KNOWLEDGE-BASED EXPERT SYSTEMS: AN EMERGING TECHNOLOGY FOR CAD IN CHEMICAL ENGINEERING

by

D. Sriram, R. Banares-Alcantara, V. Venkatasubramanian,
A. Westerberg and M. Rychener

DRC-06-76-84

December, 1984

Knowledge-Based Expert Systems: An Emerging Technology for CAD in Chemical Engineering

D. Sriram, R. Bañares-Alcántara, V. Venkatasubramanian,
A. Westerberg, and M. Rychener

Design Research Center,
Carnegie-Mellon University,
Pittsburgh, PA 15213

ABSTRACT

Knowledge-Based Expert Systems (KBES) systems deal with difficult, *ill-structured,* problems in complex domains for which no straight-forward, *algorithmic,* solutions exist. The solution process often involves skillful manipulation of large quantities of knowledge, in a trial and error fashion, starting out with certain assumptions and hypotheses and revising them when it is necessary till the solution is achieved. A number of problems encountered by chemical engineers can be efficiently solved using the KBES approach. This paper focuses on a number of issues regarding the application of this emerging field to chemical engineering problems.

KBES: An Emerging Technology for CAD In Chemical Engineering

1. INTRODUCTION

Artificial Intelligence (AI) is concerned with the development of computer programs that emulate the intelligence of humans. In this respect AI is deeply concerned with the understanding of human problem-solving strategies, in particular the problem-solving in specified domains by experts. Expert problem-solving involves solving problems using large amounts of specialized knowledge, called *domain knowledge* using rules of thumb (called *heuristics)* learnt and refined over years of problem-solving experience in that domain. The amount of knowledge manipulated is often vast and the expert rapidly narrows down the search by recognizing patterns and thereby using the appropriate heuristics. Designing a computer program to do this is the study of *Knowledge Based Expert Systems (KBES),* an important practical aspect of AI which has recently caught a lot of attention because of its enormous potential to solve problems in the engineering industry.

KBES systems deal with difficult, *ill-structured,* problems in complex domains for which no straight-forward, *algorithmic,* solutions exist. The solution process often involves skillful manipulation of large quantities of knowledge, in a trial and error fashion, starting out with certain assumptions and hypotheses and revising them when it is necessary till the solution is achieved. Although a number of papers have been published on the applications of KBES in engineering [29], there has not been much activity in the chemical engineering field. However, the potential of KBES for chemical engineering applications has been recognized by a number of researchers [19,26,32], The purpose of this paper is to present various aspects of this emerging field of CAD to chemical engineers.

The organization of a KBES is described in Section 2. Section 3 discusses the range of applications for KBES in chemical engineering. An example is provided in Section 4. This is followed by a discussion of existing systems in Section 5 and a few potential applications are outlined in Section 6. Section 7 deals with the development cycle of a KBES.

UNIVERSHYUBRAWES
CARNEGJE-MELUNf UNIVERSITY
PITTSBURGH; PENNSYLVANIA 15213

2. KNOWLEDGE-BASED EXPERT SYSTEMS

KBES are interactive computer programs incorporating human expertise developed over a number of years and provide advice on a wide range of tasks. These systems typically consist of the following three components (see Figure 2-1):

1. Knowledge-base is a collection of general facts, rules of thumb and causal models of behavior of the problem domain. A number of formalisms have been used to represent knowledge and the most widely used one is the *production system* model. In this formalism, the knowledge is encoded in the form of antecedent-consequent pairs or IF-THEN rules and uncertainty in the knowledge is represented by means of confidence factors [30]. Other forms of representations commonly used are logic and frame-based schemes. The application of these schemes to engineering design is addressed in [18].

2. Context is a workspace for the problem constructed by the inference mechanism from the information provided by the user and the knowledge-base.

3. Inference Engine is used to monitor the execution of the program by using the knowledge-base to modify the context. A number of problem solving strategies exist in current expert systems. A detailed description of these strategies can be found in [24,31].

Apart from three main modules described above the system should also be provided with a graceful user interface, explanation facility and a knowledge-acquisition module. For a more detailed discussion of the organization of KBES, the reader is referred to [13, 34].

Conventional Programs vs KBES

Conventional programs consist of a set of statements whose order of execution is predetermined. These programs are very inflexible and any updates to the program needs considerable effort, because the programmer has to locate the appropriate place to update in the predefined sequence. The programmer must ensure *completeness,* that is the program performs the correct actions for all possible combination of conditions, and *uniqueness,* that is the output must be unique for a certain set of conditions, of the program.

KBES alleviate the problems posed by conventional programs by making a clear distinction between the knowledge-base and the control strategy. This allows for incremental addition of knowledge, without manipulating the overall program structure; the programmer need not guarantee *completeness.* Further, by associating confidence factors with the IF-THEN rules the system can be made to provide a number of alternate solutions, ranked on the basis of their likelihood, to a set of input conditions; thus relaxing the *uniqueness* constraint.

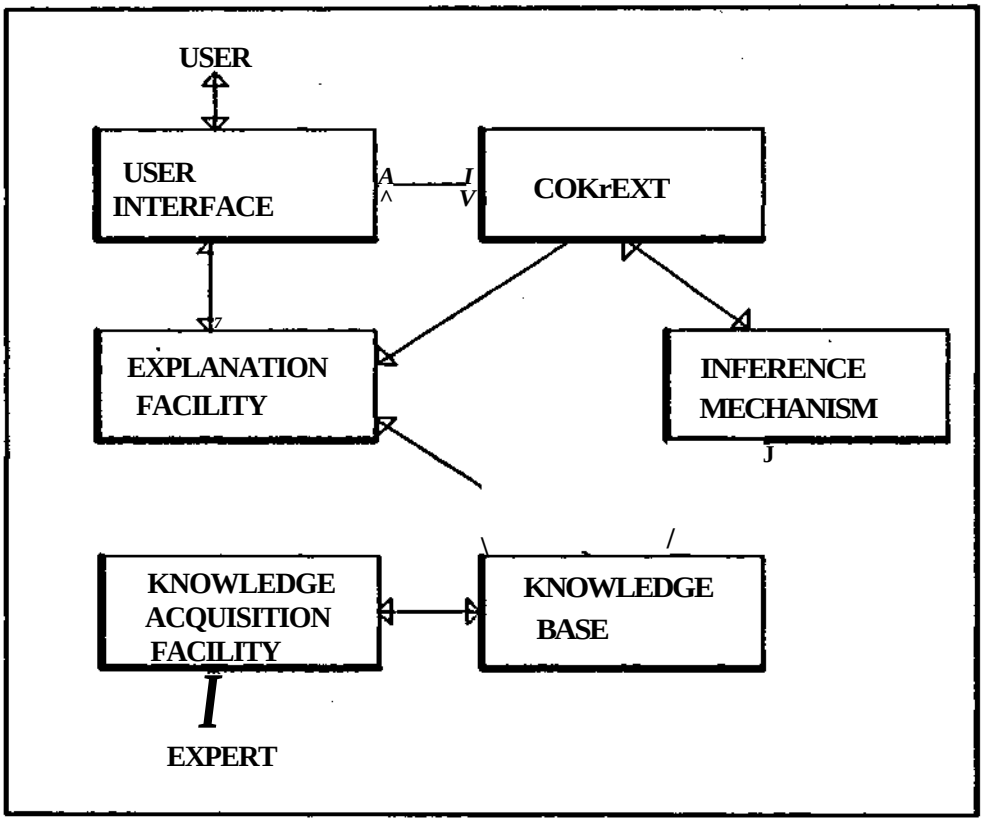


Figure 2-1: A Schematic View of a Complete Knowledge-Based Expert System [28]

3. RANGE OF KBES APPLICATIONS

The range of KBES applications in chemical engineering (or any other branch of engineering) can be bounded by the *derivation-formation* spectrum. In derivation problems, the problem conditions are specified as parts of a solution description[1] (the possible outcomes exist in the knowledge-base). The knowledge base is used to complete the solution. Essentially, the solution to these problems involves the identification of the solution path. In formation problems, problem conditions are given in the form of properties that a solution must satisfy as a whole. Most real-life problems fall between these two categories. Chemical engineers normally encounter the following types of problems at the derivation end of the spectrum [13]:

1. Interpretation. The given data are analyzed to determine their meaning. The data are often either unreliable or erroneous or extraneous. Hence the system should be able to eliminate candidates based on incomplete information.

2. Diagnosis. The problem areas or faults are identified based on potentially noisy data. The

[1] William Ctancey [6J uses the term classification for derivation type of problems. However, in the present work classification-type problems are grouped with diagnosis-type problems.

diagnostisian must be able to relate the symptoms to the appropriate fault. The task may involve reasoning based on incomplete and inexact data, faulty sensors, causal interpretation of the system.

3. Repair. A set of actions to rectify the faults in the system. A first step in this task is to diagnose the problem area.

4. Monitoring. Signals are continuously interpreted and alarms set whenever required.

Problems encountered at the formation end of the spectrum are:

1. Planning. A program of actions are setup to achieve certain goals. The course of actions should be setup such that excessive resources are not expended or constraints are not violated.

2. Design. Systems or objects that satisfy particular requirements are configured. This involves satisfying constraints from a variety of sources. Large design problems are usually solved by dividing the problems into a number of subtasks. The designer must be able to handle the interactions between these tasks properly.

4. AN EXAMPLE

A typical case of what can be expected from an interaction with an expert system is exemplified with an illustration of CONPHYDE (CONsulant for PHYsical property DEcisions) [3], a prototype expert system whose area of expertise is the selection of the most appropriate thermodynamic model for the evaluation of physical properties in vapor-liquid equilibrium calculations. The system is constructed using the general purpose utilities of KAS [23], a domain independent expert system framework. Although, most of the knowledge in CONPHYDE was the author's interpretation of text-book knowledge, it serves to illustrate various components of a KBES.

In the following examples the user's answers and commands are underlined boldface italics (e.g. ***RUN***, ***YES***, ***2***, etc.), and comments about specific features of the system are enclosed between curled parenthesis and in italics (e.g. {*This is a comment*}). The text printed by CONPHYDE is the rest. Section 4.1 shows the interaction of the user with CONPHYDE. The explanation facility is described in Section 4.2, while the process of knowledge acquisition by the system is addressed in Section 4.3.

4.1. Interaction with the User

The user and the system communicate by means of *certainty measures*. These measures range from ***-5.0***, which indicates that the assertion is false; to ***+5.0***, which indicates a truth value for the assertion. A value of ***0*** means that there is no information about the assertion. The system also has the capability of asking questions that can be answered with a **YES/NO** reply, and questions that must be responded to with a numerical value and its corresponding certainty (e.g. value of a polarity,

a temperature, etc.).

```
>> RUN
--------------------
This version of CONPHYDE contains models for the following subsystems:

1 A physical property prediction method that uses activity coefficients
  in the liquid phase. (the parameters of the model are estimated, it
  contains the ASOG, UNIFAC and Regular Solution Theory)
2 A physical property prediction method that uses activity coefficients
  in the liquid phase with adjustable parameters. (contains the
  Margules, Van Laar, Wilson, UNIQUAC and NRTL methods)
3 A physical property prediction method that uses equations of state of
  the virial type (contains the Virial equation of state)
4 A physical property prediction method that uses equations of state of
  the multiparametric type (contains the Benedict Webb & Rubin, its
  Starling modification and the Lee Kesler equations)
5 A physical property prediction method that uses equations of state of
  the Van der Waals type (two parameters).(contains the Soave and
  Peng & Robinson modifications of the Redlich-Kwong equation)

Different components have properties that make the use of a prediction
method appropriate or not.  To help in selecting a subsystem to focus
on, it would be useful for me to know any characteristics of the
chemical system that are immediately available.

        {Volunteering stage}

Do you have information about the conditions?
(Not in operation now).                          NO
Do you have information about components present in the mixture?
(This can include items that are uncertain or vague.)          YES

--------------------
     Please give a component and its associated certainty value.
       (If you need instructions, type HELP; terminate by typing DONE.)
--------------------
1 system contains hydrogen. (RULE 57). TAXONOMY: HYDROGEN
2 the mixture has a wide boiling temperature range (e.g. mixtures
that contain quantum gases etc.). (RULE 56). TAXONOMY: QUANTUM GASES.
3 There are electrolytes present in the mixture. (RULE 14). TAXONOMY:
ELECTROLYTE.
4 Polymers are present in the mixture.  (RULE 13). TAXONOMY: POLYMER.
--------------------
1:TRACE {It is possible to trace which nodes are affected by the user's answers }

OK
1:1 2 {Item 1 is present with a certainty of 2
        (i.e. we are more or less sure that hydrogen is present)}

 Changing the certainty of (* system contains hydrogen. (RULE 57).
TAXONOMY: HYDROGEN )
        from 0.0 to 2.0
Changing the score for:  SPACE-BENEDICT.BENEDICT  from  0.0  to  2.0
```

Changing the certainty of (•
tha use of the Soave eqn. of state 1s advisable)
from 0.0 to -.798864

Changing the certainty of (*
the use of the Benedict Webb and Rubin equation of state 1s advised)
from 0.0 to -.8

Changing the certainty of (*
the mixture has a wide boiling temperature range (e.g. mixtures
that contain quantum gases etc.). (RULE 66). TAXONOMY: QUANTUM GASES.)
from 0.0 to 2.0

{andsoon...}

21*NOTRACE {stoptracing}*

OK
2:2 1 *{Quantum Gases present with a certainty of 1, the fact that hydrogen is a subset of quantum gases is contained in CONPHYDE's database. CONPHYDE checks for consistency and informs the user of his error.}*
** This certainty seems Inconsistent with your previous statements:
1 — 2) system contains hydrogen. (RULE 67). TAXONOMY: HYDROGEN
(certainty 2.0)

Do you want to changes the certainty for Statement 2 ? *YES*
New certainty * ? *2*
2) the mixture has a wide boiling temperature range (e.g. mixtures
that contain quantum gases etc.). (RULE 66). TAXONOMY: QUANTUM GASES.
(certainty 2.0)
Z:DQNE

Presently, the subsystems 1n CONPHYDE and their certainties are:
1) BENEDICT model (certainty 1.6)
2) VANDERWAALS model (certainty 0.0)
3) PREDICTABLE model (certainty 0.0)
4) VIRIAL-TOP model (certainty 0.0)
6) ADJUSTABLE model (certainty 0.0)
On the basis of the Information you gave me, I would work first
on the first hypothesis.

Which (1f any) of these hypotheses would you prefer to pursue? jf
OK. (Use the STAY or SWITCH commands to Intervene.)

3 -- To what degree do you beiieva that the use of equations of state
1s preferred over the use of activity coefficient methods ? 0

{Sincetheuserhasnodirectinformation CONPHYDEtriesto establishitfrommorebasicevidences(seeVANDERWAALSmodel).}

4 -- To what degree do you believe that the system 1s entirely composed
of a mixture of nonpolar or slightly polar compounds, normal-fluids
(rare gases, N2, $O2_f$ CO or hydrocarbons), C02, H2S or H2. ? *zl*

{CONPHYDE switches its current hypothesis and continues.}

```
18 -- If polar components are present In the mixture, what 1s the
polarity of the most polar of the components? (Answer with a range 1n
polarity values, e.g. for slightly polar components the answer would
be: 0.0  0.1) (RULE 12)  ? _L3 JL5 {Range of polarities of most polar component}
What 1s your confidence 1n this estimate? 2J>

20 -- Polymers are present 1n the mixture.  (RULE 13).
TAXONOMY:  POLYMER.      ? NO   {Some nodes can be asked with a YES/NO question.}

          .   {and  soon...}
          .
```

At the end of a session it is possible to ask for a list of all the nodes (or hypotheses) in the system with their corresponding certainty values. Following is the extracted list of all the physical prediction methods rated by CONPHYDE according to the example session:

```
PENG-ROBINSON<VANDERWAALS>  4.91847
SOAVE<VANDERWAALS>  4.648716
UNIQUAC<ADJUSTABLE> 4.44114
LKPKP<BENEDICT> 3.90707
MARGULES-KADJUSTABLE>  3.069636
WILSON<ADJUSTABLE>  2.926204
MARGULES-2<ADJUSTABLE>  2.611833
VAN-LAAR <ADJUSTABLE>  2.611833
NRTL<ADJUSTABLE>  0.2618606
STARLING<BENEDICT>  -.498039
VIRIAL<VIRIAL-TOP>  -2.68096
ASOG<PREDICTABLE>  -3.22088
UNIFAC<PREDICTABLE>  -3.22088
REGULAR-SOLN<PREDICTABLE>  -3.27887
BWR<BENEDICT>  -3.80039
```

4.2. Illustration of the Explanation Facility

In the solution of problems where one has to rely on heuristic knowledge, the explanation capabilities of a system are of utmost importance. Even if an answer is right, a human user would treat it with suspicion if it is not accompanied by an explanation of how the program arrived at those conclusions. A delineation of the steps followed is usually a sufficient explanation of the program actions. In addition to this it is possible to attach a justification text for each step.

At the end of a consultation CONPHYDE reports the results, and it is possible to examine how the answers from the user support its recommendations. Notice that this kind of facility eases the debugging and fine-tuning of the expert system, and accelerates the iterative process of building an

expert system.

{Explanatory functions can justify CONPHYDE's choice.}

On a scale from -6 to 6, my certainty that
1) A physical property prediction method that uses equations of state of the Van der Waals type (two parameters).(contains the Soave and Peng & Robinson modifications of the RedHch-Kwong equation) 1s now 4.91847.

There are two favorable factors; 1n order of Importance:
1: 1) the use of the Peng & Robinson equation of state 1s advisable (certainty 4.91847) •• establishes 1 ••
1: 2) the use of the Soave eqn. of state 1s advisable (certainty 4.648716)

For which of the above do you wish to see additional Information? *1*

I suspect that
1.1) the use of the Peng & Robinson equation of state 1s advisable (certainty 4.91847).

There are several favorable factors; 1n order of Importance:
1.1: 1) the use of equations of state 1s preferred over the use of activity coefficient methods (certainty 3.080829)
1.1: 2) the set of properties that confirm or rule out the use of a . Van der Waals equation of state (certainty 4.646161)
1.1: 3) You were sure that the operating conditions are near the critical region. (RULES 1, 25, 61) (certainty 4.0)

There 1s one negative factor with neutral effect that, if positive, could have been significant:
1.1: 4) You denied that working conditions are 1n large temperature range. (RULES 40, 63) (certainty -4.0)
•• would have helped 1f positive ••

There are two factors that have not yet been considered; 1n order of Importance:
1.1: 6) the accuracy 1n the prediction of liquid densities (or properties 1n the liquid phase) 1s Important. (RULE 36) (certainty 0.0)
1.1: 6) the accuracy in the calculation of vapor pressures should be stressed. (RULE 36) (certainty 0.0)

For which of the above do you wish to see additional Information? ±

I suspect that
1.1.1) the use of equations of state 1s preferred over the use of activity coefficient methods (certainty 3.080829).

There are two favorable factors; in order of Importance:
1.1.1: 1) You were sure that the operating conditions are near the critical region. (RULES 1. 25, 61) (certainty 4.0)
1.1.1: 2) You suspected that supercritical components are present In the mixture (e.g. N2, $CH4_t$ C2H6, C02 at ordinary temperatures) (RULE 3) (certainty 2.0)

There 1s one unfavorable factor:
1.1.1: 3) the system 1s entirely composed of a mixture of nonpolar or slightly polar compounds, normal-fluids (rare gases, N2, 02, CO or hydrocarbons), C02, H2S or H2. (certainty -1.0)

For which of the above do you wish to see additional Information? *NONE*

{and so on...}

An unique quality of an expert is his ability to justify his actions for partial solutions. The explanation facility in a KBES should be able to answer the user's query as to *why* it is asking a certain question. The ability of CONPHYDE to perform this task is described by means of an extract from the consultation session.

3 — To what degree do you believe that the use of equations of state 1s preferred over the use of activity coefficient methods ? *WHY*
{Use of explanatory functions during questioning.}
If you have any reason to believe that either method 1s better than the other, 1t would not be necessary to Investigate whether the use of equations of state 1s more appropriate than the activity coefficient methods given the nature of the system.

4.3. Example of Interaction with the Knowledge-Acquisition Facility

In this section, the process of knowledge input to the KAS program is illustrated through the development of a part of the VANDERWAALS model. The inference tree for this model is depicted in Figure 4-1, which is formed by AND, OR and RULE links. For a detailed discussion of this model building process, see **[1].**

» em *VANDERWAALS* *{edit model VANDERWAALS }*
Creating the top space SPACE-VANDERWAALS.VANDERWAALS for **model** VANDERWAALS
« * » (VANDERWAALS)

•• *connect from SOAVE PENG-ROBINSON to VANDERWAALS*

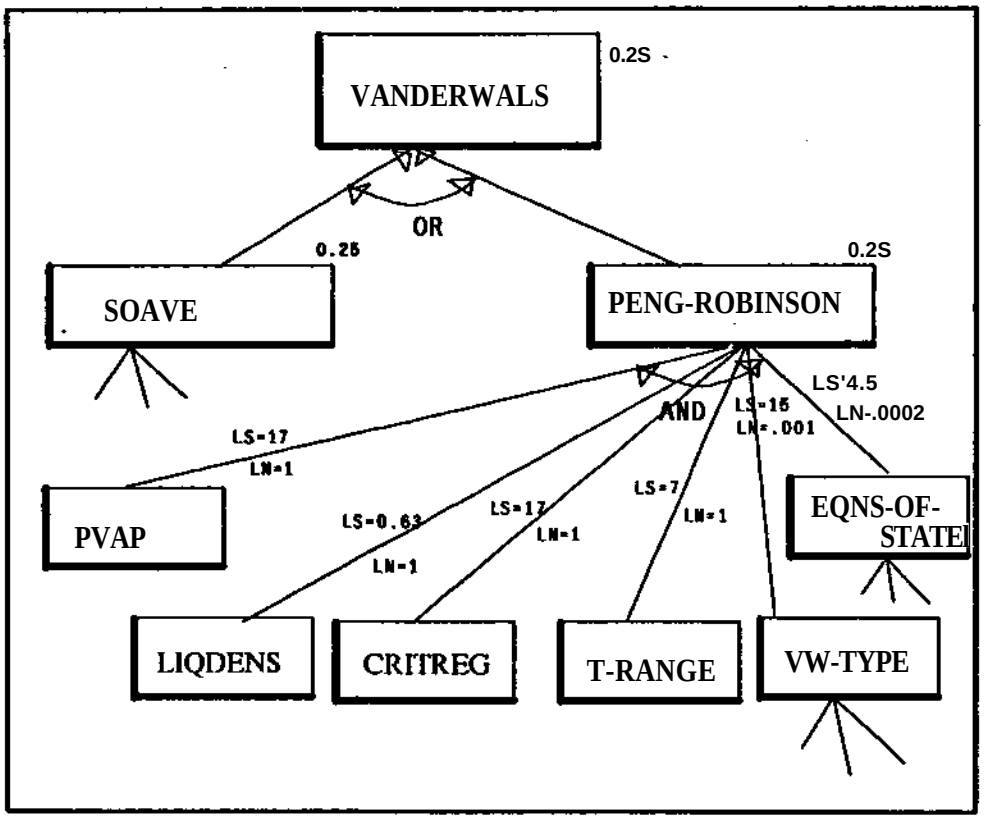


Figure 4-1: Partial Inference Network for the VANDERWALS Model

•• *connect from PVAP LIQDENS CRITREG T-RANGE VW-TYPEEQNS-OF-STATE to PENG-ROBINSON*

•* *ctvoe OR VANDERWAALS*

{type of connection to VANDERWAALS is OR}

*• *ctvpe RULES PENG-ROBINSON*

{type of connection to PENG-ROBINSON is RULES}

RULE-PVAP:PENG-ROBINSON.VANDERWAALS
RULE-LIQDENS:PENG-ROBINSON.VANDERWAALS
RULE-CRITREG:PENG-ROBINSON.VANDERWAALS
RULE-T-RANGE:PENG-ROBINSON.VANDERWAALS
RULE-VW-TYPE:PENG-ROBINSON.VANDERWAALS
RULE-EQNS-OF-STATE:PENG-ROBINSON.VANDERWAALS

•• ***Ptree VANDERWAALS*** *{print tree VANDERWAALS}*

```
VANDERWAALS OR SOAVE
                PENG-ROBINSON RULES PVAP
                                    LIQDENS
                                    CRITREG
                                    T-RANGE
                                    VW-TYPE
                                    EQNS-OF-STATE

(VANDERWAALS)
```

•• ***complete all*** *{prompt for all missing values}*

```
SPACE-VANDERWAALS.VANDERWAALS
          Description:
```
A physical property prediction method that uses equations of state of the Van der Waals tvoe (two parameters).(contains the Soave andPenQ & Robinson modifications of the Redlich-Kwong equation)

```
SPACE-SOAVE.VANDERWAALS
          Prior -  0.25

          Description:
```
the use of the Soave eon. of state is advisable

```
SPACE-PENG-ROBINSON.VANDERWAALS
          Askable ?  no
          Prior -  0.25

          Description:
```
the use of the Pena& Robinson equation of state is advisable

```
SPACE-PVAP.VANDERWAALS
          Description:
```
the accuracy in the calculation of vaoor pressures should be stressed.(RULE 35)

{and so on...}

```
RULE-PVAP:PENG-ROBINSON.VANDERWAALS
                LS »  17.0
```
{likelihood of hypothesis given presence of the evidence}

```
                LN -  UO
```
{likelihood of hypothesis given absence of the evidence}

{and *so* ***on...}***

```
** quit
Do you want to dump the models? yes
SPACE-CRITREG.VANDERWAALS
SPACE-EQNS-OF-STATE.VANDERWAALS
SPACE-LIQDENS.VANDERWAALS
SPACE-PENG-ROBINSON.VANDERWAALS
SPACE-PVAP.VANDERWAALS
SPACE-SOAVE.VANDERWAALS
SPACE-T-RANGE.VANDERWAALS
SPACE-VANDERWAALS.VANDERWAALS
SPACE-VW-TYPE.VANDERWAALS
VANDERWAALS.KAS
```

5. EXISTING APPLICATIONS IN CHEMICAL ENGINEERING

The KBES technology has been applied to relatively few systems in chemical engineering. However, there are applications in closely related areas of chemical engineering. In this section, existing expert systems (or closely related systems) are reviewed according to their location in the *derivation-formation* spectrum.

5.1. Interpretation

HEURISTIC DENDRAL[4,17] is a KBES, written in INTERLISP, that aids the chemist in the determination of the molecular structure of organic compounds. It proved the feasibility of developing KBES and laid the foundation for the development of a large number of rule-based expert systems.

The following steps, which constitute the *plan-generate-test* sequence, are involved in the determination of the molecular structure.

1. Plan:
   - The DENDRAL program is supplied with the chemical formula and the mass spectrum of the unknown compound.
   - The program determines a set of constraints from the mass spectrum.
2. Generate:
   - This information is provided to the CONGEN (CONstrained GENerator) module which generates all *plausible* 3-D structures subject to the constraints.
   - Mass spectrum is determined for each generated structure by a simulator. The simulator applies a set of cleavage rules to predict the bond breakage points.
3. Test:
   - The mass spectrum of each compound is tested against the actual mass spectrum.

- The simulated spectrum that closely matches the actual spectrum will provide the most likely structure.

The HEURISTIC DENDRAL program incorporated a number of rules obtained from experts in the structural-elucidation of organic compounds. A sample rule used in the interpretation of the mass spectra is given below:

IF	The spectrum for the molecule has two peaks ar masses x_1 and x_2 such that a. $x_1 + x_2$ = Molecular Weight + 28, and b. $x_1 - 28$ is a high peak, and c. $x_2 - 28$ is a high peak, and d. at least one of x_1 or x_2 is high
THEN	The molecule contains a KETONE group

The above rule is used to constrain the structure-generation algorithm to produce a KETONE group for the given conditions. Similar rules were developed for various phases of the program.

5.2. Diagnosis/Classification

FALCON [5],currently under development, is a KBES for diagnosing faults in a process plant. It combines both the causal model and the (surface) production rule approach. FALCON takes a set of observed effects of the plant as input and lists a number of possible faults, with a confidence factor associated with each fault. The process plant is represented by means of *components*, *variables* associated with the components, and probable *failures*. The causal model is used to propagate constraints or disturbances across various components of the system. For example, if the input flow of a pipe in the plant is reduced due to a leak in the pipe then this reduction is propagated to the adjacent component, for example, a pipe or tank. Currently the system assumes that the process, being monitored, is in a steady state and the time delay information is not included in the model. It is being developed on VAX 11/780 in Franz LISP.

CONPHYDE, described in the previous section, contains about 37 heuristic rules, considering 6 different equations of state and 9 activity coefficient equations. The CONPHYDE system is composed of 5 KAS models, with a total of about 70 spaces or nodes. CONPHYDE prompts the user for data on the composition, concentration, and physical conditions of the chemical mixture, and for data such as desired accuracy and limitations on run time (such data may be uncertain). The program's output consists of a list of physical property prediction methods in order of applicability to the specified chemical system, and an interactive explanation of the given advice. An example of a CONPHYDE rule is:

```
(25) IF the operating conditions are near the critical region,
     THEN the Peng&Robinson equation of state is advised with
          a rule strength of LS = 17, and LN = 1.0
          (Peng&Robinson has easier convergence at these conditions).
```

which implies that the Peng&Robinson equation of state is appropriate if the operating conditions are near the critical region.

5.3. Planning and Design

HEATEX [12] is a KBES for aiding in the construction of networks that minimize energy requirements by allowing the exchange of heat among various process streams. An expert has rules of thumb for arranging the network and the heat exchangers, and the program embodies some of this expertise. OPS5 [10][2] is used in this case, for its general-purpose programmability. The mode is mixed: in an initial phase of execution, the program alone produces an initial feasible configuration and evaluation of a given network. Then there is an interactive phase, where the human user can suggest evolutionary changes that might improve the overall objective of minimum energy use. HEATEX updates the network according to the user's suggestions, but it has no rules for selecting a good rule to try from among the set of evolutionary rules. The resulting ensemble performs a hill-climbing search, where enthalpy constraints are used to restrict the operations. The following paraphrases a rule from the first, constructive phase of HEATEX:

```
RULE 088
  IF   The goal is to do step 0 in creating a network,
       AND There are two streams of opposite types,
       with compatible heat-loads
  THEN Create a possible match between the two
       streams, computing the mean, the coefficient
       of the match and the delta-min value of the
       match
```

HEATEX turned out to be a useful tool for developing the theory and methods for problems of its type, resulting in the formulation of a better algorithm and especially of evolutionary rules for improvement of networks. It grew to a system of 115 production rules. The system did not, however, execute rapidly enough to be practical for further work on the problem (much faster OPS dialects have come along in the meantime). One obvious continuation would incorporate more rules into the program to suggest where the existing evolutionary rules might be applicable, further automating the

[2] Actually, an experimental, extinct dialect called OPS3RX.

improvement of networks.

DECADE (Design Expert for CAtalyst DEvelopment) [2] is a KBES, currently under development, to aid a user in the selection of a catalyst for a specified single-step reaction. The implementation of DECADE is influenced by the Blackboard architecture described in [9]; it consists of a number of Expert Knowledge Sources (EKSs) that communicate through a Blackboard. A few EKSs are enumerated below (for a more detailed description see [2]):

1. ~~Focus of Attention~~ resolves conflicts among various EKSs.
2. ~~Problem Specification~~ specifies the problem by interacting with the user.
3. ~~Thermodvnamic Consistency~~ checks the feasibility of target and other possible reactions.
4. ~~Classification of Reaction~~ classifies the target reaction.

DECADE is being implemented in SRL [37], a frame-based knowledge-representation language, and OPS5 [10], a rule-based language from the OPS group of languages.

A number of other systems have been developed using concepts from Artificial Intelligence. A brief description of some of these systems is given below; the first four applications deal with organic reaction path synthesis, while the last two deal with flow process sheet synthesis.

1. LHASA (Logic and Heuristics Applied to Synthetic Analysis) [7] was developed at Harvard, to propose a variety of synthesis routes for a given *target* molecule.
2. SYNCHEM[11], developed at SUNY Stonybrook, is similar to LHASA and uses some tree-pruning heuristics.
3. SECS (Simulation and Evaluation of Chemical Synthesis) [36], has a similar structure to LHASA, but additional information about steric and stereoelectronic effects is included in the knowledge/database.
4. The REACT program [20], developed by Govind and Powers, generates synthetic routes to industrial chemicals.
5. AIDES [27] is a heuristic program for the synthesis of process design. AIDES incorporates adaptive adjustments based on external evaluations.

6. POTENTIAL APPLICATIONS IN CHEMICAL ENGINEERING

The potential applications are discussed according to their location in the *derivation-formation* spectrum.

6.1. **Interpretation**

The development of intelligent user interfaces is essential to the success of any CAD program. KBES approach can be effectively used to provide a graceful user interface. These intelligent interfaces should [35]:

- draw conclusions based on complex calculations;
- have built in understanding of the design process; and
- provide explanations for generated results, such as why some calculations failed.

Other potential applications include interpretation of test data and providing intelligent interfaces to databases.

6.2. Diagnosis **and Monitoring**

An important aspect of Process Control is the monitoring of the various units, thereby detecting faults and diagnosing them. Here the expert system must have knowledge about the various types of faults, the probabilities of their occurrences, their symptoms, causes and cures. The system should be sensitive to faults without causing too many false alarms. A comprehensive system is bound to be complex and the traditional *fault-tree* analysis approach (see [21]) can be extended to develop KBES. Diagnosis is one of the areas where KBES have demonstrated their success convincingly through programs such as MYCIN, DART, etc. Since the task nature is identical it is perceived that KBES can play an important role in the diagnosis of chemical engineering problems. For example, the methodology reported by Himmelblau [14] can be utilized to develop a number of KBES in this area.

6.3. Planning

The planning of various phases of **a** chemical plant leading to its complete design, **a** study known as the *Project Engineering of Process Plants,* is another area where KBES would have considerable impact. The project engineering of a process plant typically involves a concerted application of knowledge from various fields of engineering and business administration. It is a highly knowledge-intensive process using a lot of *heuristics* from different fields [22]. The project engineer, **an** individual with the required diverse background to coordinate this effort, performs this duty by guiding the engineering through detailed planning of the various phases of work. KBES developed **for** this area should be able to assist the project engineer in an interactive way to plan the various aspects like the plant location, preliminary data for construction projects, process evaluation and cost estimates, basic design and flow diagrams, project scheduling, business and legal procedures etc.

6.4. Design

Process design is a complex process where the design evolves from a preliminary stage to the final stage via a trial-and-error fashion, repeatedly revising and refining the initial assumptions and restrictions. Typically, in the preliminary stage one makes a quick evaluation of the process by doing a mass balance. Then one performs the more complex task of preparing a detailed process flow sheet, complete with material and thermal balances. It is in the later stage, which is highly knowledge-intensive, KBES will be very useful. KBES applied here must have the necessary domain knowledge of the process such as the detailed models for the performance of the important unit operations. This knowledge could be stored as modules in its knowledge base and could be updated and revised when necessary.

Another aspect of process design where KBES could be used is the design of pipework and layout which can cost as much as 30% of the total design cost. This task is typically solved using heuristics and hence is idealy suited for KBES approach.

A third important aspect of process design involves the selection and design of plant equipment. In the initial stages this is more or less done using *heuristics*- that is using rules of thumb to arrive at a rapid and reliable specification of equipment type, size and cost [33]. These shortcut methods are an invaluable component of the process synthesis; they save time and they lead to a reasonably quantitative estimate of equipment size and cost, even in the absence of sufficient process data. One such hueristic approach was adapted in the selection of a distillation sequence selection, where some heuristics were used for the preliminary design stage [16]. Similarly in estimating the cost of an equipment one often relies on heuristics which deal with parameters such as equipment classification, size criterion, operator conditions extremities, materials of construction, etc [33]. These are a few representative examples of the abundant usage of heuristics in process design and hence there is a vast potential for the development of KBES. However, it must be noted that the development of these systems is not trivial and an important stumbling block is the issue of knowledge representation; Motard[19] points out that successful applications, such as heat exchanger networks, used an uniform representation of knowledge and complex domains may have to be subdivided so that each subsystem can be encoded with an uniform knowledge representation or new techniques may have to be devised to deal with the non-uniformity of knowledge representation.

7. DEVELOPMENT OF A KBES

A number of issues regarding the development of KBES are addressed in this section. Before a major developmental effort is undertaken, the appropriateness of a KBES approach to the problem should be established. Davis [8] suggests the following critera:

- a purely algorithmic solution is not appropriate;
- the task domain must have established experts;
- the experts should be much better than amateurs;
- the task should not be too easy or too difficult for the expert;
- the problem should not require commonsense (currently there are no formal methods to incorporate commonsense in a computer program); and
- the use of the KBES should result in considerable savings.

After the descision is made to pursue the knowledge-based approach, the major steps involved in the development of a KBES are[3] (See Figure 7-1) given below.

1. Identification. Important aspects of the problem are characterized. This involves identifying the relevant experts, the resources needed for the task, and the knowledge engineers (KEs); currently, there are no systems that perform knowledge-acquisition and hence the services of the knowledge engineer is required to translate the expert's knowledge into some formalism. In the initial stages, a small part of the domain is considered to study the feasibility of the system.

2. Formalization. The concepts identified in the previous stage are mapped into more formal representations based on the available tools and representational schemes. Essentially this involves the selection of a knowledge representation scheme and the appropriate tools for building the KBES (for a discussion of the tools and techniques used for this purpose see [18]). This selection requires that the KE be familiar with the domain; he may have to read a text book or two for this purpose. Once he famalirizes himself with the domain, he performs a few preliminary interviews with the expert.

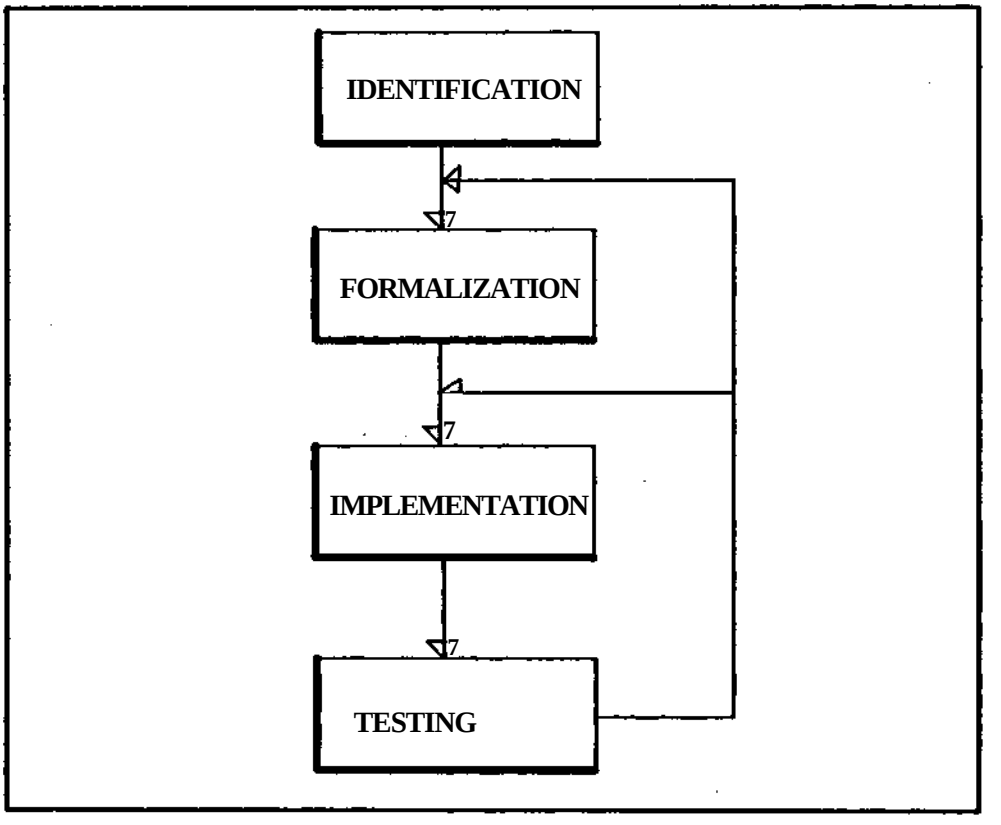
3. Implementation. This involves encoding the knowledge obtained in the previous step into the chosen tool. Sometimes, the problem (as in the case of design) may require the development of a tool for the specific purpose. In the initial stages, the knowledge-base of the system may consist of the KE's understading of the problem. A prototype system is developed in this stage.

4. Testing and Refinement. The prototype system is taken to the expert and tested. A number of examples are run and the weaknesses in the knowledge-base and the inference mechansim are identified. Often the expert would disagree with the system's

[3] Only a brief discussion is presented here. For a more complete treatment see [13].

solution and would point out the knowledge structures that lead to the inconsistencies. This type of knowledge-acquistion can be viewed as *acquistion through experimentation.*

Currently, the major bottle-neck in the development of these systems is the knowledge-acquisition process. There is an acute shortage of KEs and the knowledge engineering process is quite expensive. Further, a full scale implementation requires a close interaction between the domain expert (DE) and KE. The lack of availability of a DE, when required, often leads to expensive delays in the project. It is the authors opinion that the development of a KBES should not be undertaken unless the services of a DE are readily available to the KE.



Figu re 7 - 1 : Development Process of a KBES

8. SUMMARY

In this paper, a number of applications of KBES were described and the potential of KBES to problems in chemical engineering was addressed. KBES provide solutions to the *ill-structured* problems encountered by the engineer. This would allow the engineer to solve problems that were previously regarded as intractable. KBES also facilitate the retaining of in-house expertise. Although, these systems offer solutions to a wide range of engineering problems it is important to study the

feasibility of these systems before embarking into a major developmental effort. These systems are ideally suited for the following domains [8,15]:

- task is in a narrow domain, primarily cognitive, but requires no common sense;
- tasks take minutes to hours for an expert;
- there are recognized experts who are better than amateurs;
- skills are routinely taught to novices;
- knowledge-acquisition process should be feasible;
- experts must be readily available;
- the system structure should facilitate explanation capability; and
- there should be a high payoff for the product.

So far these systems have been successful in diagnostic-type of problems and considerable research is being undertaken to solve design-type problems. Further, a close coupling of data-bases and algorithmic programs to knowledge-bases is essential for many engineering problems.

There are a number of implications for KBES in education. These systems can be used as a training aid, thus reducing the cost of many training programs. KBES can also be used as pedagogical tools, permitting the student to organize and formalize his thought processes.

9. ACKNOWLEDGMENTS

Portions of this paper appeared in [3], [25] and [28]. We would like to thank Professor Fenves and Professor Maher for their suggestions.

10. BIBLIOGRAPHY

1. Banares-Alcantara, R., *Development of a Consultant for Physical Property Predictions,* Master's Thesis, Chemical Engineering Department, Carnegie-Mellon University, May 1982.

2. Banares-Alcantara, R, *DECADE: Design Expert for CAtalyst DEvelopment,* Ph. D. Thesis Proposal, Department of Chemical Engineering, Carnegie-Mellon University, 1984.

3. Banares-Alcantara R., Westerberg A. and Rychener M., *Development of an Expert System for Physical Property Predictions,* To be published in Computers & Chemical Engineering, 1984.

4. Buchanan, B. G. and Feigenbaum, E.A., "DENDRAL and Meta-DENDRAL: Their Applications Dimensions," *Artificial Intelligence,* Vol. 11,1978,.

5. Chester, D., Lamb, D. and Dhurjati, P., "An Expert System Approach to On-line Alarm Analysis in Power and Process Plants,[11] *Computers in Engineering - Advanced Automation: 1984 and Beyound,* Gruver, W. A., ed., A. S. M. E., A. S. M. E., 1984, pp. 345-351.

6. Clancey, W., "Classification Problem Solving,[11] *Proceedings 4th NCAI,* William Kaufmann Inc., AAAI, 1984, pp. 49-55.

7. Corey, E. J., Wipke, W. T. and Cramer, R. D., "Techniques for Perception by a Computer of Synthetically Significant Structural Features in Complex Molecules ," J. *Am. Chem. Soc,* Vol. 98,1976, pp. 210-?.

8. Davis, R., *Tutorial on Expert Systems,* Presented at AAAI-83.

9. Erman, L. D., Hayes-Roth, F., Lesser, V. R., and Reddy, D. R., "The Hearsay-ll Speach Understanding System: Integrating Knowledge to Resolve Uncertainty," *Computing Surveys,* Vol. 12, No. 2,1980, pp. 213-253.

10. Forgy, C. L., *OPS5 User's Manual,* Technical Report CMU-CS-81-135, Carnegie-Mellon University, July 1981.

11. Gelertner H., Sridharan N. S., Hart A. J., Fowler F. W. and Shire H. J., *An Application of Artificial Intelligence to the Problem of Organic Synthesis Discovery,* Technical Report, Department of Computer Science and Chemistry, SUNY at Stony Brook, August 1971.

12. Grimes L, Rychener M.and Westerberg A., "The Synthesis and Evolution of Networks of Heat Exchangers that Feature the Minimum Number of Units," *Chemical Engineering Communications,* Vol. 14,1982, pp. 339-360.

13. Hayes-Roth, F, Waterman, D., and Lenat, D., editors, *Building Expert Systems,* Addison-Wesley Publishing Company, Inc., 1983.

14. Himmelblau, D. M.,f *Fault Detection and Diagnosis in Chemical and Petrochemical Processes,* Elsevier Scientific Publishing Company, 1978.

15. Kant, E., *Notes on Task-Oriented Intelligent Systems,* Class notes for the graduate course in Artifical Intelligence, Carnegie-Mellon University, 1983.

16. King, J. C, *Separation Processes,* McGraw-Hill Book Company, New York, New York USA, 1980.

17. Lindsay, R., Buchanan, B., Feigenbaum, E. and Lederberg, J., *Applications of Artificial Intelligence for Chemical Inference: The Dendral Project,* McGraw-Hill Book Company, 1980.

18. Maher, M.L., Sriram, D. and Fenves, S.J., *Tools and Techniques for Knowledge-based Expert Systems for Engineering Design,* To Appear in Advances in Engineering Software, U. K., October, 1984

19. Motard, R. L, "Computer Technology in Process Systems Engineering," *Computers & Chemical Engineering,* Vol. 7, No. 4,1983, pp. 483-491.

20. Govind, R. and Powers, G. J., "Studies in Reaction Path Synthesis," *AIChE Journal,* Vol. 27, No. 3, 1981, pp. 429-442.

21. Powers G., Tompkins C. F. and Lapp S. A., "A Safety Simulation Language for Chemical Processes: A Procedure for Fault Tree Synthesys," *Reliability and Fault Tree Analysis. Theoretical and Applied Aspects of System Reliability and Safety Assesment,* Barlow E. Richard, Fussel B. Jerry and Singparwalla D. Nozer, eds., Society for Industrial and Applied Mathematics, 1975, pp. 57-75.

22. Rase, H. F. and Barrow M. H., *Project Engineering of Process Plants,* John Wiley & Sons, Inc., New York, New York USA, 1957.

23. Reboh, Rene, *Knowledge Engineering Techniques and Tools in the PROSPECTOR*

Environment, Technical Note 243 SRI Project 5821, 6415 and 8172, Artificial Intelligence Center. SRI International, June 1981.

24. Rich, E., *Artificial Intelligence,* McGraw Hill, 1983.

25. Rychener, M. D., "Expert Systems for Engineering Design: Experiments with Basic Techniques," *Automating Intelligent Behavior: Applications and Frontiers,* IEEE, May 1983, IEEE Trends and Application Conference, Wastington D. C. (See also DRC technical report DRC-05-02-83, Carnegie-Mellon University)

26. Sargent, R. W. H., "Process Systems Engineering: Challenges and Constraints in Computer Science," *Proceedings of the Secon International Conference on Foundations of Computer-Aided Design,* Westerberg, A. *W.* and Chien, H. H., eds., 1983, pp. 23-46.

27. Siirola, J. J., Powers, G. P., Rudd, D. F., "Computer-Aided Synthesis of Chemical Process Designs," *AIChE Journal,* Vol. 17,1971, pp. 677-?.

28. Sriram, D., Maher, M. L. and Fenves, S. J., *Knowledge-based Expert Systems in Structural Design,* Presented at NASA conference on Advances in Structural Mechanics, October 1984, Washington, D. C. (to appear in Computer and Structures, January 1985.)

29. Sriram, D., "A Bibliography On Knowledge-Based Expert Systems for Engineering," *SIGART newsletter,* Vol. 89, July 1984, pp. 32-40.

30. Sriram, D. and Mullarkey, P., *Inexact Inference Methods for Expert Systems: A Tutorial,* Forthcoming technical report, Department of Civil Engineering, Carnegie-Mellon University, 1984

31. Stefik, M. and Martin, N., *A Review of Knowledge Based Problem Solving as a Basis for a Genetics Experiment Designing System,* Technical Report STAN-CS-77-596, Computer Science Department, Stanford University, March 1977.

32. Umeda, T., "Computer Aided Process Synthesis," *Computers & Chemical Engineering,* Vol. 7, No. 4, 1983, pp. 279-309.

33. Valle-Riestra, J. F., *Project Evaluation in the Chemical Process Industries,* McGraw-Hill Book Company, New York, New York USA, 1983.

34. Weiss, S. and Kulikowaski, C, *A Practical Guide to Designing Expert Systems,* Rowman & Allanheld, 1984.

35. Westerberg A. W. and Rychener M. D., *An Intelligent Interface for a Complex Design Program,* Unpublished,Carnegie-Mellon University

36. Wipke W. T., Brann, H., Smith, G.,Choplin, F. and Seiber, W., "SECS-Simulation and Evaluation of Chemical Synthesis : Strategy and Planning,[11] *Computer-Assisted Organic Synthesis,* ACS Symposium Series, 1976, pp. 61.

37. Wright, J. M. and Fox, M. S., *SRL/1.5 User Manual,* Technical Report, CMU Robotics Institute, June 1983.