



BUILDING USEFUL MODELS OF COMPLEX REACTION SYSTEMS IN PETROLEUM REFINING

R. J. QUANN and S. B. JAFFE

Mobil Technology Company, Paulsboro, New Jersey 08066

Abstract - The identity and composition of petroleum's countless molecular components is beyond the resolution of analytical techniques. Nonetheless, models of petroleum refining processes must have a molecular basis to be useful in predicting the molecular composition of products and their physical and quality properties required by industry specifications. Strategies for solving this seemingly intractable problem are presented.

1. INTRODUCTION

The accuracy and reliability of industrial process models is best served by the incorporation of fundamental chemistry, thermodynamics, kinetic rate processes, and physical transport phenomena. However, the complexity of some reaction systems in terms of the number of components and reactions precludes a determination of all process fundamentals. This paper will review modeling techniques for complex reaction systems and the methodology for developing useful process models for petroleum refining. We define complex reaction systems as those having a large number of *both* reactants and reactions - so large that it is not possible to measure the molecular composition or study the kinetics in sufficient detail to allow a complete molecular and fundamental treatment.

For a perspective on complexity, we need to contrast modeling in the traditional chemical industry with that in the petroleum refining industry. The typical chemical process will have only a few molecular species and a sufficiently detailed understanding of the chemistry, thermodynamics, and kinetics of the reaction system. This is not meant to imply that the reaction system is trivial or that the controlling rate equations easily resolved, rather that it is *feasible* to obtain this fundamental information. In petroleum refining, the number of hydrocarbon species is enormous, and the number of chemical reactions typically an order of magnitude greater. Despite significant advances in recent years, it is quite *infeasible* to measure the complete molecular content of petroleum, obtain all molecular properties, or study the kinetics of all chemical or catalytic reactions. Petroleum processes are generally intractable from a fundamental standpoint.

Yet the need for process models to predict accurate and detailed product yields in terms of molecular species may be greatest in the petroleum refining industry. A chemical plant will generally process the same feedstock with the intent of making the same product, and under well established conditions. The molecular composition of crude oil, a refinery's feedstock, may change on a weekly basis depending on source, as might the most economical slate of product options due to market conditions. Crude oil composition can affect product yields and quality, and market prices can influence process operating strategies. A refinery process might therefore require frequent, composition dependent optimizations. Further, environmental regulations of

the 21st century will place restrictions on the molecular composition of petroleum products, in addition to their traditional quality specifications.

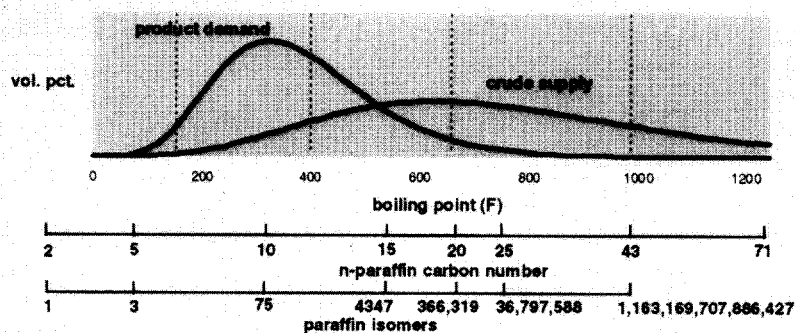
Given the fact that modeling simplifications will have to be made, how are we to proceed? The correct approach is to analyze the system in reverse order, from our needs rather than from our knowledge. The first step is to define the model's purpose - what should its output be? The major petroleum processes convert crude oil into lower boiling, and more valued, gasoline, jet fuel, diesel, fuel oil, and lubricating oil - and they must do so by meeting a multitude of performance or quality specifications as shown in Figure 1. The model's purpose must therefore be to predict boiling point yields *and* product quality, as a function of plant operating conditions, catalyst, and the feedstock source. The next step is then to assess how much molecular detail is required to achieve this, given that yields are determined by process chemistry and that all properties have a molecular basis. Petroleum has a broad boiling point or carbon number spectrum, with many molecular types and infrangible isomer complexity for C_{10+} molecules (Figure 1). Obviously, a molecular-based model must resort to some form of approximation because of the sheer number of structural permutations - only a thousand billion for C_{43} paraffin isomers alone! Finally, we need to formulate a working, kinetic description of the complex reaction system consistent with the level of molecular detail.

2. ASSESSING THE MINIMUM LEVEL OF MOLECULAR DETAIL

Only a few specific molecules currently require actual measurement or prediction in petroleum processing because of their economic value or environmental concern. These include paraffins and olefins in the C_1 - C_5 carbon number range, and benzene, toluene, and xylenes. Fortunately, they are easily measured by standard gas chromatography (GC) techniques. Far more molecular speciation is required if models are to incorporate process chemistry and compute product properties. We need to first examine the molecular basis of petroleum, its mixture properties, and its refining chemistry, and then develop a modeling strategy with molecular representation that is consistent with our needs and within the capabilities of analytical methods.

Petroleum's Molecular Detail

Petroleum can be divided into three categories based on complexity and resolution by analytical methods: (1) material boiling below 450K, whose molecular content is completely measurable; (2) material boiling between 450K and 800K, where classification of molecular types is possible but with ambiguity or uncertainty in discerning exact molecular (isomeric) structure; and (3) material boiling above 1000K where complexity can exceed even the type classification. Modern GC techniques (Figure 2) routinely identify and quantify most of the several hundred hydrocarbon components in the gasoline range (<450K). Much of the gasoline product is formed by conversion of the larger, higher boiling molecules of petroleum in processes such as hydrocracking and fluid catalytic cracking (FCC).. To predict their yield using kinetic models, the parent structures in the starting petroleum, or at least a representation of them, are needed as well. Our emphasis here is to review what is measurable and known about the molecular structure of petroleum for the purpose of developing a molecular description of 450K+ material. A comprehensive review of analytical methods has been given elsewhere (Altgelt and Boduszynski, 1994).



Processes (feedstock boiling range)



Products



Physical Properties



Performance properties

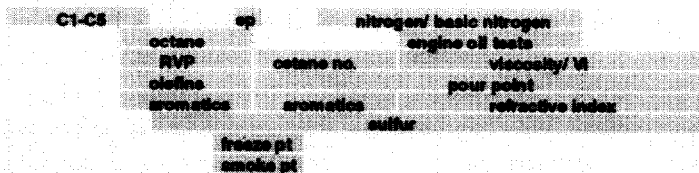


Figure 1. The boiling point distribution of petroleum, process feedstocks, and refinery products with their relevant physical and quality properties.

Although we can identify only a few specific molecules in the 450K+ material, analytical techniques, primarily chromatography and mass spectroscopy, have enabled a molecular characterization of petroleum through the use of two concepts - the *homologous series* and the *molecular class*. The simplest homologous

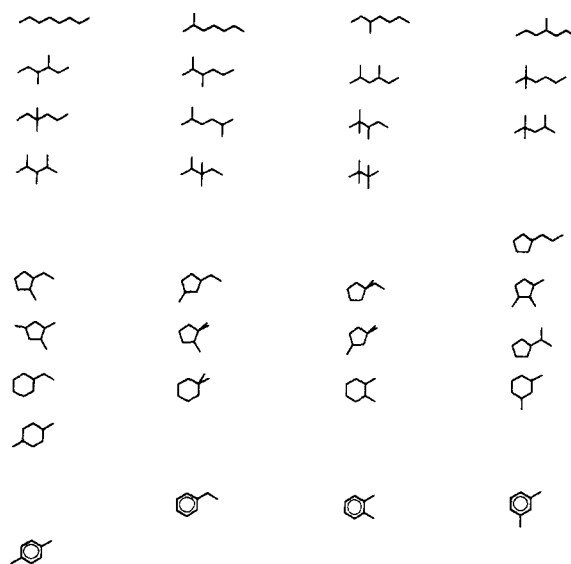


Figure 2. The C₈ components of petroleum from gas chromatography analysis.

series, the *n*-paraffins, is shown in Figure 3. The structure of molecules in this series is identical except for the addition of one -CH₂- (methylene) group for each successive member of the series. This series is common in petroleum and especially prevalent in waxy petroleum.

We can conceive of other homologous series for ring compounds as constructed from a core ring group and a linear chain of -CH₂- groups as shown in Figure 3 for alkylbenzenes. In this case there are a number of structural possibilities for the base group involving additional substituents. A *structural* homologous series has the same base but with linear alkyl chain varying in carbon number. Using chromatography, Ostoukhov et al. (1983) identified these seven homologous series in petroleum in the C₁₂ - C₃₀ range, but found other alkylbenzene structures as well. The other structures, accounting for most of the alkylbenzene mass, were unidentifiable because of the numerous isomeric possibilities resulting from branching on the alkyl chain or additional substituents. Using mass spectroscopy, we can measure the entire alkylbenzene series as a *mass* or *carbon number* homologous series (Figure 3), but with no information on the structural arrangement of substituent groups or of the longer alkyl chain. In this case all isomers at a specific carbon number are summed or lumped, and the alkyl group structure defined only by its total carbon number.

How diverse are the isomeric structures of alkyl groups? We simply don't know since chromatography can only resolve isomers in exceptional cases such as the *n*-paraffins, the low boiling range, and biomarkers. Alkyl group structure may vary with respect to the number of groups on a ring, the carbon number of the groups and the branching structure of the groups. NMR measurements of the local chemical environments of carbon atoms can reveal information on the 'average' structural arrangements of alkyl groups. Mair and Barnewall (1964) applied this technique to a physically isolated monoaromatic fraction of petroleum and deduced that the average alkylbenzene in the C₁₃ - C₁₅ range had one methyl substituent and one long chain with a methyl branch. Average, however, does not imply predominant. From a fascinating comparison of the mass spectroscopy fragmentation patterns of both synthetic and petroleum alkylbenzenes in the C₂₀ - C₄₀ range, Hood et al. (1959) deduced that the alkyl group structure is not as diverse as might be

structural homologous series in carbon number

n - paraffins

CH₄

structural homologous series for alkylbenzenes

n - alkylbenzenes



o - methyl alkylbenzenes



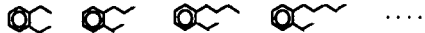
m - methyl alkylbenzenes



p - methyl alkylbenzenes



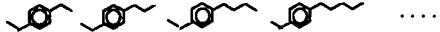
o - ethyl alkylbenzenes



m - ethyl alkylbenzenes



p - ethyl alkylbenzenes



mass homologous series of alkylbenzenes

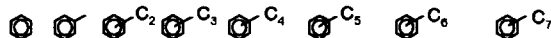


Figure 3. Examples of structural and mass homologous series.

expected given the number of possible isomeric permutations. They observed that petroleum alkylbenzenes are composed almost entirely of those having a single long chain and 0 to 4 methyl substituents on the benzene ring. Two sizable alkyl groups on a single ring is virtually absent in their study. However, Kohnen et al. (1990) found alkylthiophenes containing two substantial alkyl groups with carbon numbers in excess of 12.

Variation in alkyl group structure among different petroleum sources is also not well understood, but one might expect the linearity of alkyl chains to correlate with the wax content or *n*-paraffin content of a crude oil. Homologous series are not just a default organizational scheme resulting from analytical limitations, but

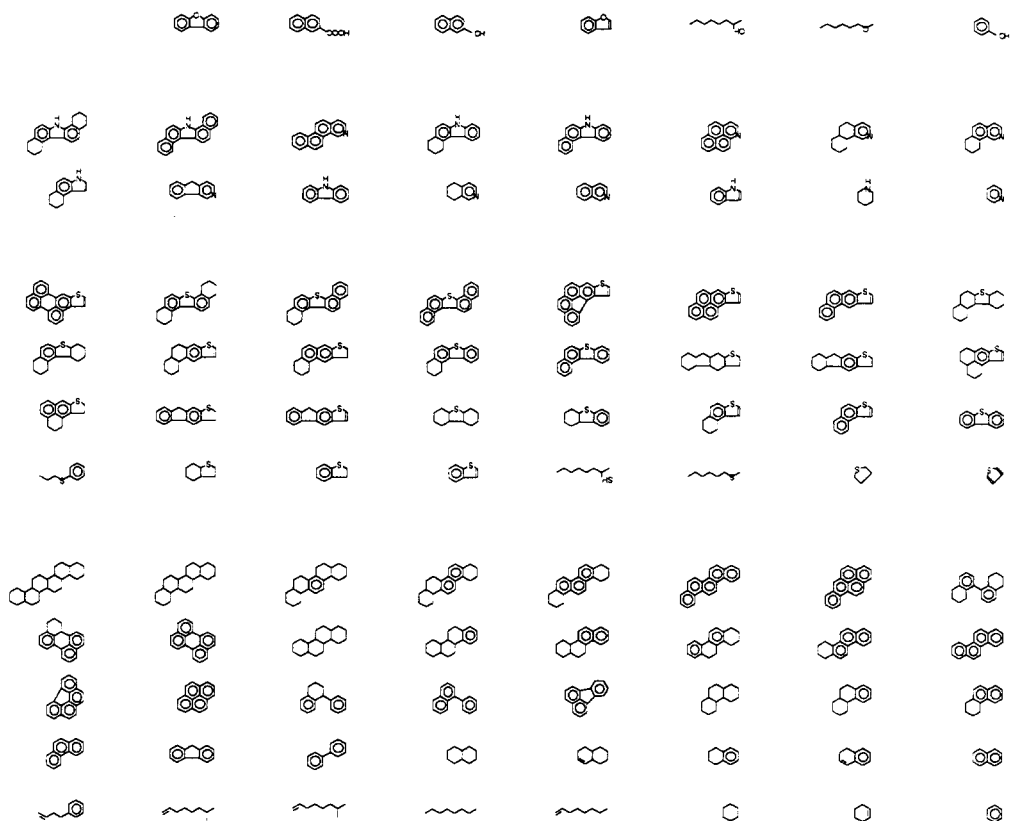
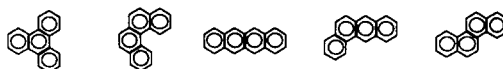


Figure 4. Common molecular classes in petroleum.

have chemical and physical significance as well. They generally exhibit smooth distributions that, combined with their isomeric structure, reflect petroleum's biological origin and maturation processes.

The concept of homologous series allows us to organize molecular classes. There are many molecular classes, each having their own homologous series, but distinguished by the structure of the ring group. For example, naphthalenes, benzothiophenes, and cyclohexanes are all present in petroleum as homologous series and each constitutes a molecular class. Other molecular classes found in petroleum (Chasey and Aczel, 1991; Boduszynski, 1988; Ronningsen and Skevrak, 1990) are given in Figure 4. As with alkyl groups, the core ring group may also have more than one configuration at a given carbon number. The juxtaposition of rings, the location of heteroatoms within rings, and the carbon number of naphthenic rings (at constant total carbons), are all indistinguishable by mass spectroscopy. The four ring aromatic compound with a mass of 228.30 a.m.u. may have, for example, five isomeric configurations:



In this case, the cata-condensed chrysene is likely to be the most prevalent form based on the origin of ring structures in petroleum (Tissot and Welte, 1984). The source of most ring structures in petroleum is believed to be the terpenes and steroids which exhibit the cata configuration.

In general though, a molecular class's homologous series is composed of both ring isomers and alkyl group isomers at each carbon number, with the only measurable information for the mixture being their molecular masses and their total concentration at each mass. Our working assumption is that the isomers of a molecular class, at a given carbon number, exhibit similar physical, chemical, thermodynamic, and performance properties, since this is the limit of our analytical detail. As the boiling point exceeds 800K, however, there are many more molecular classes, and many of these have masses too close for even high resolution mass spectroscopy to distinguish. These overlapping series may have significantly different chemical structure, reaction chemistry, and physical properties, beyond that of ring or alkyl structural isomers. Thus the molecular classification scheme degenerates further than just the lumping of isomers. In practice, even the 450-800K material must first undergo some chromatographic separations prior to mass spectroscopy to isolate certain homologous series that would otherwise have coincidental masses (Boduszynski, 1988).

Process Chemistry

For more than half a century researchers have accumulated a vast amount of fundamental knowledge on the chemistry and kinetics of petroleum processes, primarily through the study of model compound reactions, but also from real feedstock behavior. This information is significant for the development of refinery process models, but too extensive to review here. Only some essential aspects of refining chemistry are examined here to highlight the necessity for molecular detail in process models.

Conversion processes such as hydrocracking and FCC reduce the boiling point distribution of feedstocks through successive breaking of carbon-carbon bonds. Determining a final boiling point distribution requires an accounting of this process by carbon number and hence at a molecular level. Further, the composition of the lower boiling products in general do not bear resemblance to the feedstock composition for several reasons relating to the complex reaction chemistry. The cracking process tends to generate additional paraffins or olefins from the alkyl chains of the aromatic and naphthenic classes, and from opening of rings on naphthenic compounds. Hydrocracking also converts aromatic rings to naphthenic rings by hydrogenation. In FCC, a reverse reaction converts naphthenic rings to aromatic rings through bimolecular hydrogen transfer reactions. Thus, a significant shift in the distribution of molecular classes resulting from transformations of the core ring groups accompanies the reduction of molecular size.

The composition of the products are also affected by the relative reactivity of molecules. No two molecules have the same reactivity due to steric or electronic differences. The products of those with higher reactivity will tend to concentrate in the lower boiling range. In addition, hydrocarbon molecules have many parallel reaction pathways, or reaction networks, leading to different products at rates dependent on catalyst properties and operating conditions.

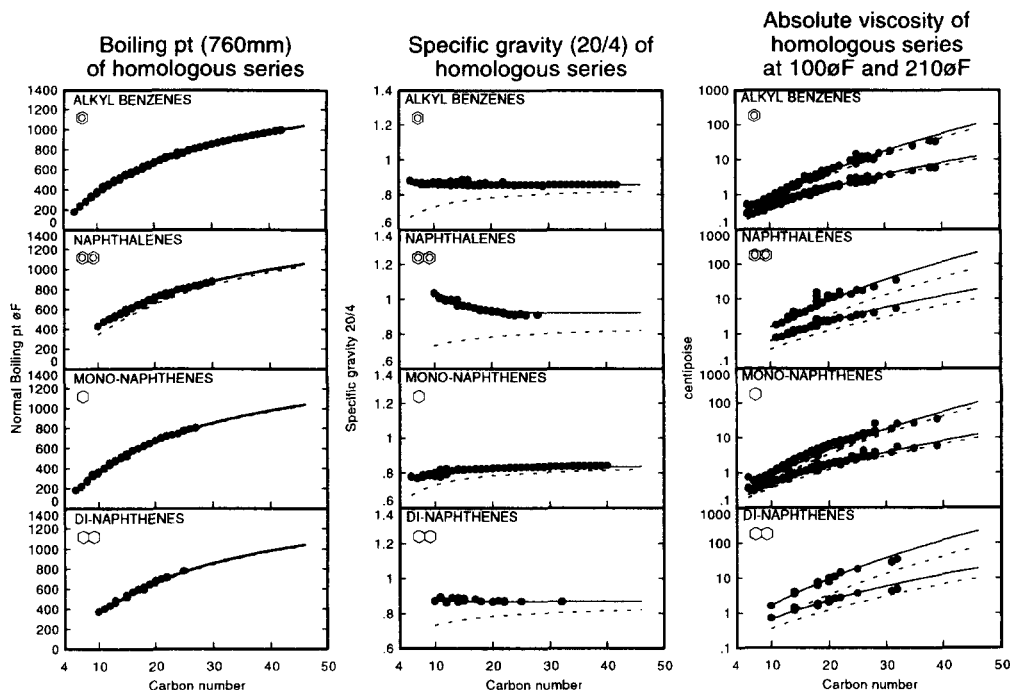


Figure 5. Physical properties of homologous series. Dashed line is the normal paraffin series.

There are also specific aspects to process chemistry that warrant molecular speciation. For example, desulfurization in hydroprocessing operations is strongly dependent on the molecular distribution of sulfur. Sulfide molecular classes are orders of magnitude more reactive than dibenzothiophenes. Even the arrangement of alkyl substituents within the dibenzothiophene class can have a significant effect on reactivity (Ma et al., 1994; Froment et al., 1994). In FCC, the sulfur apportionment to H_2S , gasoline, and cycle oil products is environmentally significant and very dependent on the molecular distribution of sulfur in the feedstock due to complex process chemistry. Another example of the significance of molecular structure is the poisoning effect of certain molecules on catalyst activity (Ho et al., 1992).

Estimating Properties of Petroleum Mixtures

The quality of petroleum derived products are determined by certain mixture properties that characterize combustion quality, fluid characteristics or polluting potential, and include octane number, cetane number, pour point, aniline point, smoke point, cloud point, refractive index, viscosity, vapor pressure, surface tension and sulfur and nitrogen content. Each mixture property is based on the collective properties of the individual molecules. If both the mixture's molecular composition and the molecular properties were available

or predicted, the mixture property could be calculated. One strategy would therefore be to develop molecular structure-property correlations for every property. This approach is certainly feasible with some assumptions or approximations regarding isomeric structures and their impact on molecular properties. However, the petroleum industry literature is replete with empirical and well tested correlations to calculate many of these properties using readily measured mixture properties, primarily the mean normal boiling point (T_b), the specific gravity (spg), and sometimes limited composition data. Critical temperatures, critical pressures and acentric factors required for equation-of-state calculation methods such as the Peng-Robinson or Soave-Redlich-Kwong can also be estimated from boiling point and specific gravity (Lee and Kessler, 1975). With molecular composition and the capability of predicting T_b and spg from this composition, we can calculate many quality and thermodynamic properties of petroleum streams. The alternate strategy is then to develop molecular-based correlations for some essential properties, specifically boiling point and specific gravity, calculate their mixture properties from stream compositions, and estimate all other properties using the wealth of literature correlations. Molecular structure-property correlations for boiling point and specific gravity are also needed for calculating volumetric boiling point distributions from molecular composition.

The homologous series can form the basis of molecular structure-property correlations. Building successful property correlations for hydrocarbons captures the property variation with molecular weight or carbon number and from molecular class to class. Molecular properties generally exhibit a smooth variation with carbon number for a homologous series. Fisher (1982) surveys the various algebraic forms used to represent the variation of properties of *n*-paraffins with chain length. He includes density, refractive index, boiling point, vapor pressure, melting point, dielectric constant, surface tension and solubility parameters.

Extension to other molecular classes is straightforward, using the property of the starting member (core ring group) of the homologous series and successive dilution by the contributions of additional $-CH_2-$ groups as shown in Figure 5. Properties such as viscosity and specific gravity are strongly affected by the molecular class. The ultimate source of pure component property data is the Beilstein Institute for Literature of Organic Chemicals which began publication of its encyclopedia in 1881. This contains data for all relevant molecular classes.

Some mixture properties, however, are not well represented by homologous series correlations, such as the octane number of gasoline (Jaffe, 1981). In this case the composition of the mixture can significantly affect the contributions of individual molecular properties. A molecular approach would enable a proper accounting of non-linear, composition dependent blending phenomena.

3. DESCRIBING THE CHEMISTRY AND KINETICS OF COMPLEX SYSTEMS

The study of the kinetics of petroleum processes has been on two divergent paths for the last sixty years - (1) the fundamental mechanisms in the chemistry and kinetics of model compound reactions and, (2) the modeling of complex mixtures. Unfortunately, the findings of the former could not be incorporated into the framework of the latter. Modeling the kinetics of complex mixtures involved major simplifications that eliminated molecules as a basis. Two approaches received considerable interest and effort - *lumping* and *continuous mixtures*. The lumping approach reduced complexity by grouping the entire set of molecules into a relatively small, manageable number of lumps which could correlate some process yield results (Weekman, 1968; Jacob et al. 1976). The continuous mixture approach treated molecules as a continuum, whose existence, concentration, properties, and change due to reactions were characterized by a continuous function of a measurable variable (DeDonder, 1931; Aris and Gavalas, 1966; Krambeck, 1994; Chou and Prausnitz,

	A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	H	A_A	S	RS	AN	NN	RN	O	RO	O=
Increment Stoichiometry																						
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	1	1	0	0	0	0	0	
N:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
O:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1

Structural Assembly

Structural Assembly

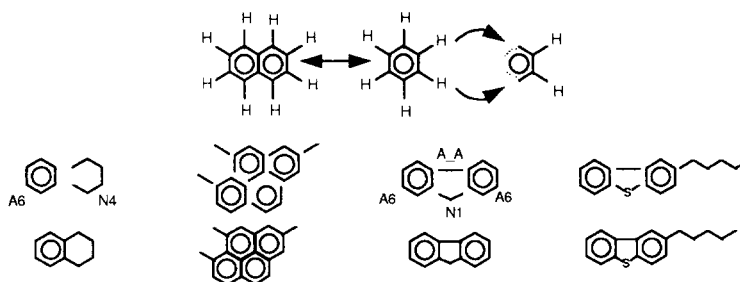


Figure 6. Constructing molecules using Structure Oriented Lumping (SOL) increments.

1986) such as boiling point or carbon number. Neither method incorporates realistic process chemistry or allows for the simultaneous calculation of many properties. Only for the reforming process, where GC methods could measure the low boiling molecular components, were process models using chemistry and molecular properties (Ramage et al., 1987). Over the past several years we developed a method to extend the molecular basis of chemistry and kinetics of complex reaction systems to the entire boiling range of petroleum. We call it Structure Oriented Lumping (Quann and Jaffe, 1992).

Structure Oriented Lumping

Structure Oriented Lumping (SOL) was developed in response to the need for incorporating molecular detail in petroleum chemistry to predict product compositions and properties. It is essentially a group contribution method to describe the structure of molecules, but one that facilitates *both* molecular property estimation and a description of process chemistry. It is also consistent with the limits of our analytical capability to determine molecular detail.

The basic concept of the SOL approach is that any petroleum molecule can be described and represented by a set of certain of structural features or groups. The SOL method organizes this set as a vector, with the elements of the vector representing the number of specific structural features. These vector elements are referred to as *structural increments*, because each element exists as an incremental part of a complete molecule, and not by itself. Molecules are built from these increments. The structure vector and its twenty-two structural increments are shown in Figure 6. The increments consist of three aromatic (A6, A4, A2) and six naphthenic (N6-N1) ring structures, a -CH₂- group (R) to specify the carbon number of alkyl chains or paraffins, bridging between rings (A-A), hydrogen deficiency (H), and heteroatom groups or substitutions in










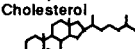

	A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	H	A_A	S	RS	AN	NN	RN	O	RO	O=
2,3,4 Trimethyl pentane																						
	0	0	0	0	0	0	0	0	0	8	3	0	1	0	0	0	0	0	0	0	0	0
Benzene																						
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Naphthalene																						
	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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
Fluoranthene																						
	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																						
	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0
H2O																						

Figure 7. Examples of SOL representation of molecules.

rings and alkyl chains. Note that each group in the vector has carbon, hydrogen, sulfur, nitrogen, and oxygen stoichiometry and thus also molecular weight.

Examples of the use of this vector are shown in Figure 7. The number of increments in each group are indicated by the vector elements. For example, benzene is constructed as a single **A6** ring, with the first element of the vector having a value of one and all others zero. The increments represent what structures are present in a molecule, but not its arrangement. In other words, a vector may represent an ensemble of isomers, all having the same increments. The four ring aromatics discussed previously (chrysene) are an example of identical structure vectors, one **A6** ring and three **A4** rings. We must therefore adopt some conventions in assuming what are likely structures or assume that we can treat the isomer group as a pseudo-molecule for estimating properties. There are two additional groups to distinguish certain isomeric features of molecules or an 'average' isomeric structure. These are number of branches in a chain (**br**) and the number of methyl substituents (**me**) on rings, respectively. Again, the method as shown here doesn't assign the location of branches or methyl groups. We have, however, built additional descriptors into the SOL method to locate branches and methyl groups as needed.

A complex mixture in SOL is represented as a set of vectors, with each molecule or ensemble of structural isomers represented by one of the vectors. The composition of the mixture is expressed in terms of a weight percent for each vector. The method is also consistent with the molecular class/homologous series organization of petroleum, and therefore also with our analytical techniques. The homologous series of a specific molecular class is represented as a set of vectors, each having identical structure increments except for the length of the alkyl chain **R** and possibly the branching and methyl substituent indicators. If we consider

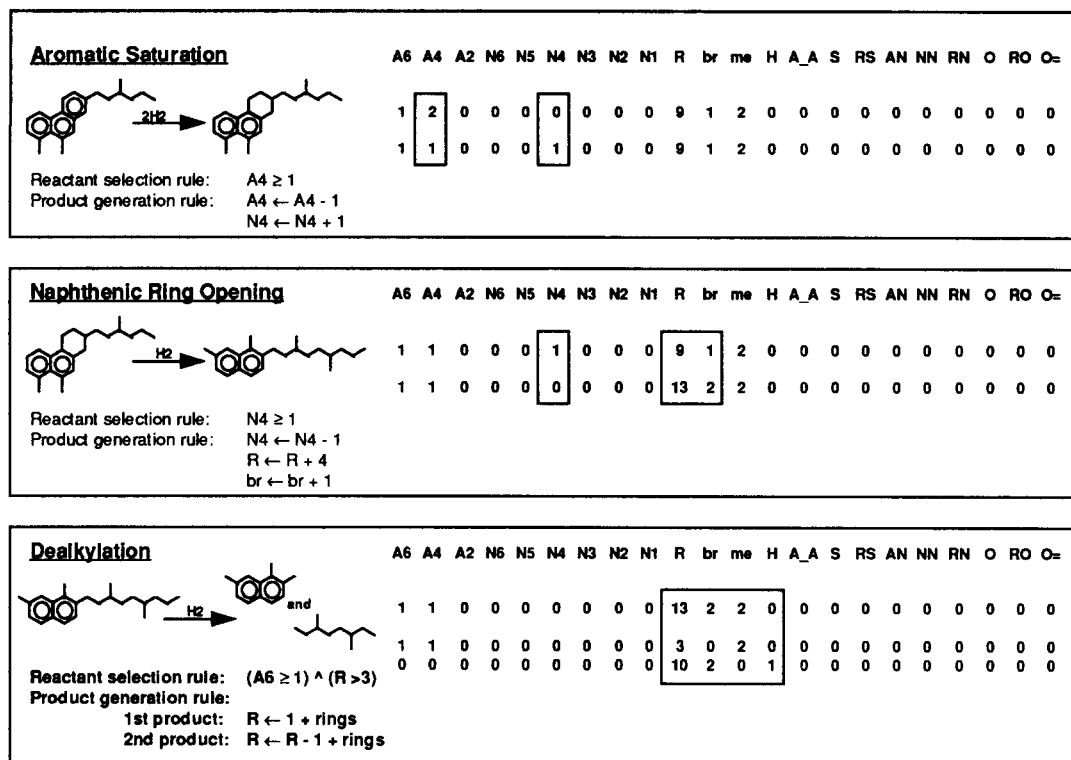


Figure 8. Examples of reaction rules using the SOL approach. The logic expressions are in the APL programming language.

all common ring structures of Figure 4 as homologous series up to about 800K, plus paraffins and olefins, a SOL mixture has about 5000-6000 components.

The questions arise as to why use structural increments, and why these particular increments? First, the increments are convenient and physically relevant for developing molecular property correlations via homologous series of molecular classes. Each increment has a certain contribution to a total molecular property. This is the basis for the property correlations shown in Figures 5. Further, other group contribution methods, including Benson groups for thermodynamic properties and UNIFAC groups for molecular interactions, can be mapped into the SOL vector. Second, and most importantly, these increments correspond to the structural entities that are rearranged in chemical reactions. They allow us to develop a mathematical or algorithmic description of process chemistry.

Using a limited set of structural groups to describe thousands of components enables the use of a limited set of *reaction rules* to establish the rather large number of reactions that constitutes the complex reaction networks of mixtures. We define a *reaction class* as all those molecules in the mixture that undergo the same specific rearrangement of a specific structural increment(s), as determined by the reaction rule. Examples of reaction rules for aromatic saturation, ring opening and dealkylation are given in Figure 8. A reaction rule consists of a (1) *reactant selection rule* to first determine, using logical constructs, if a molecule

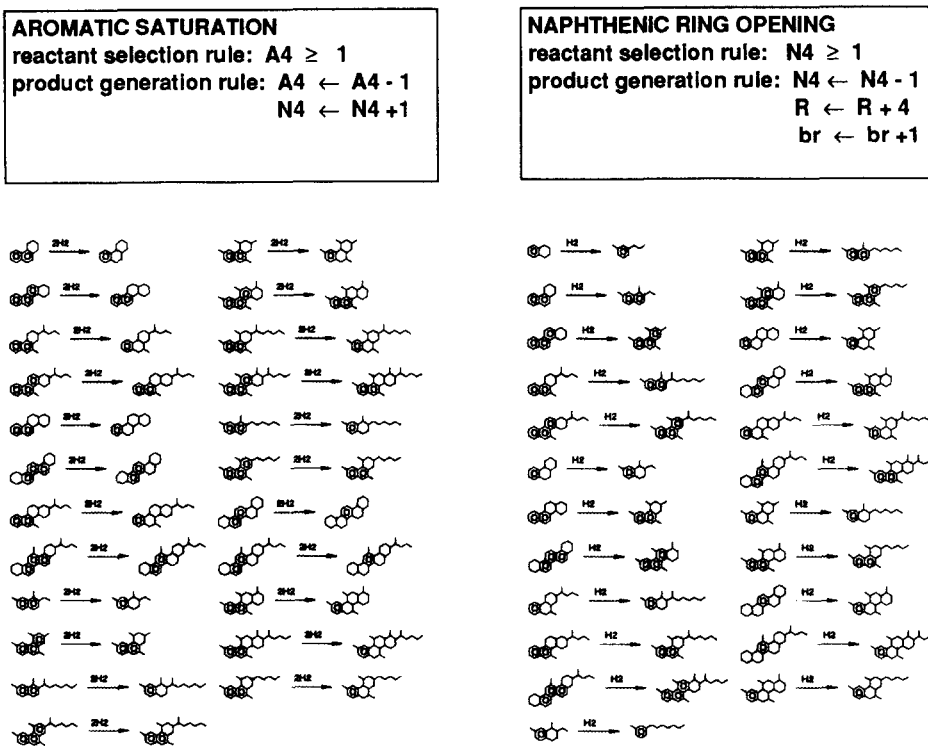


Figure 9. Application of SOL reaction rules to a mixture.

has the increment(s) necessary for the reaction, and (2) a *product generation rule* to convert the reactant structure vector to the product structure vector. For the aromatic saturation reaction shown in Figure 8, the selection rule determines if a molecule has an **A4** ring and the product generation rule converts this ring to an **N4** naphthenic ring through a simple mathematical operation. This particular rule decreases the structure vector's **A4** rings by one and increases the **N4** rings by one.

Applying a reaction rule to all molecules in the mixture generates a reaction class for the specific chemistry represented in the rule, as shown in Figure 9. In general, a reaction rule will find many reactants with the appropriate increments and each with a corresponding product molecule. Further, each molecule may satisfy the criteria of more than one of the reaction rules. This would construct parallel reactions for a molecule. Application of all rules to all components of the mixture generates the entire reaction network representing the chemistry of the process. Application of the rules to just one molecule demonstrates the generation of a small piece of the complex network in Figure 10. Petroleum process chemistry can be typically represented by 20-30 different reaction classes resulting in over 50,000 distinct chemical reactions. Computer programs generating the entire network use sorting procedures to automatically construct the differential rate and energy balance equations for reactor modeling.

An advantage of the SOL approach is that all petroleum processes have the same representation of a complex mixture, i.e. the set of structure vectors. It is the reaction rules that distinguish each process. It is

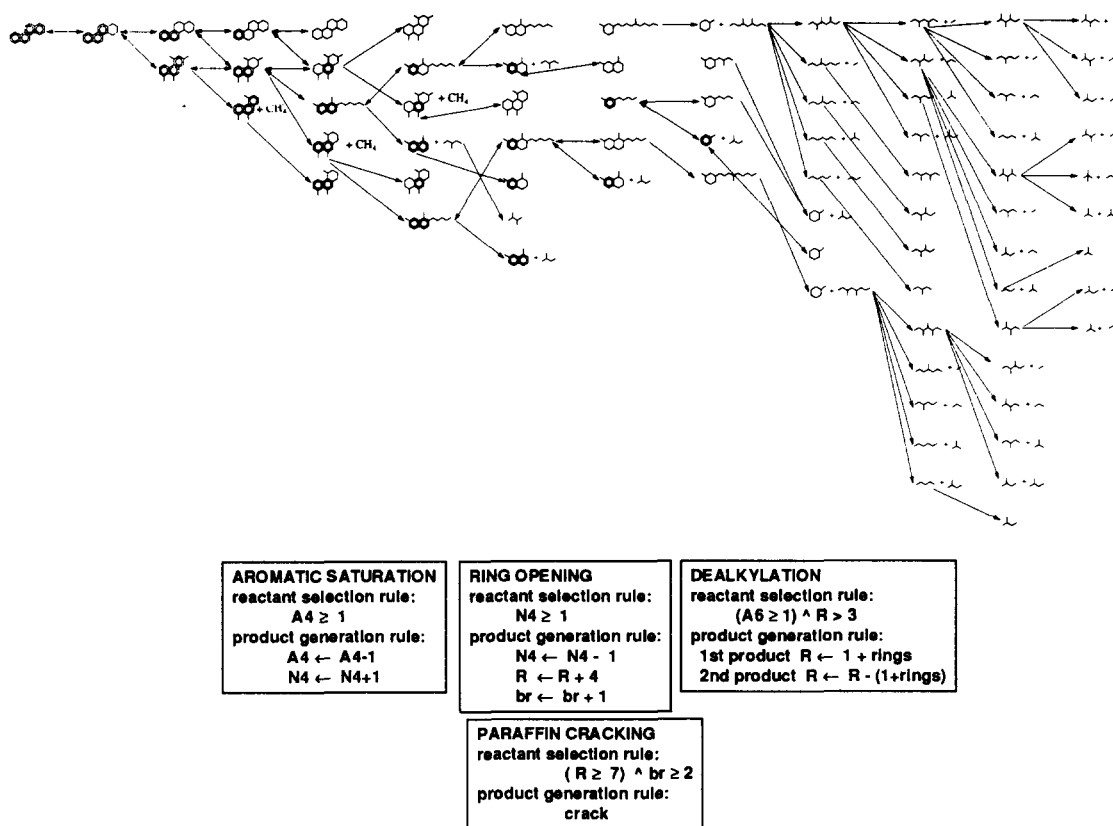


Figure 10. Application of SOL reaction rules to generate a complex reaction network from one starting molecule.

possible therefore to develop a molecular-based model of the entire refinery by integrating individual process models.

Other approaches for molecular modeling of complex mixtures and their reaction networks have been proposed as well. Klein et al. (1991) have developed a stochastic approach for constructing the molecular structure of thousands molecules and their fundamental reactions in the high-boiling portion of petroleum using a set of structural attributes whose frequencies are determined by bulk property data. Liguras and Allen (1989) incorporate fundamental chemistry in an FCC model by using several hundred selected archetypal molecules, called pseudocomponents, whose concentration in a feedstock is calculated from mass spectroscopy and carbon NMR spectroscopy. Froment and coworkers (Clymans and Froment, 1984; Baltanas and Froment, 1985; Hillewaert et al., 1988; Svoboda et al., 1995) have taken a most detailed approach by using a Boolean matrix representation of a molecule and its complex sequence of elementary reaction steps involving radicals in thermal reactions or surface intermediates in catalysis. A reaction is decomposed into a series of elementary steps, and elementary steps into 'single-events'. Ultimately, all reactions result from some combination of a limited set of single-event possibilities.

The SOL approach and the other schemes described above employ different strategies in representing molecular structure, mixture composition and process chemistry, but with a common principle. That principle is to reduce complexity by using fundamental chemistry, not by drastically reducing the number of components

in the mixture through lumping. The same limited set of fundamental reactions, elementary steps, or single events happen again and again in the complex reaction network of a hydrocarbon mixture because of the repeating occurrence of molecular structural groups.

Determining the Kinetics

Now that the emphasis of modeling is on process chemistry and incorporating its details, we must find the kinetic parameters of chemical (catalytic) reactions. The required kinetic parameters include rate constants, activation energies, and adsorption constants on catalyst sites, as a function of process conditions and catalyst properties and for real feedstocks. Obviously, this information will not come from an experimental study of 50,000 reactions. We must return to the SOL concept of the limited number of reaction classes.

The first approximation is to assume that each reaction in a class has the same kinetic parameters since they undergo the same intramolecular rearrangement. This is generally insufficient and a perturbation on reaction rates with molecular structure will be required within each class. Of the numerous model compound studies of petroleum process chemistry, including those with emphasis on reaction kinetics, few have examined molecular structure effects on rate of a reaction. Much of the past research in this area has focused on how catalyst properties influence the reaction rates of a single molecule, not on how molecular structure influences reaction rates of different molecules.

Structure-reactivity relationships in chemistry are not new. Hammett (1937) organized reaction families and substituent effects for homogeneous systems. The Bronsted and Evans-Polyani relationships are other well-known examples. The premise behind structure-reactivity relationships is that the rate constants for a reaction class can be correlated with a calculable molecular property of the molecules. One form of the relationship is

$$\ln k_i = a + b \text{RI}_i$$

where k_i is the rate constant for molecule i , a and b are constants to be determined, and RI is the reactivity index for molecule i . This equation is known as a linear free energy relationship (LFER) because it can be derived from transition-state theory, where RI coincides with the differences in the free energy of the transition-states for the reactants, and correlates with some calculable molecular property. If constants a and b can be determined (experimentally) from several reactions, the rate constants for all members of the reaction class can be calculated from reactivity indices. The choice of RI reflects the controlling mechanism of the reaction, such as carbenium ion stability in cracking chemistry, free radical stability in thermal reactions, or electronic properties of aromatic molecules in hydrogenation reactions (Neurock and Klein, 1993). Mochida and Yoneda (1967) first demonstrated the use of LFERs for catalytic reactions. Klein and coworkers (Korre et al., 1994; Broadbelt et al., 1994) have used the LFER extensively for correlating rates of catalytic hydrogenation and cracking reactions and for free-radical chemistry. Structure-adsorption relationships have also been developed for adsorption and inhibition of molecules on catalytic sites (Ho et al., 1994).

The single-event approach pioneered by Froment and coworkers (Baltanas et al., 1989; Vynckier and Froment, 1991) enables the kinetics of reactions in a large network to be calculated from only a few elementary step rates, which in turn are combinations of more basic molecular rearrangements. For example, only 17 single-event rate coefficients are needed to calculate the rates of over 400 elementary steps involving reactants, products and surface intermediates in the hydrocracking of n -octane (Svoboda et al., 1995). These single-event parameters are then applicable to the elementary reaction steps of other molecules as well.

Limited experimental work can thus provide the kinetic information required for larger, more complex reactions systems.

The use of structure-reactivity relationships or single-event methods to estimate reaction rates presumes that the exact structures of the reactants, products and their intermediates are known. Exact structures are, of course, generally unavailable and all isomers are not identified or included in modeling real feedstocks containing 400F+ material, as discussed earlier. These reactivity estimation techniques need to be developed for isomeric ensembles of molecules. The homologous series concept used in the development of physical property correlations may also find application here in developing structure-reactivity relationships. We must also contend with sterically hindered isomers, as in the case of methylated dibenzothiophenes in desulfurization reactions. The significantly reduced reaction rates of the sterically hindered isomers in a reaction class are difficult to predict. Both the amount of such isomers and their kinetic parameters must be determined experimentally. Whatever the method for assigning rate constants, it must result in accurate prediction of product yields for real feedstocks.

4. DELIVERY TECHNOLOGY

Process models are developed for two purposes: (1) to provide insight for research and development, and (2) to provide a useful tool for economic optimization of commercial operations. We will address the issues of the latter, less understood application of process modeling and develop the criteria for a working model.

A commercial operation in a petroleum refinery receives a feedstock and generates products composed of molecules. The operations staff, however, have no direct knowledge, nor use for that matter, of the molecular constitution of feedstocks and product streams or of the complex chemistry. They take measurements of flowrate, pressure and temperature directly, and sample feed and product streams for later laboratory analysis of boiling point distributions, impurities (i.e. total sulfur), quality properties, and sometimes light component (C_1 - C_5) yields to assess and optimize process performance. The process model should have the same configuration and operation as the commercial unit and simulate the same on-line and laboratory information. In other words, the model's working foundation is a molecular representation of streams and complex reaction networks, but its function and output should be familiar to the users and designed for their needs.

Petroleum process models and their user-interface should have three features: (1) feedstock selection, (2) the model's representation of the process, and (3) product characterization. Most petroleum processes have feedstocks whose composition can vary on a daily basis because of variations in crude source or upstream operations. Model feedstocks need to be presented to the user and selected on the basis of conventional characterization methods available in the refinery, e.g. feed source, boiling point, sulfur content, specific gravity, etc.. Algorithms hidden from the user translate the characterization data into model feedstocks having molecular composition. The source of the molecular compositions is a library of previously analyzed feedstocks.

11

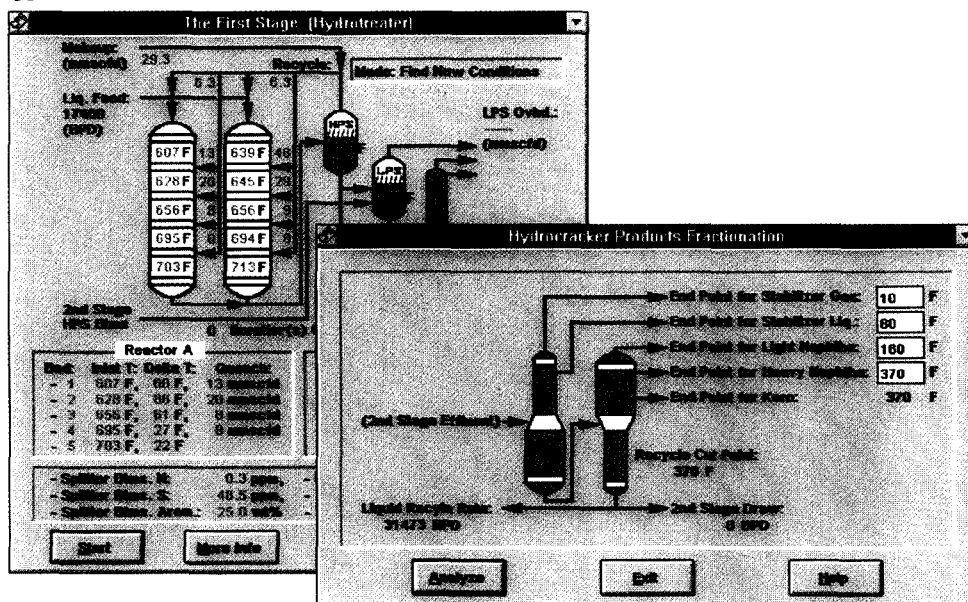


Figure 11. User interface for a model of a commercial hydroprocessing operation.

12

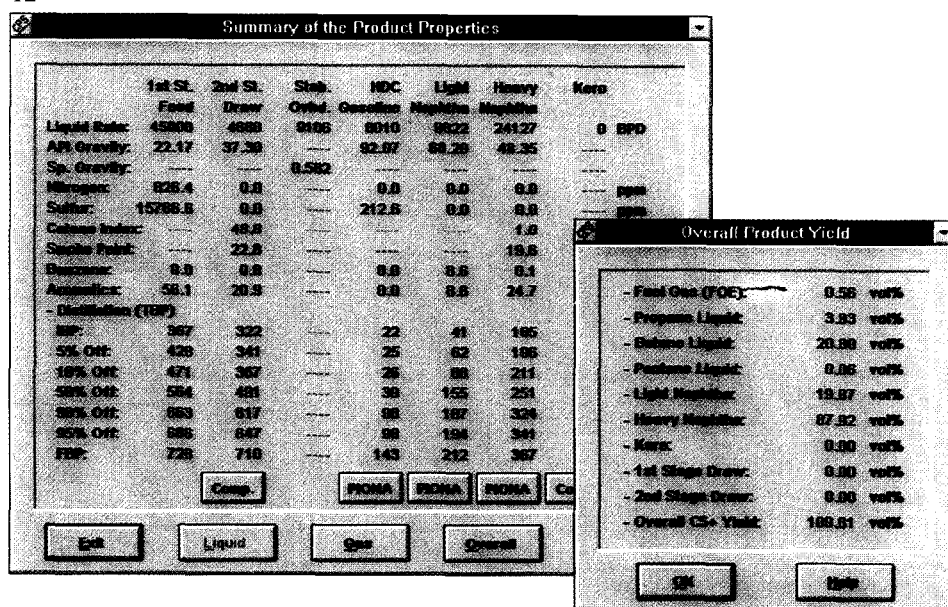


Figure 12. User interface reports model results in familiar formats.

The process model should be configured to operate the same way as the commercial process, with the same set of integrated unit operations as shown in Figure 11 for a specific hydroprocessing operation. The model tracks all of the molecular transformations in reactor beds, their disposition in separators and distillation units, and their concentration in various process streams. The molecular property correlations allow a compression of this molecular detail into relevant operational information such as liquid and gas stream flow rates, reactor temperatures, exotherms, quench gas requirements and impurity levels. Similarly, the model's output streams are composed of thousands of molecular components, but results are reported in the format of standard characterization methods, overall yields, and quality specifications as shown in Figure 12. Molecular compositions of streams are obviously available if needed.

A model must also incorporate other features specific to the commercial operation. Most processes operate under constraints such as flow rate, hydrogen availability, reactor pressure drop, flow rates, metallurgical limits, to name a few. The user must have the capability to specify this information and the model must determine if they are exceeded.

Catalytic processes present a significant challenge. Kinetics determined in the laboratory may not be representative of commercial operation due to catalyst differences, catalyst deactivation or reactor physical factors such as hydrodynamics. The model must have the capability of adjusting or tuning kinetic parameters to match commercial performance. Confidence in a model's ability to predict future operation is based on its accuracy in matching current operation.

5. CONCLUSIONS

Early models of complex reaction systems in petroleum refining drastically simplified the representation of mixture composition and chemistry, resulting in models with limited scope and reliability. Process models today are required to predict not just the yields of major products, but also the composition and quality of products, detailed boiling point distributions, and commercial operating data. The best approach has a molecular basis, for molecules are the common foundation for property calculations, process chemistry, feedstock composition variability, reaction kinetics and thermodynamics.

The difficulty with a molecular approach is that the exact composition and even exact molecular structures can not be determined for most of the petroleum. We are forced to work with identifiable classes of molecules and their ensemble of isomers. Fundamental process chemistry and molecular property and reactivity correlations can be built into the structure of the model, but must be tempered with the uncertainties in petroleum's molecular structure.

Acknowledgments - The authors thank L. A. Green and W. T. Mo for their assistance with this manuscript.

REFERENCES

- Altgelt, K. H. and Boduszynski, M. M., 1994, *Composition and Analysis of Heavy Petroleum Fractions*. Marcel Dekker, Inc.
- Aris, R. and Gavalas, G., 1966, On the theory of reactions in continuous mixtures. *Proc. Royal Soc., London*, Vol. 260, 351-393.

- Baltanas, M. A. and Froment, G. F., 1985, Computer generation of reaction networks and calculation of product distribution in the hydroisomerization and hydrocracking of paraffins on Pt-containing bifunctional catalyst. *Comput. Chem. Eng.* **9**(1), 71-81.
- Baltanas, M. A., Van Raemdonck, K. K., Froment, G. F. and Mohedus, S. R., 1989, Fundamental kinetic modeling of hydroisomerization and hydrocracking on noble-metal-loaded faujasites. *Ind. Eng. Chem. Res.* **28**, 899.
- Boduszynski, M. M., 1988, Composition of heavy petroleum. 2. molecular characterization. *Energy & Fuels* **2**(5), 597-613.
- Broadbelt, L. J., Stark, S. M. and Klein, M. T., 1994, Computer generated pyrolysis modeling: on-the-fly generation of species, reactions, and rates. *Ind. Eng. Chem. Res.* **33**, 790-799.
- Chasey, K. L. and Aczel, T., 1991, Polycyclic aromatic structure distributions by high-resolution mass spectroscopy. *Energy & Fuels* **5**(3), 386-394.
- Chou, G. F. and Prausnitz, J. M., 1986, Adiabatic flash calculations for continuous or semicontinuous mixtures using an equation of state. *Fluid Phase Equil.* **30**, 75.
- Clymans, P. J. and Froment, G. F., 1984, Computer generation of reaction paths and rate equations in thermal cracking of normal and branched paraffins. *Comput. Chem. Eng.* **8**(2), 137.
- Dedonder, Th., 1931, *L'Affinite*, Second Partie, Gauthier - Villars, Paris.
- Fisher, C. H., 1982, Equations correlating *n*-alkane physical properties with chain length. *Chemical Engineering*, Sept. 20.
- Froment, G. F., Depauw, G. A. and Vanrysselberghe, V., 1994, Kinetic modeling and reactor simulation in hydrodesulfurization of oil fractions, *Ind. Eng. Chem. Res.*, **33**, 2975-2988.
- Hammett, L. P., 1937, The effect of structure upon the reactions of organic compounds. benzene derivatives. *J. Am. Chem. Soc.* **59**, 96-103.
- Hillewaert, L. P., Dierickx, J. L. and Froment, G. F., 1988, Computer generation of reaction schemes and rate equations for thermal cracking. *A.I.Ch.E. Journal* **34**(1), 17-24.
- Ho, T. C., Katritzky, A. R. and Cato, S. J., 1992, Effect of nitrogen compounds on cracking catalysts. *Ind. Eng. Chem. Res.* **31**, 1589-1597.
- Hood, A., Clerc, R. J. and O'Neal, M. J., 1959, The molecular structure of heavy petroleum compounds. *J. Inst. Petrol.* **45**(426).
- Jacob, S. M., Gross, B., Voltz, S. E. and Weekman Jr., V. W., 1976, A lumping and reaction scheme for catalytic cracking. *A.I.Ch.E. Journal* **22**(4), 701-713.
- Jaffe, S. B., 1981, Control of gasoline manufacture. US Patent 4,251,870.
- Klein, M. T., Neurock, M., Nigam, A. and Libanati, C., 1991, Monte Carlo modeling of complex reaction systems: an asphaltene example. In *Chemical Reactions In Complex Mixtures, The Mobil Workshop*, Sapre, A. V. and Krambeck, F. J., Eds., 126-142, Van Nostrand Reinhold.
- Kohnen, M. E. L., Shinninghe Damste, J. S., Rijpstra, W. I. C. and deLeeuw, J. W., 1990, Alkylthiophenes as sensitive indicators of palaeoenvironmental changes. In *Geochemistry of Sulfur Fuels*, Orr, W. L. and White, C. M., Eds., ACS Symposium Series 429, American Chemical Society, Wash. D. C., 444-485.
- Korre, S. C., Neurock, M., Klein, M. T., and Quann, R. J., 1994, Hydrogenation of polynuclear aromatic hydrocarbons. 2. quantitative structure/reactivity correlations. *Chem. Eng. Sci.* **49**(24A), 4191-4210.
- Krambeck, F. J., 1994, Thermodynamics and kinetics of complex mixtures. *Chem. Eng. Sci.* **49**(24A), 4179-4189.
- Lee, B. I. and Kesler, M. G., 1975, A generalized thermodynamic correlation based on three parameter corresponding states. *A.I.Ch.E. Journal* **21**, 510.

- Liguras, D. K. and Allen, D. T., 1989, Structural models for catalytic cracking. 1. model compound reactions. *Ind. Eng. Chem. Res.* **28**, 665-673.
- Ma, X., Sakanishi, K. and Mochida, I., 1994. Hydrodesulfurization reactivities of various sulfur compounds in diesel fuel. *Ind. Eng. Chem. Res.* **33**, 218-222.
- Mair, B. J. and Barnewall, J. M., 1964, Composition of the monoaromatic material in the light gas oil range, low refractive index portion, 230 to 305C. *J. Chem. Eng. Data* **9**(2), 282.
- Mochida, I. and Yoneda, Y., 1967, Linear free energy relationships in heterogeneous catalysis. 1. dealkylation of alkylbenzenes an cracking catalysts. *J. Cat.* **7**, 386-392.
- Neurock, M. and Klein, M. T., 1993, Linear free energy relationships in kinetic analysis; applications of quantum chemistry. *Polyc. Arom. Compounds* **3**, 231-246.
- Ostoukov, S. B., Aref'yev, O. A., Pustil'nikova, S. D., Zabrodina, M. N. and Petriv, Al. A., 1984, C₁₂ - C₃₀ *n*-alkylbenzenes in crude oil. *Petrol. Chem. U.S.S.R.* **23**(1), 1-12.
- Quann, R. J. and Jaffe, S. B., 1992, Structure-oriented lumping: describing the chemistry of complex mixtures. *Ind. Eng. Chem. Res.* **31**(11), 2483-2497.
- Ramage, M. P., Graziani, K. R., Schipper, P. H., Krambeck, F. J. and Choi, B. C., 1987, KINPIR (Mobil's kinetic reforming model): a review of Mobil's industrial modeling philosophy. *Advances in Chemical Engineering* **13**, 193.
- Ronningsen, H. P. and Skjevrak, I., 1990, Characterization of North Sea petroleum fractions: aromatic ring class distributions. *Energy & Fuels* **4**(5), 608-626.
- Svoboda, G. D., Vynckier, E. Debrabandere, B. and Froment, G. F., 1995, Single-event rate parameters for paraffin hydrocracking on a Pt/US-Y zeolite. *Ind. Eng. Chem. Res.* **34**, 3793-3800.
- Tissot, B. P. and Welte, D. H., 1994, *Petroleum Formation and Occurrence*, Springer-Verlag.
- Vynckier, E. and Froment, G. F., 1991, Modeling of the Kinetics of Complex Processes Based Upon Elementary Steps. In *Kinetic and Thermodynamic Lumping of Multicomponent Mixtures*. Astarita, A. and Sandler, S. I., Eds., 131-162, Elsevier.
- Weekman, V. W. Jr., 1968, A model of catalytic cracking conversion in fixed, moving, and fluid-bed reactors. *I.E.C. Proc. Des. Dev.* **7**, 90.