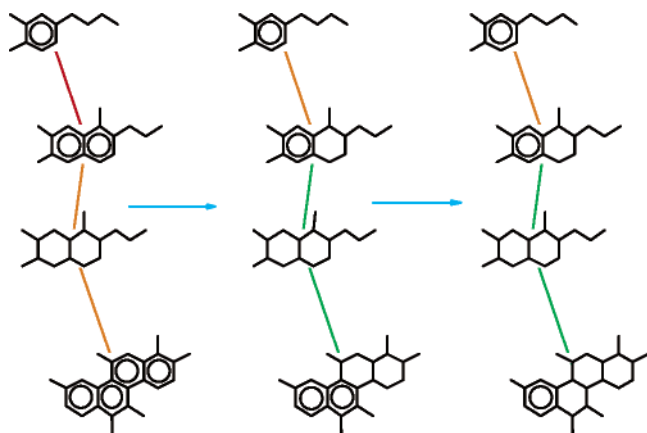
**IV.1. Aromatic Saturation.** Consider two types of aromatic saturation: **A4** and **A6** saturation.



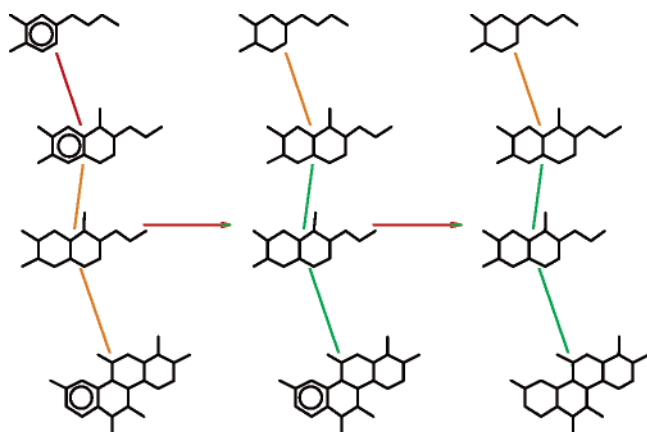
**A4** saturation is much faster than **A6** saturation and always precedes it. If we saturate one **A4** per core, several partially saturated structures will be generated;



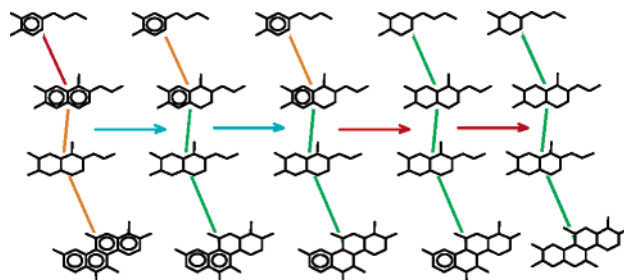
If we saturate half the **A4** increments per core in each reaction step, fewer molecules will be generated and might be a reasonable compromise:



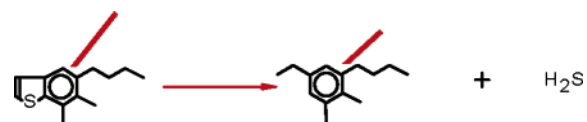
For **A6** saturation, we saturate half the available cores in each reaction step to get



Combining both steps, we get



**IV.2. Desulfurization.** Sulfur is present in petroleum in aromatic ring structures and sulfidic structures. For our purposes, aromatic sulfur is either in benzothiophene cores, dibenzothiophene cores, or disulfur aromatic ring systems. The remaining sulfur is sulfidic. These sulfur structures have a unique SOL structure and can be readily described. As an example, consider the desulfurization of a benzothiophene core in a multicore molecule:



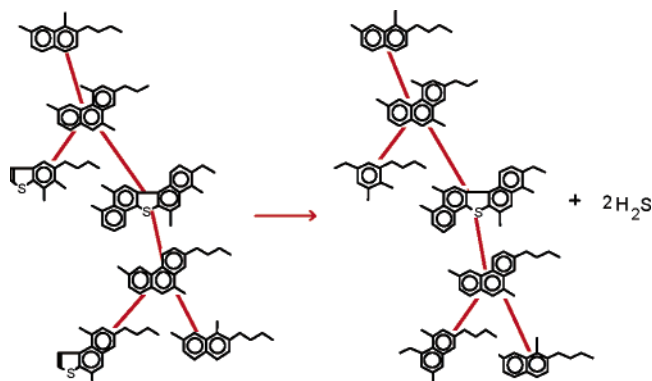
Using the same notation as that defined in Quann and Jaffe,<sup>1</sup> the reactant selection rule is

$$(\mathbf{A6} = 1) \wedge (\mathbf{N3} = 1) \wedge (\mathbf{IH} = -1) \wedge (\mathbf{NS} = 1)$$

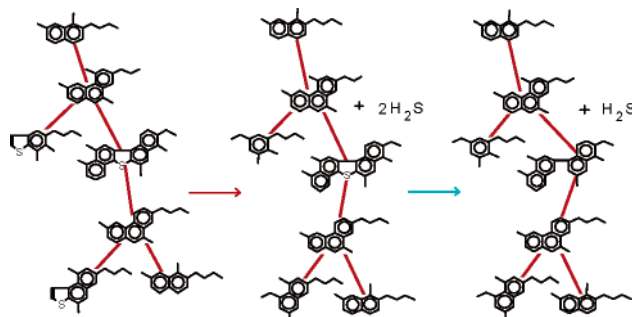
and the residua product generation rule is

$$\mathbf{N3HSR} - \mathbf{N3HSR} + {}^{-1}11^{-1}2$$

When these rules are applied to benzothiophene cores, we get the pathway

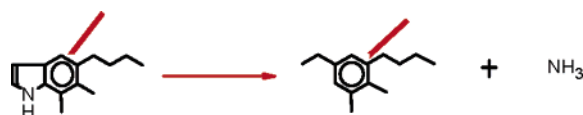


The desulfurization of dibenzothiophene cores is much slower; thus, a typical multicore pathway is



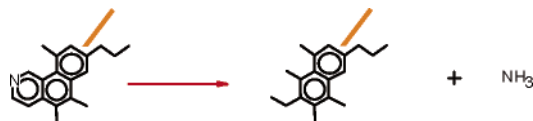
**IV.3. Denitrogenation.** Basic nitrogen and nonbasic nitrogen comprise the nitrogen species. The nonbasic

nitrogen rules for an indole core are analogous to the benzothiophene example, replacing NS with NN:



For the basic nitrogen, there are two rules, depending on whether the nitrogen is associated with an **A4** or an **A6**. The reactant selection rule for molecules with **A4** is

$$(\mathbf{A4} > 0) \wedge \mathbf{AN} = 1$$

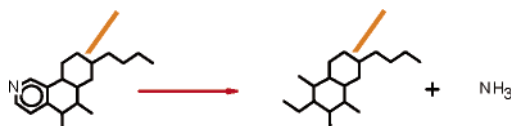


and the residua product generation rule is



The reactant selection rule for molecules with no **A4** is

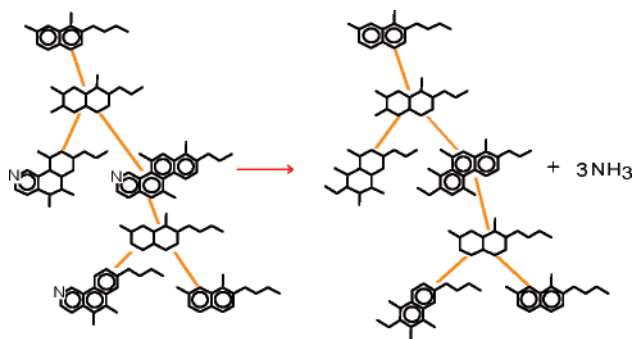
$$(\mathbf{A4} = 0) \wedge (\mathbf{N4} \neq 0) \wedge \mathbf{AN} = 1$$



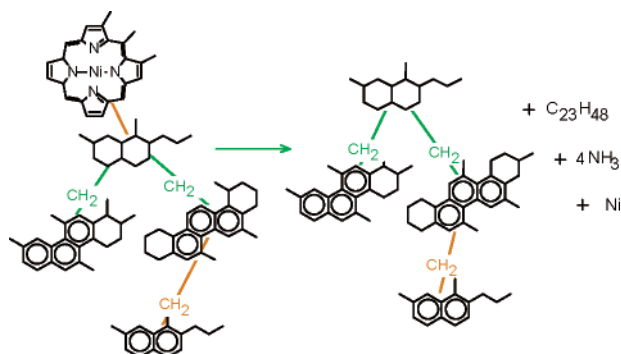
The residua product generation rule is



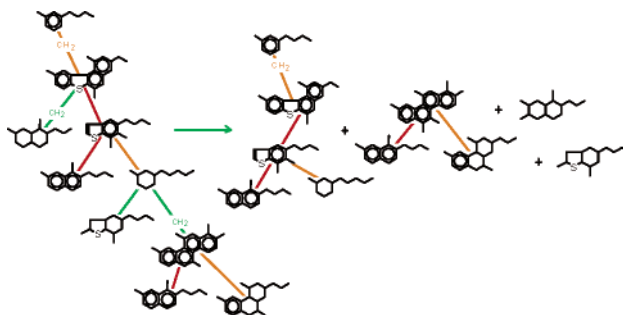
Both basic nitrogen rules are illustrated here:



**IV.4. Demetallation.** For nickel and vanadium removal, rather than producing a saturated porphyrin product, we further simplify the overall reaction by removing nitrogen and generating a fully saturated paraffin product, the metal, and ammonia. Because the metals concentration is low in petroleum, the concentration of reaction products from demetallation will also be very low, and the overall reaction network may be simplified. In a multicore molecule, the porphyrin is decomposed and anything linked is released. A simple reaction for the demetallation of a multicore porphyrin is

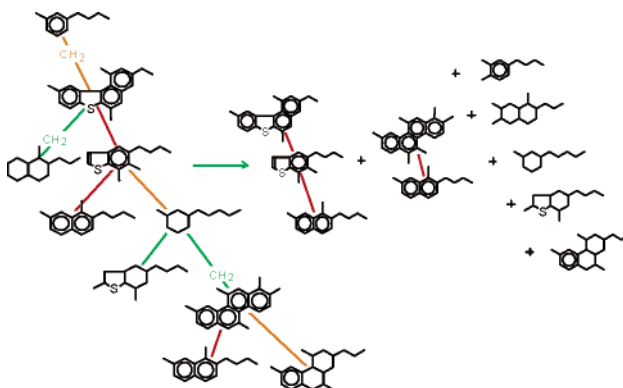


**IV.5. Multicore Decomposition.** Molecular decomposition of multicore molecules substantially decreases the molecular weight of the reactants. Molecular decomposition is the breaking of links holding the cores of a multicore molecule together, so we have the flexibility to break “weak” links or both weak and “medium” links. If there is any **embedded R** in the link, those carbons in the link are associated with one of the cores and remain with it. An example of molecular decomposition for a large multicore molecule in which we break just the weak leaks is



No other chemistry occurs in this step. In a hydroprocessing chemistry environment, the pieces that result after breaking the links are “healed” with hydrogen. In a thermal cracking environment, olefins would be created.

For molecular decomposition in which we break all but the “strong” links, such as that which occurs in hydrocracking, the reaction would be similar to the pathway depicted below.



Clearly, if we decomposed all the links, we would end up with all the individual cores comprising the original molecule as the products.

## V. Building the Ensemble of Residua Molecules

It is important to note that residua composition will be inferred. Modern techniques of chromatography



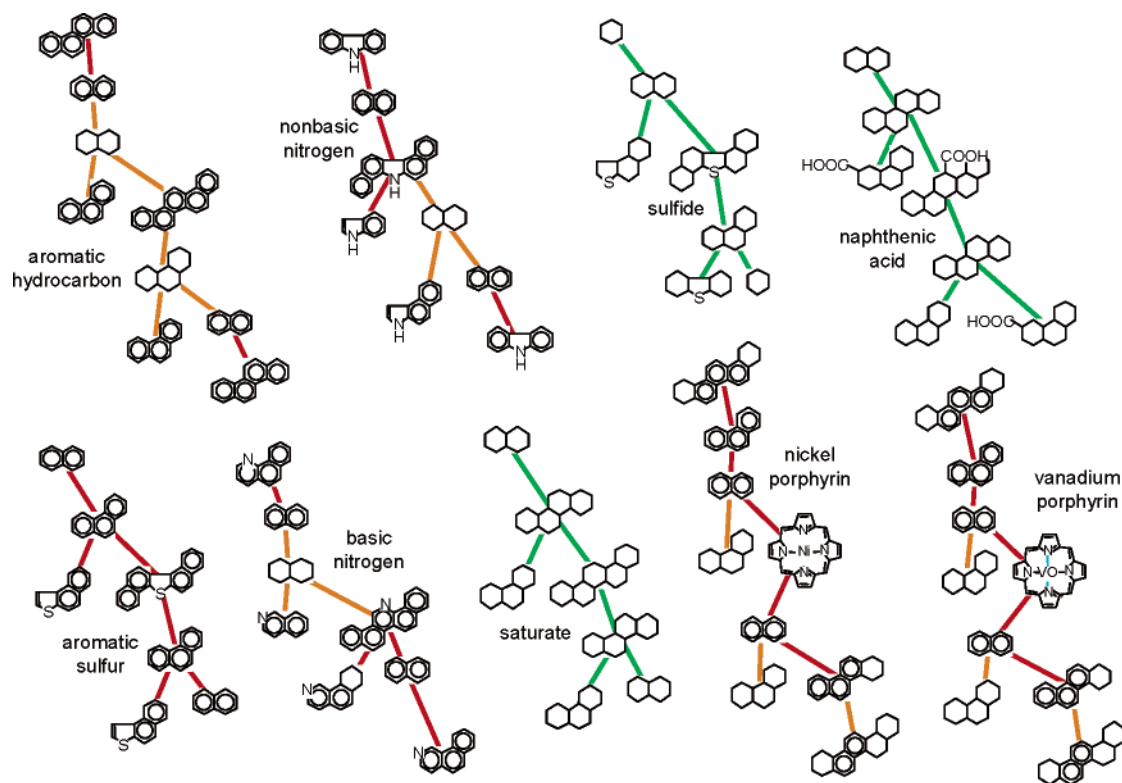


Figure 11. Seed molecules of various composition types.

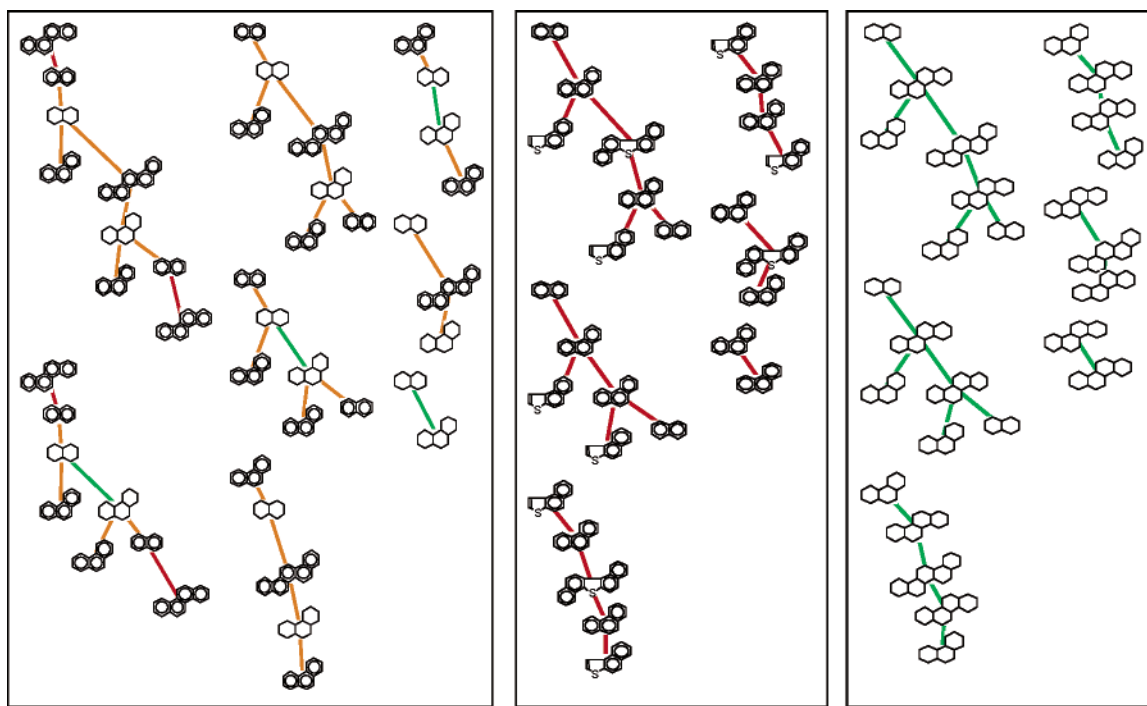


Figure 12. Smaller molecules from seeds.

(liquid and gas) and mass spectrometry are routinely used to determine the composition of petroleum fractions boiling at temperatures lower than that of vacuum residua; however, no such techniques can determine the composition of higher-boiling fractions. The best we can do for these fractions is to develop an ensemble of composite molecules that represents all the known features of the bulk residua and is consistent with the pertinent chemical transformations.

We would like the properties of the sample molecules to spread evenly across the dimensions of boiling point,

density, sulfur content, nitrogen content, aromatic content, and so on. The approach is to select a set of initial "seed" residua molecules and allow chemical transformations to generate products that populate the property and composition ranges.

**V.1. Choosing Seed Molecules.** Shown in Figure 11 are multicore molecules that span the range of composition types: an aromatic hydrocarbon, an aromatic sulfur, a sulfide, a nonbasic nitrogen, a basic nitrogen, a carboxylic acid, a naphthenic, a nickel

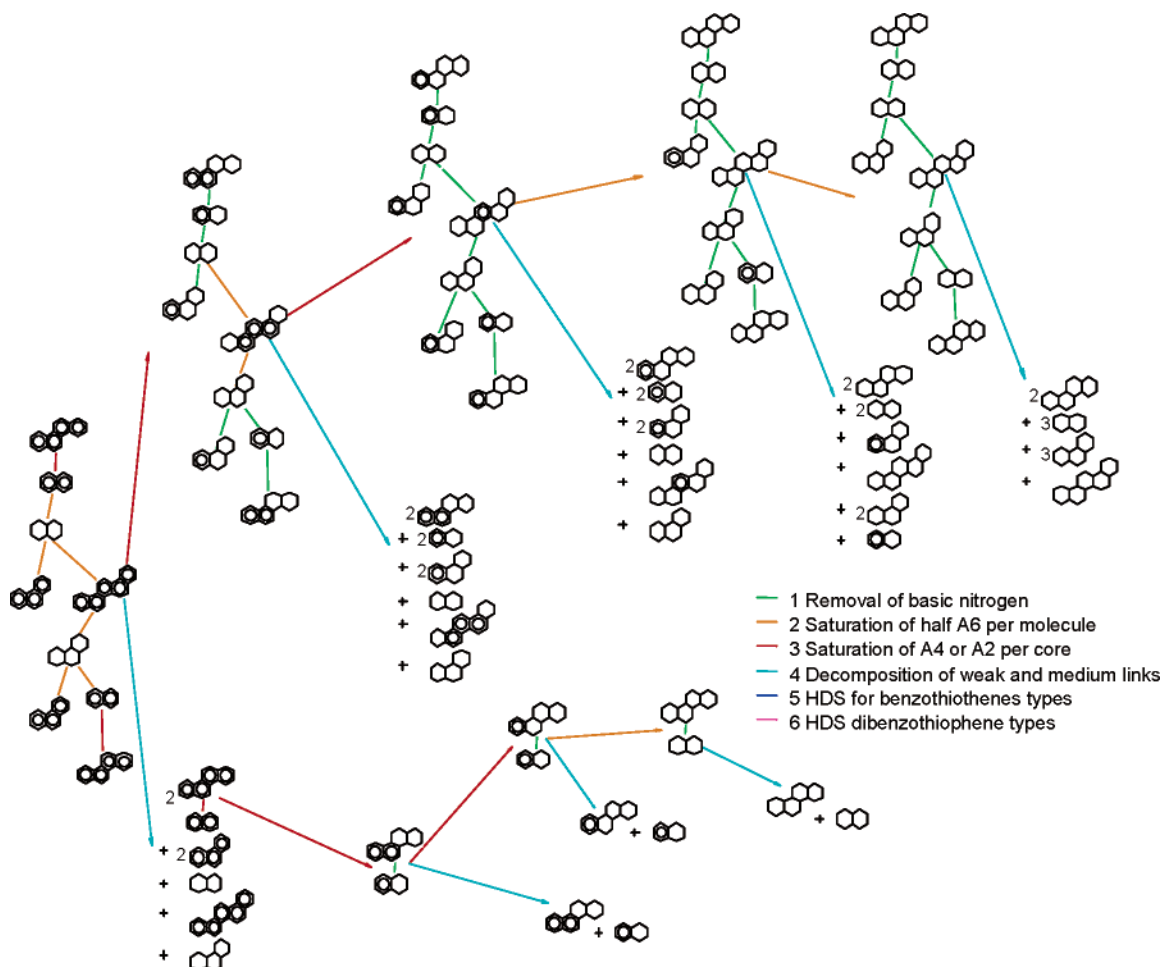


Figure 13. Reaction network from a nine-core aromatic molecule.

#### Chart 6

- |  |                                  |
|--|----------------------------------|
| 1 Removal of basic nitrogen                  | 6 HDS of sulfides                |
| 2 Saturation of A6 cores                     | 7 Removal of naphthenic acids    |
| 3 Decomposition of "weak" and "medium" links | 8 Nickel demetallation           |
| 4 Saturation of A4 or A2 cores               | 9 Vanadium demetallation         |
| 5 HDS for benzothiophene types               | 10 HDS of dibenzothiophene types |
|  | 11 Removal of non-basic nitrogen |

porphyrin, and a vanadium porphyrin. The side chains and **imbedded R** have been omitted for clarity.

To cover the molecular weight range and, thus, the boiling point, we must create molecules with the same composition type as that depicted in Figure 11 but with fewer cores. Figure 12 shows these smaller molecules, which have been prepared from the aromatic hydrocarbon, aromatic sulfur, and naphthenic seed molecules. The others may be prepared in a similar way. We have chosen 9 molecular composition types, with 7–9 cores each, and created 56 seeds.

**V.2. Building the "Chemistry Consistent" Ensemble.** To build the final ensemble, we apply the pertinent chemistry in the form of SOL reaction rules to our 56 seed molecules recursively. That is, we apply all the rules to the seed molecules and generate the products. Next, we apply the rules to the products to generate new products and so on. The process is complete when no new products are formed. The reaction rules we have used here are given in Chart 6.

Figure 13 shows the reaction network resulting from the application of the rules to the nine core aromatic hydrocarbons; 20 new molecules were created. Figure 14 shows the reaction network resulting from the

application of the rules to the seven core aromatic sulfur molecules; 17 new molecules were created, in addition to those created in Figure 13.

When we apply the reaction rules to all the seed molecules 245 molecules are generated, 44 of which have a single core. We have found that ~1500 composite molecules are sufficient to cover the compositional diversity found in residua. This approach creates a fixed set of molecules, which is a distinct advantage when designing a computer system to solve the kinetic network.

## VI. Calculating Residua Properties from Multicore SOL Molecules

The SOL representation contains enough information to calculate all the commonly measured properties of residua.

**VI.1. Residua Properties Calculated Exactly from the SOL Structure Vector.** The exact elemental composition (C, H, S, N, O, Ni, V) and the molecular weight can be calculated directly from the multicore structure vector and the stoichiometric matrix in Figure 9. The carbon fraction in aromatic rings (%CA), which is determined directly from  $^{13}\text{C}$  NMR, can also be calculated exactly. The total acid number (TAN) can be computed from the molecules containing  $-\text{COOH}$  groups.

**VI.2. Residua Properties Based on Models.** Very little model compound data are available in the residua range. Altgelt and Boduszinski<sup>3</sup> developed correlations for petroleum material, relating boiling point and specific gravity to structural properties. These correlations are based on data for gas oils and residua for which

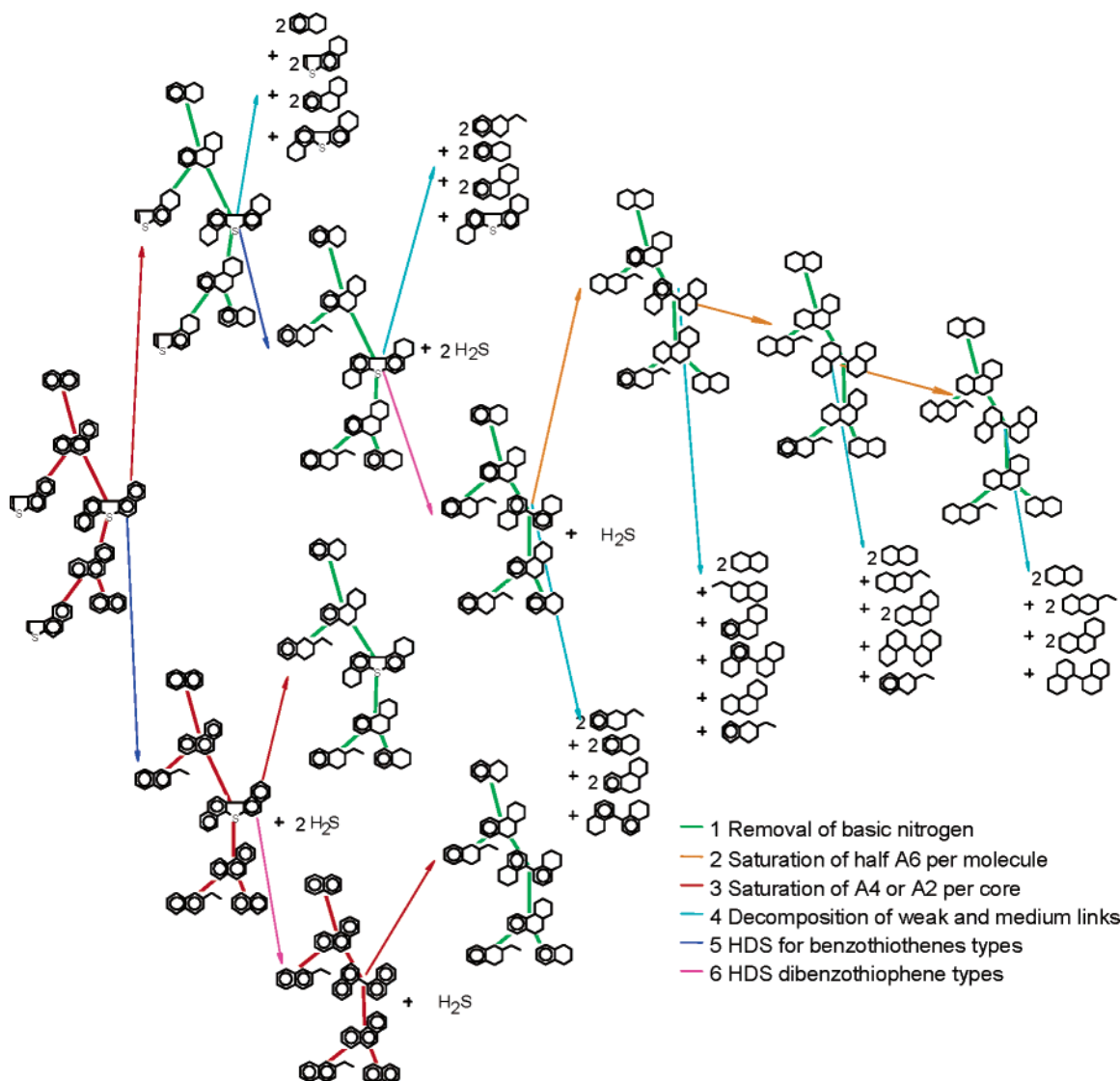


Figure 14. Reaction network from a seven-core sulfur molecule.

the authors were able to determine average structural properties of the petroleum material and boiling point from distillation data. From the SOL structure of the residua molecule, molecular weight and H/C ratio are known and the boiling point and density can be estimated from the following correlations:

$$MW = 170 + 2.67 \times 10^{-7} (BP)^3 (H/C)^{0.9}$$

where rearrangement gives

$$BP = \sqrt[3]{\frac{MW - 170}{2.67 \times 10^{-7} (H/C)^{0.9}}}$$

and

$$MW = 140 + 3.4 \times 10^{-7} \left( \frac{(BP)^3}{(SG)^{2.5}} \right)$$

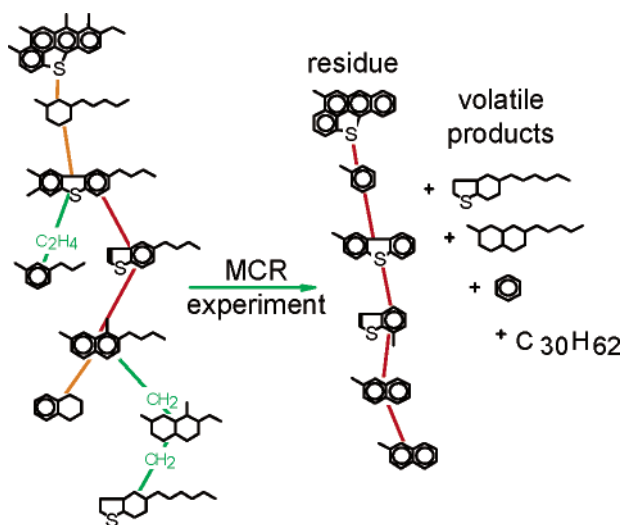
where rearrangement gives

$$SG = \sqrt[2.5]{\left( \frac{3.4}{2.67} \right) \left[ \frac{MW - 170}{(MW - 140)(H/C)^{0.9}} \right]}$$

In these relations, MW is the molecular weight of the molecule, BP is the atmospheric equivalent boiling point (in degrees Fahrenheit), and SG is the specific gravity.

**VI.3. Microcarbon Residue.** The microcarbon residue (MCR), as determined using ASTM D4530 (which is equivalent to the Conradson carbon residue (CCR) test, ASTM D 189), is a pyrolysis test that is used to characterize the coke-forming tendency of the residua. The sample is placed in a glass vial and heated to 932 °F (~500 °C) in an inert atmosphere in a controlled manner for a specific time. The sample undergoes cracking and coking reactions, and the volatiles produced are swept away by the inert gas. We have analyzed the volatile liquids and found the cut-point to be ~1030 °F (~555 °C). The residue is reported as a percentage of the original sample; this percentage is known as the MCR.

To simulate the MCR test, we will apply the pertinent thermal chemistry and determine those products that are not volatilized. This chemistry includes decomposition of the weak linkages, removal of sulfides, dealkylation of the alkyl groups attached to rings, and aromatization of naphthenic rings. The decomposition reaction has been discussed previously. The sulfide removal is dependent on the state of the sulfide. If it is a naphthenic sulfide (sulfur in a naphthenic ring), the core is liber-



**Figure 15.** Depiction of a sample molecule undergoing the microcarbon residue (MCR) process.

ated, regardless of the strength of the linkages, whereas if the sulfide is in a linkage, the linkage is broken. Furthermore, if the S atom is in a structure that can aromatize to benzothiophene, we allow that to occur. Otherwise, only the sulfur is removed. The dealkylation reaction removes most of the **R** group, leaving one methyl group per ring per core. Finally, aromatization converts all **N4** and **N6** rings to **A4** and **A6**. Any remaining molecules that boil at  $>1030$  °F are included in the MCR.

Figure 15 shows a sample molecule that is undergoing the MCR process. The weak links (marked in green in the figure) are the only linkages that are susceptible to thermal cleavage. Broken links are healed with hydrogen that comes from the aromatization reactions. Dealkylation removes all but one alkyl carbon from each core. The remaining naphthenic cores are aromatized. Thirty  $-\text{CH}_2$  groups are removed in this example. The remaining heavy molecule, which is the only product that boils at a temperature of  $>1030$  °F, represents  $\sim 48\%$  of the original molecule and is the MCR.

**VI.4. Asphaltene Content.** The asphaltene content of petroleum residua is determined by contacting the residua with an excess of solvent, often *n*-heptane. The heptane is a poor solvent for asphaltenes and causes their precipitation. Filtering, washing, and weighing yields the asphaltene content of the residua for the particular solvent used.

Asphaltene content is a function of the solubility characteristics of the molecules comprising the residua/solvent mixture. Scatchard<sup>9</sup> and Hildebrand<sup>10</sup> developed an expression for the activity coefficient of a given molecule in a particular solvent using the solubility parameter.<sup>13</sup> Relative size-effect corrections were made by Flory<sup>11</sup> and Huggins.<sup>12</sup> Barton<sup>13</sup> simplified the expression for the infinitely dilute case of a solute in a solvent, in which the activity of the asphaltene phase has been assigned a value of unity, allowing the determination of the solubility of the solute in the solvent:

$$\ln x_a = -\frac{M_a}{RT\rho_a}(\delta_a - \delta_s)^2 + \ln\left(\frac{v_a}{v_s}\right) + 1 - \frac{v_a}{v_s}$$

where  $x$  is the mole fraction solubility,  $M$  the molecular weight,  $R$  the universal gas constant,  $T$  the absolute temperature,  $\rho$  the density,  $v$  the molar volume, and  $\delta$  the solubility parameter. The subscripts "a" and "s" refer

**Table 1.** Relating Fedors<sup>14</sup> Groups for the Energy of Vaporization to SOL Increments

structure	SOL	Fedors
	A6	6 -CH=
	A6 A4	8 -CH= 2 >C=
	A6 2 A4	10 -CH= 4 >C=
	2 A6 N1 AA	8 -CH= 1 -CH2- 4 >C=
	2 A6 N1 S AA	8 -CH= 4 >C= 1 -S-
	N6 3 N4	12 -CH2- 6 -CH-

**Table 2.** Residua Properties Used To Select Residua Molecules

boiling-point distribution (wt% distilled versus boiling point)  
specific gravity (SG) versus boiling point (BP)  
sulfur versus boiling point (BP)  
basic nitrogen versus boiling point (BP)  
nonbasic nitrogen versus boiling point (BP)  
nickel versus boiling point (BP)  
vanadium versus boiling point (BP)  
total acid number (TAN) versus boiling point (BP)  
asphaltene versus boiling point (BP)  
microcarbon residue (MCR) versus boiling point (BP)  
carbon in aromatic rings (%CA) versus boiling point (BP)

to the solute and solvent, respectively. In this expression, the SOL representation allows us to determine the molecular weight, density, and molar volume of each residua molecule and the solvent. The infinitely dilute case is reasonable, because we consider the solubility of each molecule in the rest of the solution. Asphaltene character of a given molecule results when the mole fraction of that molecule in the solution is greater than the calculated mole fraction solubility. Here, the solution is the mixture of the solvent and the crude/residua. We still need to determine the solubility parameter for all the molecules in the system.

The solubility parameter is defined as

$$\delta = \left(\frac{E}{V}\right)^{0.5}$$

where  $E$  is the energy of vaporization and  $V$  is the molar volume. We use the correlation for density together with molecular weight to estimate the molar volume. Fedors<sup>14</sup> (in Barton<sup>13</sup>) presents a group contribution method to estimate the energy of vaporization, and we have related the SOL increments to these groups. Table 1 shows some examples of the relations. For composite molecules, we apply the appropriate groups for the linkages.

## VII. Determining Molecular Abundance

The final task in building a SOL model of composition is to determine the abundance for each composite molecule. The ensemble is fixed by the chemistry, so we



need only to match the abundance of the molecules to the available data. Bulk property measurements are typically available and often estimates have been made of various property distributions versus boiling point. Several researchers have reported methods to determine the molecular abundance, subject to constraints.<sup>4,15,16</sup> In our laboratory, we have used a non-negative Lagrange multiplier method, starting with an equimolar composition distribution. Table 2 lists the property objectives.

### VIII. Conclusion

We have extended the method of structure-oriented lumping (SOL) to the multicore molecules that comprise the residua. The extension merges seamlessly with the single-core representations. The multicore representations carry sufficient information so that all properties currently measured can be either directly calculated or readily modeled. Furthermore, it is possible to generate rule-based reaction pathways under program control, which capture the residua conversion chemistry.

### Literature Cited

- (1) Quann, R. J.; Jaffe, S. B. *Ind. Eng. Chem. Res.* **1992**, *31*, 2483.
- (2) Quann, R. J.; Jaffe, S. B. *Chem. Eng. Sci.* **1996**, *51*, 1615.
- (3) Altgelt, K. H.; Boduszynski, M. M. *Composition and Analysis of Heavy Petroleum Fractions*; Marcel Dekker: New York, 1994.
- (4) Neurock, M.; Nigam, A.; Trauth, D.; Klein, M. T. *Chem. Eng. Sci.* **1998**, *49*, 4153.
- (5) Sheremata, J. M.; Gray, M. R.; Dettman, H. D.; McCaffrey, W. C. *Energy Fuels* **2004**, *18*, 1377.
- (6) Strausz, O. P.; Lown, E. M. *Fuel Sci. Technol. Int.* **1991**, *9*, 269.
- (7) Strausz, O. P.; Mojelsky, T. W.; Faraji, F.; Lown, E. M. *Energy Fuels* **2002**, *16*, 809.
- (8) Gray, M. R. *Energy Fuels* **2003**, *17*, 1566.
- (9) Scatchard, G. *Chem. Rev.* **1931**, *8*, 321.
- (10) Hildebrand, J. H. *J. Am. Chem. Soc.* **1935**, *57*, 866.
- (11) Flory, P. J. *J. Chem. Phys.* **1942**, *9*, 660.
- (12) Huggins, M. L. *J. Chem. Phys.* **1941**, *9*, 440.
- (13) Barton, A. F. M. *CRC Handbook of Solubility Parameters and Other Cohesion Parameters*, 2nd Edition; CRC Press: New York, 1991.
- (14) Fedors, R. F. *J. Polym. Sci. C* **1969**, *26*, 189.
- (15) Allen, D. T.; Liguras, D. Structural Models of Catalytic Chemistry: A Case Study of a Group Contribution Approach to Lumped Kinetic Modeling. In *Chemical Reactions in Complex Mixtures, The Mobil Workshop*; Sapre, A. V., Krambeck, F. J., Eds; Van Nostrand Reinhold: New York, 1991.
- (16) Hudebine, D.; Verstraete, J. J. *Chem. Eng. Sci.* **2004**, *59*, 4755.

Received for review May 20, 2005

Revised manuscript received August 24, 2005

Accepted August 30, 2005

IE058048E