Effective Software Support for Chemical Research

Amanda Jayne Welsh

Ph.D.

University of Edinburgh

1993

for "me mam"

Abstract

Locating metal-bound hydride ligands in transition metal cluster compounds can be difficult with conventional analytical chemistry. Thus many empirical and theoretical methods have been developed for locating such ligands. One such method has been implemented in the Complete Coordinate Convergence Program (*CCCP*). However, on close examination, this tool is found to be user-unfriendly and its success is shown to be dependent on an initial location estimate from the user.

Thus the Locator was developed as a tool designed to overcome these problems. Methods are presented which exploit known chemical and spatial restraints in finding an initial estimate of the hydride ligand position(s) for more reliable subsequent optimisation by the CCCP code. The performance of Locator was tested on 178 models, with a favourable success rate of 72%. This figure rises to 84% when taking account of bonding interactions of the hydride ligand(s) provided from the user.

These improvements to CCCP were motivated by a user requirements analysis undertaken to identify computational chemistry applications which would benefit from the exploitation of software engineering and HCI principles becoming prevalent in mainstream computer science. A molecular graphics tool was also selected because it was considered that this would be useful to many researchers, and because the available molecular graphics tools were observed to be lacking in terms of human-computer interaction principles. Further, such a tool could be used as a base for other applications.

This thesis by no means suggests that the scope of physical scientific applications which would benefit from the uptake of such ideas is limited to this relatively narrow field of chemistry. Rather these tools are considered in a *proof of concept* case study within the time constraints of this work.

The adoption of the Visualisor interface and the addition of the initial estimate routine are considered to have significantly improved hydride ligand location in transition metal clusters both in terms of success and accuracy of the results and the ease with which these results may be obtained. Further, it is considered that this project has demonstrated the benefits of the application of current computer science software development principles to scientific software.

Table of Contents

1. Introduction 1

1.1 Potential Benefits of Computer Software to Chemistry 2

1.2 Barriers to a Wider Use of Computers in Chemistry 4

1.2.1 The Barriers Encountered 4

1.2.2 Causes of the Barriers 6

1.3 Proposed Solution 8

1.3.1 Objectives 8

1.3.2 Method 9

1.3.3 Overview 10

2. User Requirements Analysis 13

2.1 Introduction 13

2.2 Overview 15

2.2.1 User Resistance 15

2.3 Questionnaire 19

2.3.1 Aims 20

2.4 Results 22

2.4.1 Statistical Analysis of Responses 22

2.4.2 Section 1 — Working Practices 23

2.4.3 Section 2 — Library System 27

2.4.4 Section 3 — Current Facilities 28

2.5 Discussion 36

2.5.1 Limitations of the Questionnaire 36

2.5.2 Tool Support for User Activities 40

2.5.3 Support Relative to Origin of Software 47

2.6 Conclusions 50

2.6.1 Current State 50

2.6.2 Choice of tools to be implemented 51

2.6.3 Requirements 52

2.6.4 User Model 53

3. Design of the Tool Kit 58

3.1 Introduction 58

3.2 Overview 59

3.2.1 General Issues for Effective Design 59

3.2.2 Software Engineering Issues for Effective Internal Design 61

3.2.3 Human-Computer Interaction Issues for Effective External Design 65

3.3 Objectives 71

3.4 The Tool-Kit 74

3.4.1 The Orthogonalisor 75

3.4.2 The Visualisor 77

3.4.3 The Locator 77

3.5 Implementation 78

3.5.1 General Issues 78

3.5.2 Software Engineering Issues 78

3.5.3 HCI Issues 83

4. The Visualisor **89**

4.1 Introduction 89

4.2 Literature Review 90

4.2.1 Scientific Visualisation 90

4.2.2 Molecular Graphics 92

4.3 Objectives 96

4.4 Implementation 97

4.4.1 Files Menu 99

4.4.2 View Menu 101

4.4.3 Print Menu 102

4.4.4 Display Menu 104

4.4.5 Chemical Knowledge 104

4.5 Evaluation 105

4.5.1 Achievement of Objectives 106

4.6 Discussion 111

4.7 Conclusions 112

4.8 Further Work 113

4.8.1 Atom Number Restrictions 113

4.8.2 Loading and Saving Files 113

4.8.3 Display 114

4.8.4 PostScript Output 116

4.8.5 Keyboard Shortcuts 116

5. Hydride Ligand Location in Transition Metal Complexes **118**

5.1 Introduction 118

5.1.1 Characterisation 118

5.1.2 X-ray Diffraction 119

5.1.3 Neutron Diffraction 120

5.2 Current Methods of Hydride Ligand Location 121

5.2.1 Steric Methods 121

5.2.2 Potential Energy Calculations 122

5.2.3 Molecular Orbital Calculations 125

5.3 Analysis of CCCP 131

5.3.1 Stage 1: Failure of CCCP 135

5.3.2 Stage 2: Investigation of the area around the known "correct" position 137

5.3.3 Stage 3: Estimates using Chemical Knowledge 140

5.3.4 User Interface 143

5.4 Conclusions 148

6. The Locator **150**

6.1 Introduction 150

6.2 Objectives 152

6.3 Principles of the Initial Estimate Method 153

6.3.1 Rationalisation of the model 153

6.3.2 Initial Estimate of Hydride Ligand Position 156

6.4 User Interface 166

6.4.1 Output and Experimental Control 166

6.4.2 Hydride Ligand Location Procedures 170

6.5 Performance Evaluation of the Locator 173

6.5.1 Models Used 174

6.5.2 Evaluation of the Preliminary Stages of the Locator. 175

6.5.3 Comparative Evaluation the Locator 179

6.5.4 Absolute Evaluation of the Locator 181

6.5.5 Dihydrogen Models 189

6.5.6 HCI Evaluation 191

6.6 Conclusions 195

6.7 Further Work 196

6.7.1 Efficiency 196

6.7.2 Charge Iteration 197

6.7.3 Tuning of the orbital energy and exponent of the metal-bound hydride ligand(s) 199

6.7.4 Dihydrogen Models 204

6.7.5 The Failure of the Locator for Some Models 204

7. Conclusions and Further Work **205**

7.1 Conclusions 205

7.2 Future Work 209

7.2.1 Improvements to Current Tools 209

7.2.2 A Construction Tool 210

7.2.3 Future Developments to the Tool-Kit 211

7.2.4 Broadening the Scope of Applications 212

A. Questionnaire **213**

B. Questionnaire Results **223**

B.1 Section 1 — Working practices 223

B.2 Section 2 — Library System 228

B.3 Section 3 — Use of EMAS 229

B.4 Section 4 — User Suggestions 231

B.4.1 Library System 231

B.4.2 Applications 231

B.4.3 User Interface Issues 233

B.4.4 Support 233

C. List of Molecular Models 234

D. Molecular Models 244

E. The Locator: Raw Results 255

F. Comparison of Initial Estimate Methods 262

G. Publications 265

Chapter 1

Introduction

Research chemists are finding increasing benefits from the use of computers; the rapid decrease in the cost of computer hardware, allied with ongoing improvements in quality, has meant that computers and associated software are becoming more and more evident in chemistry. High quality workstations with bit-mapped screens and cheaper, better graphics are gradually taking over from teletype terminals connected to main-frames, enabling more sophisticated and user-friendly software to be run.

Corresponding to this rise in quality of *hardware* has been a potential rise in quality of *software*, already evident in many areas and some applications, but unfortunately not in all. There are many chemical software applications available which are functionally useful. Programs which calculate solvent-accessible surfaces of molecules [1,2] provide the chemist with valuable visualisation information for tasks such as reaction design. Databases [3–10] allow users to search for a variety of objects, such as molecular models, reaction mechanisms, references to articles *etc.*, by a variety of keys, including citations, authors and sometimes molecular sub-structure. Molecular modelling programs [11–13] predict the positions of atoms which cannot be located by conventional techniques. Some of these tools have well designed interactive graphical interfaces, but others, mainly in-house applications, pay little attention to the user interface [14–16]. These tools often have many man-years invested in them, but

where they are difficult to use because of badly designed or *ad hoc* interfaces the time invested may be wasted because of under-use.

This chapter discusses the potential benefits to chemistry of computer science — the advantages offered by current computer science technology and approaches. The potential "ideal" state is then contrasted with the current situation and reasons for the difference between the two are proposed. A set of objectives for bringing computer science and chemistry closer together, and closing the gap between the ideal state and the current state, are listed, and a method for achieving those objectives is given. Finally an overview of this thesis is given chapter by chapter, and the contents of each introduced.

1.1 Potential Benefits of Computer Software to Chemistry

The large amount of data used by chemists present an ideal opportunity for computerisation [8] and computers have been used for data acquisition and "number-crunching" for several years. Computer software can also aid chemists in the processing, collection and visualisation of results while some analytical instruments are controlled by computers. Articles, theses and reports may be produced using word-processors.

The benefits of software in supporting the activities of a chemist during research may be divided into three categories:

1. *Speed:*

   Software support may allow the activity to be performed to the same standard as if done manually but in a shorter period of time. For example, the time spent in literature searching or the construction of simple theoretical models can greatly be reduced by using a computer.

2. *Quality:*

   Software support may allow the activity to be carried out to a higher degree of quality than if tackled manually. With today's sophisticated word processors and document preparation systems articles and reports may have much better presentation than those prepared with typewriters, hand-drawn diagrams, scissors and glue. Furthermore, with the advent of powerful spelling, grammar and style checking facilities, transcription errors can be all but eliminated.

3. *Possibility:*

   Perhaps the biggest benefit computers may offer to chemists is the ability to tackle problems which would be impracticable without their aid. Several activities involve extensive numerical calculations, such as the processing of X-ray diffraction data or the location of metal-bound hydride ligands in transition metal clusters. If these tasks were performed with paper and pencil, the calculations involved in either would be too large to be practicable, because of both time and susceptibility to error.

In an ideal world, the chemist would encounter a task which could be more efficiently performed with the aid of a computer, locate the appropriate program, use it, then continue with his or her research. The programs should be tools, a means to an end. The people using them are interested only in the task to be accomplished — not in the tools they use to accomplish that task. An electrician is not interested in how an ammeter works — only in the readings it gives — and would be unlikely to use it often if a significant time were required to read a manual before taking each measurement. A person undertaking a certain task with a computer program would like the tool to be just as transparent — being interested only in the results it gives.

Thus, an ideal software tool should be transparent to the user and require little or no time to learn or re-learn how to use it. Though probably not attainable, these ideals may be valid aims. Unfortunately, the tools available to research

scientists often fall far short of these ideals, for several reasons identified in this chapter.

1.2 Barriers to a Wider Use of Computers in Chemistry

1.2.1 The Barriers Encountered

The poor quality and inappropriate user interface design of the software available in the real world may cause chemists to fail to exploit the potential benefits of computers in their research. Furthermore, a poorly designed operating system and/or on-line help system may prevent the user from finding out about appropriate software.

To be able to navigate around the on-line help system – if available – and to locate and use the appropriate tools obliges the user to surmount learning barriers at each stage. The learning barriers for a system with a well constructed, informative help system and an intuitive, well-documented operating system and tools may be small, making the use of a computer advantageous. However, if the programs are internally inconsistent and counter-intuitive much time must be spent learning or reminding oneself of complicated syntax, thus the learning barriers may be somewhat higher. For occasional users, as would be the case for most chemists, it may be necessary to repeat the learning processes each time a computer is used. In the worst case the time spent learning or re-learning how to use the program may exceed that saved by using the program itself; in such a case the novice user may decide that the barriers are not worth surmounting. Unfortunately, many programs available to chemists fall in to this category, suffering from several problems which may discourage chemists from their use:

1. *Command-line interfaces:*

   Programs tend to have "command line" interfaces — where complex al-

phanumeric commands must be entered from the keyboard, as opposed to direct manipulation interfaces — where operations are conducted with a pointing device. It is generally acknowledged [17–21] that the latter style is more effective and easier to use, particularly for occasional or novice users. Command line interfaces may cause problems for such users since complicated syntax must be remembered and used correctly, and manuals or on-line help systems must be consulted before appropriate operations are ascertained.

2. *Lack of input flexibility:*

   Often a program requires an unnecessarily rigidly structured input file. Some programs require data to be entered in certain character positions on a line, left or right-justified depending on the data type. Others require data to be solely in upper case.

3. *Lack of integration:*

   Programs tend not to be integrated in that files must be edited from one form to another. For example, the molecular model file output from *SHELX*, a crystal structure solution program, is incompatible with the input file for *CCCP*, a tool for subsequently locating transition metal bound hydride ligand(s). The users must edit output from SHELX for input to CCCP if unresolved hydride ligands are to be located.

   Programs also tend not to be integrated, in that the programs must be *used* independently. The user must wait until the crystal structure is solved or the hydride(s) located before being able to view it, whereas a three-dimensional visual representation of the current state of the molecular model might well be useful during the running of the program.

4. *Inconsistency:*

   Inconsistency is common, both internally — within one tool, and externally — between tools. As an example, in the hydrogen location tool CCCP,

some Boolean values *suppress* output where others *enable* it. A more detailed critique of the interface to CCCP is given in Chapter 5

5. *Lack of intuitiveness:*

   The lack of direct manipulation interfaces and consistency both contribute to programs being unintuitive or even counter-intuitive. Users may find it difficult to build a mental model of the way the program works and apply this model to other systems.

6. *Superfluity:*

   Programs often require information to be provided by the user which could be provided by the program, or at least assigned default values. Ortep, for example, requires atomic radii and bonding interaction from the user, information which could easily be provided by the program. It is acknowledged that sometimes the user should be able to give this information if they so wish, but defaults should be provided otherwise.

To understand the reason for the deficiencies in chemical software listed above, the use of computers in chemistry and the ways in which such software is developed must be examined.

1.2.2 Causes of the Barriers

There would seem to be little or no overlap between computer science and chemistry. The studies of software engineering and human-computer interaction are relatively new developments in computer science. Their use seems to be confined to computer scientists. Authors of chemical software are often of necessity chemists rather than computer scientists. When a chemist encounters a numerical problem which is too difficult to solve on paper, and there are no programs readily available to solve this problem, the chemist may consider writing one or adapting an existing, functionally similar program in an *ad hoc* manner as appropriate. The chemist may be untrained in computer science

practices or not consider such techniques to be applicable to the programs they write, perhaps being under the impression that documenting how to use the program in operation is preferable to making the actual program easier to use.

An example of this approach is the development of CCCP. This program was adapted, by a chemist, from *ICON8*, written by another chemist, which relies on an input file containing obscure Boolean control values and follows complex Fortran input conventions. Output data consists of a series of both individual values and matrices. CCCP uses ICON8 iteratively and retains the same input and output, with some free-formatted additions. However, instead of the matrices being printed out once (as for ICON8), they are printed out for each iteration of the program. This form of presentation of these data makes the actual data sought, the positions of the hydride ligands, difficult to locate and assimilate. Further, the program is dependent on starting point(s) required from the user, although no guidelines are provided for calculating such a point. When given the correct conditions, the program *works* but is difficult both to use and to interpret the results for even the simplest of models.

The difficulty of using such programs often circumvents their use. Functionally, these programs may be useful, having many man hours invested in them, but if they prove too difficult to use this investment will be wasted.

It is not an intention to criticise chemists for this practice; it is reasonable for chemists to be interested only in the application, the results of the program and the underlying chemistry — not in the program itself, its design or ease-of-use. Chemists are trained in chemical theory and chemical techniques but rarely in software design or human-computer interaction techniques.

This thesis shows that computer science techniques *are* worthwhile and appropriate to chemical applications.

1.3 Proposed Solution

1.3.1 Objectives

The initial, perhaps ambitious, aim of this research project was:

> *To design and implement a support environment for use within the scientific community, providing the users with a means of going about their work with minimum attention to the tools they use, allowing them to concentrate on the task in hand.*

It was soon realised that the design and implementation of a whole support environment for the entire scientific community was too ambitious. However, if the original aim was not *attainable* with the existing constraints, it was perhaps *demonstrable.* Thus a more realistic aim was formulated:

> *To design a base for an integrated tool-kit and illustrate the tool-kit by implementing some example tools to be constituents of the set. The tool-kit would be designed for general use by any scientist, but for the purposes of this research the science of chemistry, specifically in the Chemistry Department at the University of Edinburgh would be addressed as a case study. The tool-kit would be designed to bring together latest techniques in computer science* and *chemistry. This would promote the design of tools with maximum utility to chemists which allow them to concentrate on the work in hand, with minimum attention to the tools they use to complete their task.*

The reasons behind the choice of chemistry for a case study were as follows.

- The author and two of her supervisors have a background and consequent interest in chemistry. Access to chemists and current computing facilities in Edinburgh University Chemistry Department would therefore be

enhanced, and personal experience and knowledge of current problems could be applicable to the research.

- Chemistry is arguably as appropriate as any other physical science for a case study of this type, and is in fact an archetype. Chemistry has aspects of physics in physical and theoretical chemistry and biology in organic chemistry. Differing avenues of research can require numerical problems, visualisation of results or processing of experimental data. It could be argued that most aspects of science which could benefit from the use of a computer manifest themselves in some form in chemistry.

1.3.2 Method

The method by which it was hoped to achieve these objectives was as follows:

1. *Literature Survey*

   Perform a literature survey on the latest approaches to software design — specifically in the fields of software engineering for internal design, and human-computer interaction for external design. Select aspects of these approaches which are appropriate to the proposed project with a view to application to the design of the tool-kit.

2. *User Requirements Analysis*

   Obtain a more detailed picture of the working practices of research scientists and the software available to them, using the Chemistry Department at the University of Edinburgh as a case study. Crystallise requirements by carrying out a *user requirements analysis*, using the results of this analysis to select appropriate tools, and characteristics thereof, for implementation. This led to the identification of hydride ligand location as a suitable area for study.

3. *Application and Implementation*

   Implement the tools selected in stage 2, applying the latest computer

science techniques and using encapsulated chemical knowledge in the form of chemical heuristics and a "database" of chemistry facts. This led to unexpected problems with existing tools and to the exploration of new solutions to EHMO computation.

4. *Evaluation*

   Evaluate the tools in terms of their contributions to both chemistry and computer science.

1.3.3 Overview

The remaining chapters of this thesis address the following issues:

Chapter 2: User Requirements Analysis

The literature concerned with questionnaire design, user requirements analysis, user involvement and user modelling is reviewed and the importance of each discussed. The implementation of user requirements analysis *via* the distribution of a questionnaire is described. The results of the questionnaire are analysed to produce a user model, and user requirements are listed.

Chapter 3: Design of the Tool Kit

As the aim of this project is to implement tools as part of a greater whole, a tool-*kit* so to speak, it would be insufficient to describe the tools in isolation without also discussing those aspects which constitute a *kit*. This chapter reviews current literature about *human-computer interaction* (HCI) and *software engineering* techniques. The manner in which these techniques have been applied in the implementation of the software environment is discussed with respect to the user requirements and objectives introduced in the previous chapter.

Chapter 4: The Visualisor

It became apparent from the results of the questionnaire that a new molecular graphics tool would be desirable. Existing implementations of such a tool were non-interactive, their use was laborious, and a molecular visualisation tool would be utilised by a greater proportion of chemists than any other. Furthermore, such a tool could provide a graphical interface for other tools which deal with molecular models, such as that described in Chapters 5 and 6, thus promoting consistency across tools. Current literature on molecular graphics is reviewed and lists of requirements and objectives are formulated. The implementation of the molecular graphics tool is described and then evaluated in the light of the literature review and the objectives and requirements.

Chapter 5: Hydride Ligand Location in Transition Metal Complexes

The location of hydride ligands within transition metal complexes can be difficult in conventional analytical chemistry. The traditional technique of X-ray diffraction is generally considered inadequate in locating metal-bound hydride ligands, especially for larger metals or transition metal cluster compounds. Neutron diffraction provides a solution, but is expensive and results are of low precision. This chapter reviews current techniques for both characterising the whole molecule and, more specifically, for locating hydride ligands within these species. One method, specifically the Modified Extended Hückel Molecular Orbital (*MEHMO*) technique implemented in the Complete Coordinate Convergence Program (*CCCP*), is investigated more thoroughly, and some suspected limitations of the method are confirmed by a series of experiments. It was established that there was room for improvement of CCCP in two distinct areas: the human computer interaction properties of the program and, more importantly, the dependence of the final hydride ligand positions on a user-given starting point.

Chapter 6: The Locator

This chapter describes the implementation of a program, *the Locator*, which attempts to address the problems with CCCP highlighted in the previous chapter. The Locator contains a routine which analyses the spatial characteristics of the X-ray skeleton of a molecular model in order to make an initial guess of the position(s) of the unknown hydride ligand(s), before subsequent optimisation by the contained CCCP code. Further, the Locator uses the Visualisor as an interface to address the problems of usability highlighted in the previous chapter. The hydrogen location tool is evaluated with respect to the objectives and requirements detailed, both quantitatively — where the success at locating hydrides is discussed — and qualitatively — where the application of HCI and software engineering issues is discussed. Comparing the results produced by the Locator against those produced by the raw CCCP established that the Locator usually afforded a more accurate result, in a shorter space of time and more often, than without this initial estimate as a starting point. More extensive testing, on 178 models, showed that the Locator achieved a reasonable success rate on all but a small group of models, those which contain dihydrogen ligands. Further, the initial estimate at the hydride position was sometimes closer to that revealed by neutron diffraction than that arrived at after optimisation.

Chapter 7: Conclusions and Further Work

This final chapter concludes the thesis by discussing to what extent it is considered that the original objectives have been achieved and proposes possibilities for future work.

Chapter 2

User Requirements Analysis

2.1 Introduction

The first stage in the design of the tool-kit was to ascertain the current situation regarding the working practices of the sample user community — in this case research chemists in the Department of Chemistry, University of Edinburgh — and their use of computers, in order to identify the areas with the greatest potential for benefit from software support.

After approximately one year working in the Chemistry Department the author had some experience of the working practice and use of computers in fields of inorganic chemistry and X-ray crystallography. By talking and mixing with people both within and outwith this specific field *impressions* were formed of how other chemists in the Department went about their work.

The impressions gained from this short exposure to research chemistry were that whilst the software facilities available were generally adequate in terms of their functionality — they did what the user wanted them to do — many left a lot to be desired in their user interface. Features of these tools which render them more difficult to use, either solely, or in combination with other factors, include lack of intuitiveness, sometimes counter-intuitiveness, internal inconsistency, a requirement for rigidly formatted input files and a lack of a direct-manipulation

interface. It is not claimed that all the tools suffered from these problems, but there was certainly evidence that some did.

Information regarding the existence of some tools seemed to be lacking. The knowledge of tools seems to be passed by word of mouth, with new researchers being introduced to tools by their peers whilst rarely being informed of alternatives. One postgraduate chemist was introduced to a screen editor — *VECCE* — whilst some way through writing his thesis, having used a line editor — *Edit* — to that point. He considered the new editor to be a great improvement, and expressed frustration that he had not been made aware of the tool previously. He firmly believed that the use of the screen editor accelerated the completion of his thesis.

It did not seem appropriate to rely on personal experience or that of personal contacts as being sufficiently typical for extrapolation across the rest of the department to yield a representative *user model*. The author's impressions required corroboration to be certain that the same problems were being experienced by a significant number of other researchers. More information was required, specifically from an investigation into the working practices of members of the *whole* user community — inorganic, organic, physical and analytical researchers alike. It also seemed necessary to establish those software facilities currently available, the standard of these facilities, both in functionality and ease-of-use, whether they were being fully exploited by the user community, and if not — why not? A questionnaire was therefore composed for distribution to all members of the user community in order to answer these queries and establish the validity of the informal intuitive perceptions.

This chapter reviews the literature addressing *user requirements analysis*, the term given to the investigation of the user community as part of the design process, and describes how this was achieved with a questionnaire. The results of the questionnaire are discussed, and a user model of a generic research chemist constructed therefrom is presented.

2.2 Overview

2.2.1 User Resistance

Manifestations

Many instances of user-resistance to the installation of new software have been documented. This resistance can be manifested in many forms ranging from reluctance to be involved with the change to outright sabotage. Hirschheim [22] reviews four articles in which many categories of resistance to the introduction of new software are identified. Dickson *et. al.* [23] have identified three classifications of behaviour exhibited by users experiencing such change:

aggression — a behaviour which represents an attack (either physically or non-physically) with the intent of injuring or causing harm to the object presenting the problem. An example is described where workers reputedly placed paper-clips and poured honey in the entry slots of source recorders meant for identity badges [24]. Moreover, some source recorders were mysteriously run over by fork-lift trucks.

projection — a behaviour exhibited when the system is blamed for causing difficulties. As an example, employees may claim that they have no time to work if they spend all their time punching jobs in and out of a transactor.

avoidance — occurs when people defend themselves from the system or withhold therefrom.

Similarly Fried [25] identifies regression, aggression and hostility and the tendency to blame others, whilst Sanders [26] suggests that users react by withholding data, providing inaccurate data, distrusting computer output and exhibiting lowered morale.

Causes

Hirschheim suggests many causes for resistance:

1. *Innate conservatism:*

   Resistance may be based on the desire for continuity and a loyalty to the old methods.

2. *Lack of felt need:*

   If the present system is perceived to be satisfactory it will be difficult to persuade users to change to another.

3. *Uncertainty:*

   Users may suffer from uncertainty of the effect which introduction of new software will have their working life — for example, on job security and prestige — particularly if they lack confidence of being able to acquire the necessary skills.

4. *Lack of involvement in the change:*

   Lack of involvement in the change may cause resistance to both the decision to change and participation in the development.

5. *Redistribution of resources:*

   The redistribution of resources may disrupt the status quo of power over their distribution.

6. *Lack of organisational validity:*

   New software may not match the work-patterns of the people who are obliged to use it. [27]

7. *Lack of management support:*

   If management is not seen to support and encourage the change, organisational workers are unlikely to be willing advocates of the system.

8. *Poor technical quality:*

   Resistance is more likely to occur in systems which are cumbersome, "unfriendly", unreliable, slow or lack functionality.

9. *Personal characteristics of the designer:*

   The designer must remember that prospective users of the system will not necessary have the same fascination with technology as themselves.

Not all aspects of user resistance apply to this thesis. The form of user resistance most likely to be encountered here is *avoidance,* or *lack of use.* There is a significant difference between the current project and those discussed above in that any tools will be developed for the *convenience* of the user should they choose to take advantage of this service, rather than being imposed on users or replacing established facilities. Resistance is therefore unlikely to be caused in research chemists by any of *uncertainty* (3), *lack of involvement* (4) or *redistribution of resources* (5). Another reason for the perceived lack of involvement being of limited significance is the high turnover of researchers in academic departments. Academic staff constitute about a quarter of the University of Edinburgh Chemistry Department and tend to remain for extended periods. However, the post-graduate and post-doctoral researchers constituting the remaining three quarters of the Department tend to stay for a only limited period. Only a few users of a system in the latter category will therefore have been present at the time of implementation and hence have had any potential for expectations of involvement. It is modestly hoped and certainly intended that *poor technical quality* (8) would not cause concern. *Lack of management support* (7) is not particularly relevant since the Chemistry Department does not rely on a managerial system for research.

Those causes of user-resistance which are considered to be the most important here are *innate conservatism, lack of felt need* and *poor organisational validity.*

Let us consider each of the remaining points in turn with respect to how, if possible, they may be counteracted:

1. *Innate conservatism* is unlikely to be overcome. However, by concentrating on counteracting the other points it is hoped that innate conservatism will be insufficient to sustain resistance.

2. *Lack of felt need* — Users may be persuaded to change if the new tool can be shown to be more powerful than existing methods in that the user can perform the task to a greater extent and/or a better quality and/or with less cognitive effort. This may be ensured by *better technical quality.*

6. *Lack of organisational validity* — if the designer has a reasonable knowledge of the working practices and environment of the prospective users, resistance arising from lack of organisational validity may be minimised.

User involvement

It is widely accepted [22, 23, 27–31] that the involvement of users in the design process (proactive design) is important and contributes to the acceptance of a system by prospective users, since this reduces:

6. *Lack of organisational validity:* the software is more likely to fit current working practices.

Generally software developers design and implement programs outwith their own field of expertise and work environment. The speciality and work environment of designers are typically "software design" and "software house" respectively. Thus there needs to be a meeting of the expertise of the designer with a knowledge of the application and the environment in which it is to be used. According to Franz [29] "Users feel they already understand current problems and know their needs well. Designers require a more thorough comprehension because they are not regular users."

A human-computer interaction *(HCI)* literature survey performed by Alty *et. al.* [17] supports the importance of user involvement. The adage 'Use the users'

model" was supported by 10 of the 15 papers they reviewed, meriting first place in their list of guidelines. Hansen [32] advocates "know the user — watch him, study him, interact with him, learn to understand how he thinks, why he does what he does" as reviewed by Foley *et. al.* [19].

2.3 Questionnaire

Two options were open for the investigation of user requirements to construct a model of the prospective users of the system. The first was a series of interviews with prospective users. It was thought that this would be too intrusive and that it would be too difficult to persuade a significant sample to participate. The preferred second option was the distribution of a questionnaire to all relevant members of the Department. A questionnaire could reach everyone in the Department, the participants would be able to respond in their own time, and would not be intrusive.

At the time of compiling this questionnaire, most of the computer services in the University were provided centrally by the Edinburgh University Computer Services (*EUCS*). Some departments did possess and run their own machines — for example the Computer Science Department had a network of Sun Workstations and Vaxes whilst the Biology and Physics Departments each had a Vax — but most departments were serviced by EUCS. The main EUCS machine was an IBM-compatible mainframe running the locally developed operating system Edinburgh Multi-Access System (*EMAS*). The EMAS service has since been withdrawn, being replaced by a Sequent Symmetry known as *Castle* running Unix.

Having consulted with a sociologist with some expertise in the field, a questionnaire was designed. This was circulated around the staff and research students in the Chemistry Department. Appendix A contains a copy of this questionnaire.

2.3.1 Aims

A general impression gained from working in the Chemistry Department before the distribution of the questionnaire was that there were four distinct levels of computer users therein. These corresponded to:

1. People who did not use a computer at all, for any purpose. These people seemed to be concentrated in the organic section.

2. People who used one of the two on-line data-bases for literature searches.

3. People who used various utility packages in addition to on-line literature searches.

4. People wrote and used their own programs in addition to the other resources provided.

The questionnaire was designed so that only those questions relevant to each respondent's level of usage needed to be answered. People in category 1 above need only answer questions addressing everyday working practices. People in category 2 would then go on to answer the questions relating to the on-line library catalogues. The third section contained questions concerned with the other tools relevant to those in categories 3 and 4. The aims of each section are detailed below:

Section 1: Addressing chemists, working practices and potential areas for support.

This section was considered to be the most important, giving the most opportunity for voicing opinions and suggestions for support. The aims of this section were:

- to gain information about the working practices of the research chemist for the construction of a user-model, and

- to ascertain which of these practices could be supported by software, and to what extent.

Section 2: Addressing the on-line library catalogue systems.

At the time of compilation there were two on-line library catalogues available, *GEAC* and the Edinburgh University Library Catalogue (*EULCAT)* a multi-user account on EMAS. The aim of this section was to investigate both and establish:

- their relative merits,
- if either was preferred above the other,
- if one system was preferred, why this should be.

Section 3: Addressing tools currently available to the research chemist on EMAS.

This section was designed to establish

- the tools already available to the user,
- the level of user awareness of the existence and/or the function of these tools,
- how any such information was obtained,
- to what extent these tools were used,
- when there were several competing tools why one was preferred to another.
- the perceived advantages and disadvantages of individual tools and
- suggestions for improvement.

To ensure that every aspect was covered and that respondents would not feel limited in any way, specific areas of the questionnaire were provided to voice

opinions on all aspects of the tools — including the user interface, general observations and possible improvements.

2.4 Results

The responses to the questionnaire are shown in Appendix B. The results shown in this chapter have been processed from these responses.

2.4.1 Statistical Analysis of Responses

Table 2–1 shows the distribution of the replies, both from different sectors within the Department — organic, inorganic and physical — and in the hierarchical levels — academic staff, post-doctorates and post-graduates. It can be seen that out of 140 questionnaires distributed, 35 were completed and returned, an unexpectedly high response rate of 25%.

	Staff	P-Ds	P-Gs	*total*
Organic	11	8	35	54
Inorganic	11	12	22	45
Physical	10	6	22	38
Others	3	0	0	3
total	35	26	79	140

(a) Number of questionnaires sent

	Staff	P-Ds	P-Gs	*total*
Organic	3	1	4	8
Inorganic	6	2	6	14
Physical	0	1	11	12
Others	1	0	0	1
total	10	4	21	35

(b) Number of replies received

	Staff	P-Ds	P-Gs	*total*
Organic	27%	13%	11%	15%
Inorganic	55%	17%	27%	31%
Physical	0%	17%	50%	32%
Others	33%	—	—	33%
total	29%	15%	27%	25%

(c) Percentage response for the individual categories

Table 2–1: Statistics, showing the response to the user survey.

The distribution of replies across the different sectors was unsurprising. The response rates from the inorganic and physical sectors were approximately equal at just over 30%. These were expectedly higher than those from the organic sector, where the use of computing facilities was known to be less.

It was however surprising that slightly more staff than post-graduates responded. The preconception was that post-graduate students might be more comfortable with the use of computers, more familiar with current computer technology and more enthusiastic about the capabilities of computers than members of staff. A possible reason for the lower response rates from post-graduates could be that because of their temporary situation students would have little to gain from long-term benefits.

2.4.2 Section 1 — Working Practices

All respondents filled in this part of the questionnaire, detailing the hours spent in activities connected with chemical research. Responses to this part varied greatly in the total number of hours attributed and in the degree of detail given:

Time attributed:

A break-down of the time spent per day was requested. However, since the time attributed by individual respondents varied from three hours to eighteen and a half hours and averaged at just under seven, it was decided to interpret the hours more liberally and take them as a *proportion* of the total time attributed. Since only those activities concerned with research were of primary interest, time attributed to administration, teaching and such was ignored when calculating these proportions.

Activity detailed:

The amount of detail listed in the break-down of research time varied greatly — from generic activities such as "experimental" to specific activities such as "washing glassware" and "running chromatograms". Thus

an attempt was made to interpret the activities into the specific categories of

- synthetic chemistry,
- analytical chemistry,
- theoretical chemistry,
- literature work,
- document preparation and
- interpretation.

Ambiguous activities such as "experimental" were categorised by taking account of the respondent(s) involved and their field of interest.

The results of this processing and interpretation are summarised in Figure 2–1.

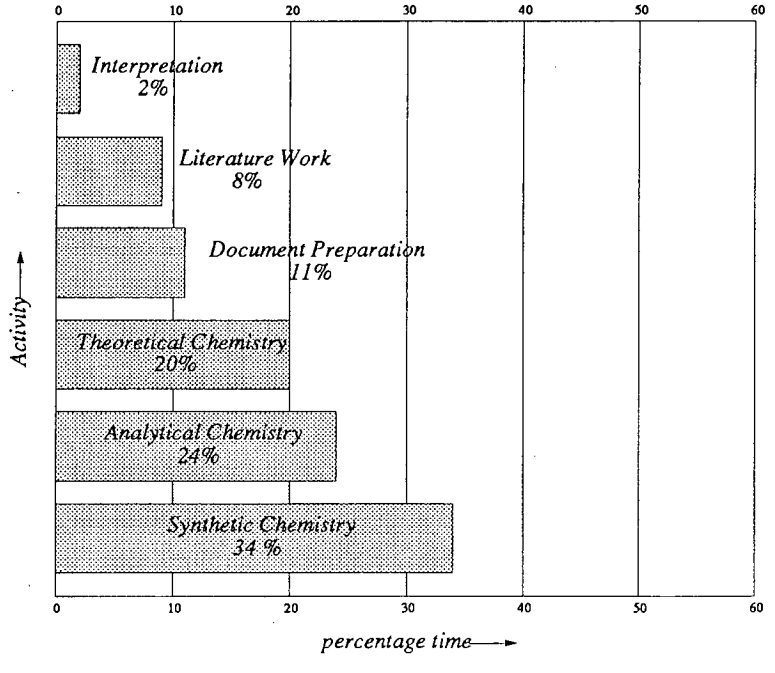


Figure 2–1: *A histogram illustrating the average distribution of time spent on the activities associated with chemical research.*

Synthetic Chemistry

Synthetic Chemistry comprises those parts of chemistry involved in the synthesis of a new compound from reagents, such as the actual synthesis, purification and crystallisation. According to the questionnaire responses and the interpretation thereof approximately $34\% \pm 5$[1] of the working day is spent by the average research chemist in synthetic chemistry.

The respondent cited categories (see Appendix B) which were interpreted as constituting synthetic chemistry were

- Bench-work ($\approx$34%) and
- Washing glassware (<1%).

Analytical Chemistry

Observations made during synthesis may give the chemist some *indication* of the structure of the compound synthesised and the reaction mechanism by which it was formed. However, these are not sufficient to form conclusions. Thus the synthesised compound(s) — and possibly observations made during the experiment — must be passed on for more formal analysis by *Analytical Chemistry*. There are several analytical techniques open to the chemist such as X-ray diffraction, mass spectrometry, infrared (*IR*) spectroscopy, nuclear magnetic resonance (*NMR*) spectroscopy and micro-analysis. Each gives different information and is appropriate for different compounds. By applying a few of these techniques appropriate to the compound being analysed, the analytical chemist can collect enough information to construct a theoretical model of that compound.

Those activities interpreted as instances of analytical chemistry are listed below. These comprise approximately $24\% \pm 4$ of the working day of an "average" research chemist responding to the questionnaire.

---

[1] data are presented as percentage mean $\pm$ standard error.

- Data Collection (≈2%)
- Experimental (≈4%)
- Running Spectra (≈12%)
- Running Chromatograms (≈2%)
- Electrochemistry (≈1%)
- Crystallographic Structure Solution (≈1%)
- Analysing Spectra (≈2%)

Theoretical Chemistry

The term *theoretical chemistry* here encompasses both the formation of a theoretical model to explain empirical results and the application of accepted chemistry theory to predict the results of experiments as yet undone. Theoretical chemistry constitutes approximately $20\% \pm 4$ of the working day of a research chemist. Only one activity was interpreted as theoretical chemistry:

- Calculations (≈20%)

Interpretation

In this context, interpretation is the term given to the processing of information from whatever source. Chemists interpret information to reason about and explain experimental results, build theoretical models and decide what further action is appropriate. Interpretation comprises approximately 2% of the time spent by an "average" chemist responding to this questionnaire.

- Data Analysis (≈1%)
- Computer Graphics (≈1%)

Document Preparation

This category would seem to be self explanatory. According to the questionnaire, the time spent on document preparation by the average chemist is approximately 11%.

Literature Work

This category would seem to be self explanatory. Questionnaire responses indicate that approximately 8% of a researcher's time is spent on activities connected with academic literature.

2.4.3 Section 2 — Library System

20 respondents answered this part of the questionnaire. Table 2–2 shows the distribution of the users of each system. The two library systems appeared to be used equally frequently, although apparently for different reasons. Users of GEAC tended never to have been exposed to EULCAT, whereas users of both tended to prefer EULCAT.

	EULCAT	GEAC	*total*
(a) Used solely	3	7	10
(b) Both used — this preferred	7	2	9
total	10	9	

One user expressed no preference

Table 2–2: *The distribution of users of the two on-line library catalogue systems, EULCAT and GEAC..*

Of the 17 people who expressed an opinion of the ease-of use of GEAC, 9 claimed it to be tedious. Of the 13 people who expressed an opinion of the ease-of use of EULCAT, however, the greater proportion at 6 thought it reasonable and only 3 thought it tedious. GEAC was considered to be too slow and unreliable.

The service was often unavailable, particularly at weekends. There was also a complaint about the inability to type ahead with GEAC. EULCAT, on the other hand, was criticised for being user-unfriendly and for allowing only searches on key-words rather than full titles.

Suggestions for Improvement

Some suggestions were made for features which would be appropriate for inclusion in any new system. These included the ability to output reference material to a printer or to a personal database, recognition of journal abbreviations, and access to on-line chemical, physical, engineering and citation index abstracts. Access to a "Dictionary of Chemicals" index listing IUPAC names, trivial names and physical data was also requested.

2.4.4 Section 3 — Current Facilities

Question 1 — "How often do you use EMAS? (on average)"

22 respondents answered this part of the questionnaire and, according to responses to this first question, most of these — 19 from 22 — use EMAS at least once a day.

Question 2 — "For what purposes do you find EMAS most useful?"

Performing calculations appeared to be the activity for which EMAS was considered to be most useful, being cited as such by 9 people. Table 2–3 shows the activities listed and their frequency of citation by respondents:

Performing Calculations	9	Resources	4	Data Transfer	3
Crystal Structure Solution	6	Graphics	3	E-mail	2
Word Processing	5	Programming	3	Data Analysis	2

Table 2–3: In answer to "For what purposes do you find EMAS most useful?".

Question 3 — "How often do you use each of these facilities?" (and subsequent questions).

In all twenty-three EMAS facilities were examined. When the responses to questionnaire was processed however, it was decided to investigate only thirteen of these more closely — those considered to be most relevant to this project. A measure of the extent to which each of these tools was utilised was calculated according to the following points system:

$$\text{total points} = 0 \times \text{never} + 1 \times \text{rarely} + 2 \times \text{often} + 3 \times \text{regularly}.$$

The results of this calculation for the thirteen selected tools are shown in Table 2–4 and Figure 2–2. A brief description of the function of each and the responses to questions 3–10 are discussed below. It should be noted that those respondents who gave no answer to question 3 are assumed not to have used the tool.

◦ *ICON8 and CCCP*

ICON8 and CCCP are programs which perform Extended Hückel Molecular Orbital (*EHMO*) and Modified Extended Hückel Molecular Orbital (*MEHMO*) calculations respectively. CCCP locates hydride ligands in transition metal cluster compounds with a combination of ICON8 and a simplex optimisation. MEHMO theory and its implementation in CCCP are described in Chapter 5.

These tools were only used by a single person in the Department, ICON8 often and CCCP rarely. The reasons given for the lack of use of these tools were ignorance of their existence (14 and 12 respondents respectively) and that the respondents had no need of them (5 and 7 respectively). The single person who

Tool	Frequency of Use				Points
	never	rarely	often	regularly	
Help	1	7	6	8	43
Edit	8	1	2	11	38
Scribe	7	6	5	4	28
Easygraph	9	7	3	3	22
SHELX	14	0	3	5	21
CALC	14	1	2	5	20
Pluto	14	2	2	4	18
ECCE	16	1	1	4	17
Ortep	14	1	6	1	15
VECCE	19	0	1	2	8
μEmacs	20	2	0	0	4
ICON8	21	0	1	0	2
CCCP	21	1	0	0	1

Table 2–4: *An illustration of the extent of use of some of the tools on EMAS, listed in order of popularity. The system for the calculation of the total points is described in the text.*

did use the tools found them tedious to use (ease of use) but of a "reasonable" and "good" standard (functionality) respectively.

- *CALC*

CALC is a locally developed molecular geometry tool. This tool can perform a variety of geometric operations on molecular models. Examples of this include the conversion of models from one coordinate system to another — this process, and the need for it is described in more detail in the next chapter — and the addition of atoms and functional groups.

8 of the 22 respondents used this tool, 5 of them regularly. Of the 14 people who never use CALC 9 had never heard of the tool, 3 had never needed to use the tool and 2 did not specify any reason. Of the 8 who did use the tool, 7 found out about it from a fellow student or member of staff and the remaining 1 did not answer this question. 3 found it tedious to use but 4 found it reasonable and 1 simple.

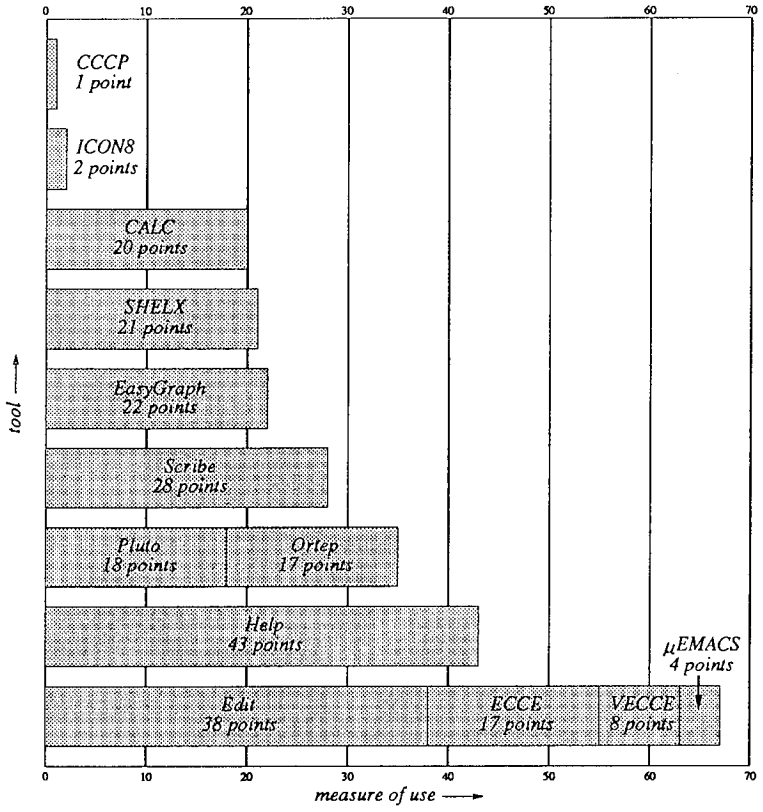


Figure 2–2: *A histogram illustrating the extent to which a selection of tools on EMAS are used.*

o *SHELX*

SHELX produces a molecular model from data output by an X-ray diffractometer to aid the interpretation thereof.

8 of the 22 respondents used this tool, 7 of whom were also users of CALC. 5 of them used SHELX regularly and 3 of them often. 5 found it reasonable to use, 1 tedious, 1 difficult and 1 simple; 3 thought the standard could be improved. Of the 14 who did not use it, including 3 who did not answer the question, 8 had never heard of it and 5 expressed no need to use it, whilst the remainder did not answer the question.

o *Easygraph*

Easygraph, as its name would suggest, is a tool for drawing simple graphs.

9 people, including 2 who did not answer the question, had never used Easygraph. Only 1 had never heard of the tool, 4 had no need to use it either at all or more, 1 had tried it and found it too difficult and 3 simply did not know how to use it. Of those who did use Easygraph, 10 had found out about it from a fellow student or member of staff, and only 3 from official EMAS services.

- *Scribe*

Scribe is a document preparation system.

Only 1 person had never heard of this facility, although 7 had never used it, including 3 who did not answer the question. Other reasons for not using it, or not using it more often, included 6 citations for having no need to use it more regularly, 4 for finding it too difficult and tedious and 2 for not knowing how to use it. Again, the majority of people at 13 out of 14 learned of its existence from fellow students or members of staff. Of the 15 who do use it, 2 found it difficult to use and 6 tedious. 2 thought the standard of the tool itself was good, 2 reasonable, 3 that it could be improved and 2 bad.

- *Ortep and Pluto*

Ortep and Pluto are molecular visualisation tools. Molecular visualisation is discussed in more detail in Chapter 4

It would seem that these two tools are in direct competition, since Ortep and Pluto have similar functionality, although Pluto can only display atoms as isotropic spheres whereas Ortep can display them as thermal ellipsoids. The tools appear to be used equally frequently with 8 users each; 3 and 2 respondents respectively did not answer this question. Of the 8 respondents who had used either of these tools, most at 6 learned of that tool from fellow students or members of staff. These respondents thought that the tools were tedious to use — 3 for Ortep, 2 for Pluto — or reasonable — 5 for Ortep, 6 for Pluto. None considered their use to be simple. Of the standard of the tools themselves, 1 thought that they were bad in each case, 3 and 4 respectively thought that they

could be improved, 2 and 3 that they were reasonable and only 2 and 0 that they were good.

◦ *Help*

Help is the on-line local help system, giving information on all the centrally provided and some local tools.

Unsurprisingly all but 1 of the respondents have used the help system at one time or another. 13 of these users learned of its existence from fellow students or members of staff. Most at 15 thought that the system was easy to use and 3 thought that it was reasonable with only 3 thinking it was tedious.

◦ *ECCE, Edit, μEmacs and VECCE*

These are the text editors available on EMAS. Edit and ECCE are line editors. Only one line of the text is viewed at a time; any updates thereto are not displayed automatically but must be explicitly requested.

Again these tools are considered together since they would seem to be in direct competition. As Table 2–4 and Figure 2–2 show, Edit was by far the most popular tool with 11 regular users. It would seem that the main reason for μEmacs not being used is that people did not know of its existence. As reasons for not using VECCE, equal numbers of respondents at 6 cited not knowing of the existence of the tool and not knowing how to use the tool. Regarding ECCE, approximately equal numbers at 6 and 5 did not know how to use this facility and had no need of it. It is interesting to note that although Edit was the most widely used tool it was also the most frequently cited as being too difficult and tedious to use.

◦ *General*

One feature which was apparent during the analysis of responses to the questionnaire was the manner in which people found out about the tools they use. To illustrate this more clearly the responses to question 7: "How did you find out about this facility?" are summarised in Table 2–5.

From fellow students or members of staff?	81%
From *User Support*?	4%
From the *EMAS* user guide?	3%
Other	12%

Table 2–5: In answer to "How did you find out about this facility?".

Suggestions for Improvement

Some of the most appropriate requests for improved and new tools are categorised and paraphrased below:

- *Graphics*

Respondents requested a three dimensional interactive molecular graphics facility with explicit orienting commands, the ability to handle at least 500 atoms and allowing simultaneous examinations of both structure and electron density maps. For more general graphics facilities, requests were made for options to produce graphs with multiple plots, graduated scale along two sides of the VDU and of mathematical functions.

- *Document Preparation Facilities*

Respondent requests included the ability to insert external files such as graphics files, tables, and data files, a draft facility, and more information on the text layout of each environment option such as "report" or "article". An integrated graphics program/word processor, a previewer, and the ability to select landscape or portrait modes were also requested.

- *Crystal Structure Solution:*

There were several requests simply for the maximum automation of SHELX and better integration of structure solving/refinement.

o *Editors*

An editor with a spell-checking facility with reference to a private dictionary which recognises "phrases" such as word-processing commands and chemical compounds and terms was suggested.

o *User Interface*

Several respondents requested better data input with more flexible formats for input files, generally improved presentation of results and the ability to pass data directly between programs.

o *Documentation / Help*

There were many requests simply for better and more readily available documentation for, and information on the availability of, the tools and devices provided — both on-line and as hard copy. More specific requests were for on-line examples classes and a layered, menu driven help system.

o *Miscellaneous:*

More sophisticated software for acquiring, storing — *e.g.* databases — analysing, plotting, manipulating and comparing experimental data — *e.g.* by spectra and chromatograms — or theoretical data was requested. Computer Aided Design (*CAD*) facilities were also requested, both for general mechanical and electronic design and drafting, as was software for calculation of "g" values for EPR spectra and refinement of experimental parameters and a better calculator.

2.5 Discussion

2.5.1 Limitations of the Questionnaire

Demise of EMAS and Non-Use of On-line Catalogue Data

Superficially, it may seem the move of the central computer service from EMAS to Castle would render much of the questionnaire irrelevant. In particular, the whole of section 3 of the questionnaire was devoted to the investigation of the tools provided on EMAS. However, responses to this section of the questionnaire are still largely relevant since:

- most of the tools provided on EMAS services have been transferred to Castle in support of the people who have been using EMAS for several years and who were expecting to be unaffected by this move, and

- the questions and more significantly the responses thereto could be interpreted more generally and apply almost equally well to Castle.

Similarly the results of that part of the questionnaire which investigated the on-line library services can be interpreted more generally. As discussed later, the on-line library catalogues were not chosen to be included in the chemical tool-kit as part of this project. However, the responses to the questionnaire were considered useful, in particular those addressing the reasons as to why one system is preferred over the other and their perceived relative merits.

Lack of Communication

The questionnaire was designed to obtain the maximum amount of information. Sometimes it was considered necessary to suggest the nature of the information sought with examples. It was hoped that these examples would be taken as

such and not queries about specific tools. The questionnaire was not completely successful in this respect. Respondents sometimes simply agreed or disagreed with example suggestions:

In answer to "Are there any of the activities mentioned above which could possibly be benefited by using a computer? *[such as a) the comparison of spectra with theoretical models or from a library of spectra, b) complex mathematical calculations, c) storage of information etc.]*"

> *"Yes — library spectra would be a good idea, as would information retrieval. I don't use complicated mathematical calculations every day."*
> Respondent 23

This tendency is illustrated by the three most frequently activities cited by respondents in section 1 of the questionnaire being those given as examples (see Appendix B).

Limited Knowledge of Potential Computer Applications

Some of the respondents tended to be limited in their requests for computer support by what they already knew to be possible — as the adage says, "you never miss what you've never had". Requests were often framed in the context of existing tools, suggesting specific simple improvements to existing tools rather than letting their imagination "run riot" as had been hoped. For example, one respondent to the questionnaire wrote:

> *"A 3-D fully interactive version of Pluto would be a great advantage, especially if it were one of the CCP4 versions which would allow the simultaneous examinations of both the structure and electron density maps. Failing this, a version of Pluto allowing at least 500 input atoms (not including symmetry-related ones) would be better than the present version."*
> Respondent 16

Similarly, where software was considered difficult to use, respondents would often request a better, more substantial manual or better information about the tool rather than more fundamental improvements in the usability of the tool itself. This tendency had been observed previously while working in the

Chemistry Department. Some of the tools available in the Department seemed to be so complex in their use that manuals have to be regularly consulted. For example, the ICON8 user manual dedicates many pages to explaining how to set up an unnecessarily complex input file. This subject is discussed in more detail in Chapter 5, Section 5.3.4. Dependency on manuals was further illustrated when a copy of the molecular visualisation tool discussed in Chapter 4 was sent to a former member of the Chemistry Department now overseas. When the tool was first offered his response was:

> [sending the program] *would be "ace" — a manual would be "acer".*

Although these effects were not entirely desirable they were not surprising and may have been owing in some part to the wording of questions. Most of the suggestions of this sort were still considered to be valid in that they could be applied more generally. The requests and suggestions regarding specific tools highlighted the specific problems with tools. Moreover, they were useful indications of pitfalls to avoid in the design of future tools. The request for manuals and more information, although sometimes valid, were less helpful. In such cases the specific problems remained unclear. There was only the indication that users were experiencing problems of an unspecified nature with these tools.

Inappropriate Respondent Suggestions

Those respondent suggestions which are inappropriate must first be eliminated to isolate those areas which would most benefit from additional software support.

Many of the suggestions for improvements are already commonplace on other systems. For example, those respondents requesting "a better calculator" would be adequately provided for by a choice of several different calculators if using Unix, the X Window system and/or Gnu Emacs, all commonly used standard software. This same standard software would also satisfy those respondents

requesting improved editors and better graphics screens; the X Window system runs on machines with bit-mapped screens. The need for improved word-processing with the ability to import external text or graphics files and most of the other word-processing suggestions would be met by the commonly available document formatting system LaTeX, with which this thesis has been prepared.

Other suggestions were considered outwith the scope of this project. These include requests such as "more regular archiving"[2], "better back-up services" and "more training in basic computing".

Finally some respondents requested software support for areas of research which, although perfectly valid as potential areas for future work, were considered too specialist to be of general use.

Difficulties of Interpretation

Some of the working activities cited in section 1 of the questionnaire proved difficult to interpret. Activities such as "experimental" could be placed into any of the *synthetic, analytical* or *theoretical* categories, whilst "calculations", one of the examples presented, could be *theoretical, analytical* or perhaps *interpretive.* Prior knowledge of the respondent could be used, but it is acknowledged that the categorisation is somewhat subjective. The resulting figures are therefore only considered to be rough estimates of true behaviour.

The questionnaire responses did nevertheless shed light on some areas which could be better supported by software and gave insight into the way in which research chemists spend their time. There were several useful suggestions for improvements to existing software and ideas for new tools.

---

[2]users of EMAS could mark a file to be archived which would then be removed and stored on tape within 24 hours. However, archiving is not a solution to a user experiencing an immediate lack of file-space

2.5.2 Tool Support for User Activities

In this section the activities mentioned previously are discussed together with the software tools which may give some support to these activities. Some tools may support more than one activity. Such tools are discussed more than once, each time in relation to the support given to the particular activity under discussion.

Synthetic Chemistry

There are no computerised tools available to research chemists at Edinburgh which are considered to support synthetic chemistry specifically. In fact, synthetic chemistry is probably that part of chemistry least open to software support. Perhaps the most feasible would be the use of molecular graphics and animation to propose and plan reaction mechanisms for later experimentation. Alternatively, computers might monitor a reaction and control temperature or conceivably react in a certain way if a change were detected. However, the manner in which reactions proceed will not always be known in research chemistry. Unless the reaction is to be repeated many times in bulk, as in a chemical engineering context, this form of support is currently not considered to be practical. Other less specific applications of computers might include word processing to maintain laboratory log books.

Analytical Chemistry

Analytical chemistry would seem to be highly automated. Analysis is often done by machines and most modern machines have the capability of producing digital output. In the Chemistry Department at Edinburgh for example, X-ray diffraction is performed on an X-ray diffractometer and the results of the analysis are presented as a computer data file. The most common output of an infrared spectrometer is a chart representation of the absorption by the compound of a continuous range of wavelengths of infrared light. However, these data are

often also sent to an outport of the machine, ready to be interpreted directly by computer. The capability for such direct data transfer opens many opportunities for subsequent processing by computer programs. The immense amount of information produced by the diffractometer may be processed automatically to produce a molecular model, *e.g.* with SHELX. Similarly, the digital storage of infrared spectra creates many opportunities to build databases thereof. This facilitates an automatic comparison to search for compounds with similar spectra to be plotted together, perhaps highlighting key areas.

o *Existing Tool Support*

SHELX is the only tool considered to offer specific support for analytical chemistry:

Crystal Structure Solution (SHELX):

There would appear to be a thin line between the support for *interpretation* and *analytical chemistry* provided by SHELX. One one hand, it could be considered to support interpretation in helping the user to interpret data produced by the X-ray diffractometer, whilst on the other it could be regarded as part of the analytical chemistry process. However, it is not entirely appropriate to discuss the support provided by SHELX to these activities separately, since it is the same *function* of SHELX providing this support in either case. Since SHELX appears to be an integral part of the X-ray diffraction process at Edinburgh it is perhaps most appropriately thought of as supporting the analytical process.

The questionnaire responses about the usability of SHELX seemed to be centred on "reasonable", although not all respondents concurred:

> *"SHELX is inflexible and cumbersome when dealing with non-routine cases ... modelling disorder in molecules containing symmetry elements is cumbersome and potentially confusing."* Respondent 3

> *"On the crystallography side this* [more readily available and accessible information] *would allow "person 1" and "person 2" to do what they are best at i.e. crystallography & the research student*

> *could learn more about computing by being able to accomplish more for him/herself"* Respondent 11

> *"SHELX: . . . more automation of the system should be available preferably in the form of questions like 'Do you want to update your channel 5 (input) file Y/N?'."* Respondent 23

Not all the atom positions are determined simultaneously in the solution of a crystal structure. The process is iterative, with several atom positions being suggested to the user after each iteration. The user then selects those positions considered likely to be correct whilst rejecting others. The selected atom positions are fed to the program for subsequent refinement in the next iteration to locate further atom positions. A list of atomic coordinates is gradually constructed in this manner. Suggested atom coordinates are presented as a labelled dot-to-dot style diagram in the SHELX implementation, with a key for cross-referencing. Appropriate atom(s) must be manually selected from the atomic coordinate list in the SHELX output file to specify which atom positions are to be retained — hence the above comment from respondent 23.

Both the presentation of information to the user and the method for selecting atoms to be retained in SHELX would seem to be unsatisfactory. Users must spend unnecessary time addressing the mechanics of the program rather than the underlying chemistry. Furthermore, the poor visualisation and unnecessary manual editing leaves great scope for error.

Better visualisation facilities could be provided by presenting users with a more natural ball-and-stick style picture which could be interactively rotated in three dimensions. Better selection facilities could then be achieved by allowing users to select an atom for retention or rejection with a simple pointing device, pressing a button on completion to initiate the next refinement iteration. On comparison, the facilities offered by the SHELX would appear to be error-prone, time-wasting and open to extensive improvement in at least two areas — automation and visualisation.

This case study indicates that the great potential for software support of analytical chemistry has not yet have been realised. For example, it would appear that SHELX is functionally useful but unnecessarily difficult to use.

Theoretical Chemistry

Theoretical chemistry is also amenable to software support. The construction of some models can require large sets of calculations. For example, EHMO calculations must consider all pairs of atoms in the molecule, whilst finite difference simulations of electrochemical systems must consider all discrete elements over discrete intervals of time. Calculations of these types may be performed manually, but would be necessarily simple or coarse and prone to human error. More complex calculations and finer iterations are necessary for more realistic simulations and models, and thus a computer is required.

- *Existing Tool Support*

Molecular Geometry (CALC):

CALC allows the incorporation of extra atoms and functional groups into existing models. The tool appears to be reasonably well used considering that it is only appropriate to a small section of the Department.

EHMO Calculations (ICON8 and CCCP):

ICON8 and CCCP support specialised areas of chemistry. The low frequency of their use is therefore unsurprising. The one respondent who did exploit these tools found them tedious to use. The author has experience of both of these tools and found their use difficult, requiring manual consultations each time they are used. The user interface to one of these tools, CCCP, is described in more detail in Chapter 5, Section 5.3.4

Again, whilst functionally useful these tools are somewhat lacking in terms of ease-of-use and visualisation.

Interpretation

Interpretation is a further section of chemistry open to software support. The style of data presentation, from whatever source, can be of the utmost importance in ensuring an efficient interpretation. Current scientific visualisation literature is reviewed in Chapter 4. Graphical data presentation is generally considered to aid interpretation more than textual presentation. A list or matrix of numbers may be best presented as a graph. A molecular model may be presented graphically to give important visual cues to the 3D spatial relationship therein, with representations such as ball-and-stick, space filling or solvent-accessible surfaces. Computer graphics may be exploited to aid the chemist in visualisation of results, proposed models and comparisons between theoretical and experimental results.

- *Existing Tool Support*

Molecular Visualisation (Ortep and Pluto)

At the time the questionnaire was distributed there were two molecular visualisation tools available in the Chemistry Department at Edinburgh — Ortep and Pluto. They seem to be rather difficult to use, requiring some computing expertise, and are not part of an integrated system. Each requires a certain style of input file and must be used in isolation. Other tools cannot directly access any visualisation facilities which these offer. At best, other tools which provide or deal with molecular models can only provide indirect visualisation facilities by being obliged to output to a file. These files may require editing to an acceptable form to be read by the molecular visualisation tools.

Molecular Geometry (CALC)

CALC does not aid interpretation graphically, but rather by providing such information as bond lengths, angles and torsions.

Graphics (Easygraph)

Despite being the only tool available to chemists for drawing simple graphs, a facility which would seem to be fundamental to the work of many research chemists, Easygraph would not appear to be particularly easy to use:

> *"Easygraph— I hate it because it often takes ages to get it to do what you want it to do."* Respondent 30

Yet again, whilst functionally useful, these tools are lacking in their ease-of-use and visualisation.

Literature Work

There are many ways in which literature work may be supported by software. For example, it should be possible to search by topic, keywords, authors or title on-line for books, journal articles, articles in conference proceedings to automatically update a personal bibliography. Copies of relevant articles could then be extracted from a global database to be stored in a personal database or for printing. This scenario might not be implemented fully, both because of legal considerations arising from copyright law and pragmatic reasons since the maintenance of such a database could be involved. However, even a partial implementation may be desirable.

Library Catalogues (EULCAT and GEAC)

The systems currently available fall far short of the ideal. Users complained about only being able to use (a maximum of two) keywords in EULCAT. This restriction can make it difficult to be adequately specific when searching for journals with titles composed solely from common words such as *"Computers and Management"*, *"Science"* or *"The Computer Journal"*

> *"Checking for titles is not easy on EULCAT if you have a common word in the title, as it doesn't appear to take full titles but only keywords"* Respondent 23

Neither of these systems makes any provision for the location of journal or proceedings articles on specific subjects. Since most up-to-date scientific research is published in such articles, this would be a highly desirable feature. Several respondents requested on-line chemical abstracts, which, it is believed, are available on other sites and would allow such searching:

> *"Without a doubt, the best improvement that could be made would be for all users to be able to "log-on" for on-line Chemical Abstracts literature search, using e.g. reference code number and keywords, phrases etc."*
> Respondent 30

Document Preparation

Document Preparation Systems (Scribe)

Although a word-processing system is available to chemists, a significant number of postgraduate students choose to write their thesis long-hand to be typed professionally. Their reasons for this could simply be an unwillingness to use computers. However, they could also arise from the standard of the single available system:

> *"I find Scribe confusing and unsatisfactory to use, and its documentation isn't clear. Other chemists have commented about problems with referencing, producing tables etc. ... I can't even centre a numbered subheading"*
> Respondent 6

Editors (Edit, ECCE, μEmacs and VECCE)

As described previously, there were four text editors available to users of EMAS. Two of these were the line editors Edit and ECCE, whilst the other two were the screen editors VECCE and μEmacs:

μEmacs:

This tool appeared to be the best both in its usability and functionality. It has a menu controlled interface and provides many functions not provided by the others, such as context sensitivity — *i.e.*, μEmacs could recognise the "type" of a file (C, Fortran, LaTeX, *etc.*) by its file name extension and adapt its "mode" accordingly.

VECCE:

VECCE, on the other hand, does not have as many facilities and relied on a command-line interface. As a result, the user had to learn basic commands before the tool could be used and the more advanced commands before the tool could be used to its full advantage. However, VECCE could be considered to be preferable to ECCE and Edit since cursor keys could be used to browse through the document and any alterations could be seen immediately — a feature recommended by HCI research.

ECCE and Edit:

Both ECCE and Edit appeared to be difficult to learn and use. Each had complex commands which had to be learned before the tool could be used. Each only displayed one line of the document at a time. The user had to explicitly request the redisplay of any modified lines of text.

2.5.3 Support Relative to Origin of Software

It appeared that the tools which were available to chemists using EMAS can be classified into two categories. Each tool may be classified into categories of "generic software tools" or "specific chemical software" according to its *origin* — that is, its developer and/or provider. Certain problems appear to be common amongst tools within each category.

Generic Software Tools

The first category contains those facilities which were provided by EUCS as a service to all users of the system, such as editors, mail facilities and graph drawing programs. These tended to be well supported by both a reasonably easy-to-use help system and hard-copy documentation. User support teams

were available to specific subject groups. All this support was provided by EUCS.

What may have been lacking was information about the *availability* of the tools provided. Given knowledge and information about the tools available for a particular activity, users may make objective decisions about the most appropriate tool to use. However, the responses to the questionnaire strongly confirmed a suspicion that by far and away the most common mechanism by which knowledge of the existence of a tool is disseminated is by *ad hoc* word of mouth. As a result, users are often unaware of alternative tools to perform a given task. This was clearly illustrated in a response to the questionnaire:

> *"The unfortunate problem with any information system is that unless you know it exists you can't exploit it ... I found the way I learned was directly from people about me, that is learned the folk-lore of all the smart tricks from the grape-vine of accrued common knowledge."* Respondent 30

For example, if the editors are ranked simply on the criteria of usability and facilities offered, it would seem that the popularity of the tools should have been, in descending order:

1. μEmacs
2. VECCE
3. ECCE and Edit.

Paradoxically, it was found that the *actual* usage of the tools was the reverse of that "predicted" above. As summarised in see Table 2–4, the questionnaire responses showed that Edit was the most commonly used editor with 38 points and a total of 14 users, second was ECCE with 17 points and 6 users, third was VECCE with 8 points and 3 users, and finally came μEmacs with 4 points and 2 users.

The reasons for the non-use of the "better" tools appeared to be a lack of awareness of their existence, or at least a lack of awareness that they *were* easier to use and provided greater functionality. The example in the introduction to this chapter and the following quote from a questionnaire respondent illustrates that

users are enthusiastic about new, better tools when introduced to them and are more than willing to use them:

> *"Not enough publicity in the Chemistry Department about this system* [VECCE]" Respondent 23

It is acknowledged that some people who had been using the same editor for several years might be reluctant to switch to a new tool. However, new users to the Department and the available software tools should have no such inflexibility. Because of ignorance of other available options, these people often use the same editor as their peers or supervisor. The original tools tend to remain the most popular whilst the newer, more user-friendly and more powerful tools remain under-exploited. Perhaps the available resources might be better utilised if the EUCS support was more *proactive* in supplying information automatically to any new user rather than passively *reactive* in only supplying this information on demand.

Specific Chemical Software

The specific software available to chemical researchers at the University of Edinburgh seems to suffer from different problems to those described above. Lack of awareness of the existence of these tools is not a problem owing to their specialist nature. These tools are often only of use to a limited number of researchers whose interests lie in a common research area. Word-of-mouth dissemination of the availability of a tool would seem to be more efficient in a smaller user community.

The major problem with the chemical software at the University of Edinburgh would seem to be of inadequate usability, as detailed in Chapter 1.

In Chapter 1 it was identified that whilst many of these tools could be functionally useful, often performing a function which would be practically impossible to achieve otherwise, their designs seemed to pay little attention to human-computer interaction or software engineering issues. Because chemists are

trained in chemistry and not computer science, the chemists writing these specialist programs may be unaware of accepted software practices in computer science. Thus, although the tools perform the desired function, they are often cumbersome and non-interactive, lack intuitiveness and have awkward input and output conventions.

2.6 Conclusions

It must be concluded that the questionnaire was of limited success and proved less useful than was expected. It was hoped that the respondents would be more imaginative, perhaps to the extent that some requests would be yet technologically or practically unachievable, such as the automatic production of solid molecular models by stereo lithography. However, respondents tended to raise specific requests for limited extensions to existing tools. Nevertheless, although of limited utility, responses to the questionnaire did shed light on some areas. Previous suspicions about the state of chemical software regarding human-computer interaction issues were confirmed. Most of the suggestions complied with accepted HCI principles. Several themes recurred throughout the requests in the questionnaire responses, such as facilities for three-dimensional graphics, interactive programs and improved documentation and help facilities.

2.6.1 Current State

The current state of software support may be illustrated with an example. Suppose a research chemist reaches a stage in a project which potentially could benefit from full or partial software support. This might involve repetitive numerical calculations or the visualisation of results. If the researcher wishes to exploit a computer to perform this task, one of several problems may currently be encountered:

1. The necessary software is not available. In this case, the chemist must spend valuable time writing custom software or revert to performing the task manually.

2. The software exists, but is inefficient, inadequate, difficult to use and time-consuming. For example, the only calculator provided on EMAS is textual and accepts only reverse Polish format. The word processor Scribe provided on EMAS has no graphics or tabulating facilities and no way of incorporating any graphics files — any diagrams or tables must be physically cut-and-pasted. The two molecular-drawing programs are non-interactive, cumbersome and difficult to use.

3. Adequate software exists but is not used by the chemist. There are two potential reasons for this. Firstly, the chemist may be unaware of the tool's existence. The availability and use of tools appears to be learned by word of mouth within the Chemistry Department. Secondly, the chemist may be reluctant to employ an unfamiliar tool in preference to an inferior one already used and understood.

4. There is minimal integration of tools for users employing more than one program. Files must be manually edited if those output from one tool are to be input to another.

2.6.2 Choice of tools to be implemented

After considering the questionnaire responses and drawing on personal experience, it was decided to concentrate on two tools which, if implemented appropriately, would enhance the software environment significantly.

Responses to the questionnaire suggest that interpretation is not particularly well supported in the Chemistry Department at the University of Edinburgh. The two tools which support molecular visualisation, Ortep and Pluto, are perceived as cumbersome and difficult to use and are non-interactive. It is

proposed that these difficulties limit the exploitation of benefits available from these tools to a lower level than their full potential.

The implementation of a molecular graphics facility was considered a priority, being a generic tool of use to most chemists. As stated earlier, respondents to the questionnaire criticised the two molecular graphics programs available on EMAS for being non-interactive, tedious and difficult to use. It would seem that a new interactive, "user-friendly" program would improve the available tool-kit. Molecular visualisation is considered an important contribution towards the interpretation of many analytical results. A molecular visualisation tool is considered to constitute the basis of the proposed tool-kit. Such a visualisation facility should not merely directly read data output from other tools but should be directly available to tools. This would enable the user to take advantage of visualisation facilities at *any* appropriate stage.

The second tool selected to be implemented was a hydrogen location facility. The reasons behind this choice were that it typified a computational chemistry program, personal knowledge of the limitations of an existing implementation of such a tool, and an interest in the underlying chemistry.

2.6.3 Requirements

A list of requirements was compiled which related generally to interface issues and specifically to molecular graphics and hydrogen location. The categorised requirements were:

User Interface

1. A direct manipulation interface.
2. Free format input files.

3. Greater integration of programs, to share I/O file formats, data structures and/or code where appropriate. As a schematic example:

Crystal structure solution	→	Structure refinement	→	Result presentation	→	Document preparation

4. Visualisation and appropriate presentation of results.

Molecular Graphics

Most of the requests relating to molecular graphics concerned user interface issues such as direct manipulation, which have already been discussed. However, there were some requests specific to molecular graphics:

5. Three-dimensional molecular graphics.
6. Ability to handle at least 500 atoms.
7. Simultaneous examination of both structure and electron density maps.
8. Ability to incorporate output within a word-processed document.

Hydride Ligand Location

There were no suggestions specifically related to hydride ligand location.

2.6.4 User Model

Figure 2–3 proposes a representation of the manner in which a chemist undertakes work. It was constructed from a combination of the responses to the questionnaire, communication with some researchers from the Department of Chemistry and personal experience. It is not meant to be a definitive statement of the manner in which research is undertaken by every chemist. No suggestion is made that behaviour which does not follow this pattern is in any way incorrect. The model is simply an attempt to encapsulate the working practices of a chemist as an aid to understanding for the benefit of the author and readers of this thesis. Some chemists may recognise their own work in the whole of this

model, some in only a part and some not at all. This model is meant to represent an "average" or "typical" chemist. It has not been constructed so that only a chemist who exactly fits this model will benefit from software support based thereon, but rather so that software support will benefit many chemists in much of their work. It is acknowledged that the model may be biased towards inorganic chemistry. However, every attempt has been made to produce a model to describe the work of a *generic* chemist, avoiding the exclusion of any particular group.

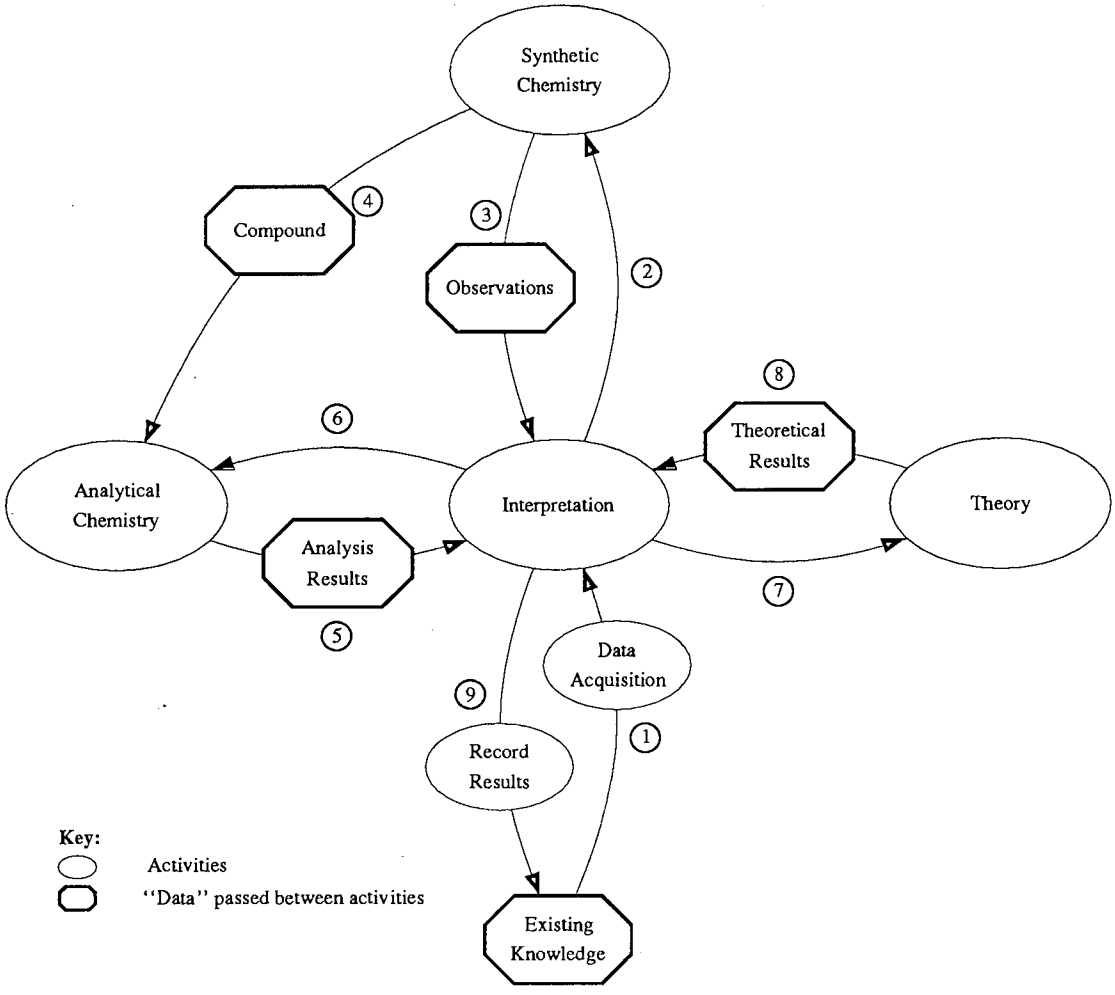


Figure 2–3: User Model.

It may be useful to illustrate the user model with an example of a possible route of the behaviour of a chemist in this model. Whilst navigating this route, the

activities undertaken (as shown by ellipses) may be described, together with the products of these activities (as shown by bold octagons) and their interactions. The sequence of numbers in Figure 2–3 indicates the route through the diagram described below:

Data Acquisition

The idea for a project or a new stage in an ongoing project will often come from a combination of existing knowledge, both internal and external. Internal information is considered to be derived from the experience accrued during the lifetime of the chemist, both chemical and otherwise. This would have been obtained over many years from a variety of sources such as previous experiments performed by the chemist or colleagues, information passed on by other chemists or gleaned from documented scientific literature. External information is considered to be that gained, for example, from documented scientific literature to supplement internal knowledge when researching a new project.

Interpretation

The chemist interprets the information encapsulated within the acquired data to plan further research and thus decide the next appropriate stage of the project. Interpretation is considered to be the mainstay of the work of a research chemist, or any researcher for that matter. Interpretation is the essence of the research undertaken by the chemist. It requires human intelligence, rationalisation and understanding. Without this activity there would be no progress in chemistry. The chemist must interpret any information obtained from whatever source to determine the next appropriate step in the research. Interpretation therefore occupies a central point in this diagram, through which all routes between aspects of chemistry pass.

Synthetic Chemistry

The further research planned during interpretation may involve the synthesis of a compound such as a transition metal cluster compound. The chemist may make observations during the synthesis such as a colour change or solubility in a particular solvent. These observations are *interpreted* to deduce the occurrence of certain processes. For example, a colour change may indicate oxidation or reduction of metal ions in solution. At this point the chemist will have an idea as to what the compound could be and from what it is formed — the reagents and the intended product are known.

Typically, the compound will then be analysed. This is commonly undertaken to confirm the constitution of the compound and to establish bonding interactions and structural features therein.

Analytical Chemistry

A variety of techniques are available for the analysis of chemical compounds, a few of which are described below. Each affords the chemist different information, allowing a picture of the structure of the molecule to be gradually constructed. Typically, there will not be any definite plan of the type or order of the exact analytical techniques to be performed. Rather, the chemist will have an idea of which techniques would be appropriate. The *interpretation* of results from one technique may suggest that undertaking another would be beneficial. Examples of such procedures are:

- *Microanalysis* — to determine the proportions of such elements as *C*, *N*, and *H*.
- *Infrared spectroscopy (IR)* — to determine the presence of certain functional groups.
- *Nuclear Magnetic Resonance (NMR)* — to gain insight into the symmetry of the compound and determine whether key functional groups such as hydride ligands are contained therein.

- *Mass spectrometry* — to determine molecular weight and information about functional groups.
- *X-ray diffraction* — to ascertain the spatial relationships between atoms contained within molecule by the construction of a molecular model with the help of tools such as SHELX.

For example, the presence of hydride ligands in a synthesised transition metal cluster compound may be inferred by IR spectroscopy and strongly supported by NMR. However the establishment of the exact molecular skeleton by X-ray diffraction followed by the use of a semi-empirical method such as Hydex or CCCP could ultimately be necessary to yield the total molecular stereochemistry.

Record Results

Armed with the knowledge so accrued, the chemist may decide to contribute to the existing recorded knowledge by documenting the results of this project and perhaps submitting a paper to an academic journal. In this process the chemist may incorporate images with the aid Ortep or Pluto into text produced with Scribe.

The chemist may now further research by repeating this process.

The distribution of a questionnaire and subsequent processing of the responses has shed light on many areas of the use or computers and associated software in the Department of Chemistry. The knowledge accrued by the user requirements analysis and the user model so constructed may now be applied to the design and implementation of the tool-kit and selected tools.

Chapter 3

Design of the Tool Kit

3.1 Introduction

Issues of both *internal* and *external* design must be addressed by designers to ensure software is of a high quality. Interest in the *internal design* of a program is likely to be limited to "designers" and "programmers" — people who may want to modify the code in some way, rather than simply use it. It determines how adaptable and maintainable a program may be — how new features may be incorporated *(extensibility)* and how easily the program may be run on machines other than that on which it was developed *(portability)*. Users of the program, however, are more likely to be interested in the external design of the program. The *external design* determines the manner of user-computer interaction, the ease and feeling of satisfaction experienced by the user in this interaction, and how these may be maximised.

The procedures for producing well designed software both internally and externally have been well researched in the fields of *software engineering* and *human-computer interaction* respectively. These procedures are discussed in more detail in the overview of section 3.2. It should be stated that this chapter is concerned only with the tool-kit as a whole — those features of the tools which bring them together to constitute a tool-*kit*. The more specific features of each tool are

discussed in the relevant chapters. A series of requirements and objectives is constructed from the issues identified as important and relevant to this research in this section, together with the results of the user requirements analysis in the previous chapter. The implementation of the tools is discussed, as is the manner in which the principles introduced in the overview have been addressed. The evaluation of the tool with regard to original objectives is left to later chapters where the implementation of the tools is discussed.

3.2 Overview

HCI and software engineering factors for both internal and external design are well documented. Whilst they may influence each other or work towards the same goals, these factors can sometimes be competing. It can be difficult to simultaneously reconcile those of HCI with those of software engineering. Figure 3–1 identifies various different factors. These are split into *goals* for those considered to be an end in their own right and *techniques* which are considered to be a means to those ends. The individual factors are discussed in more detail under the relevant headings in subsections 3.2.2 and 3.2.3.

3.2.1 General Issues for Effective Design

Technique

- *User Requirements Analysis*

User requirements analysis has been discussed in more detail in Chapter 2. Its contribution to the functionality of a tool is discussed below.

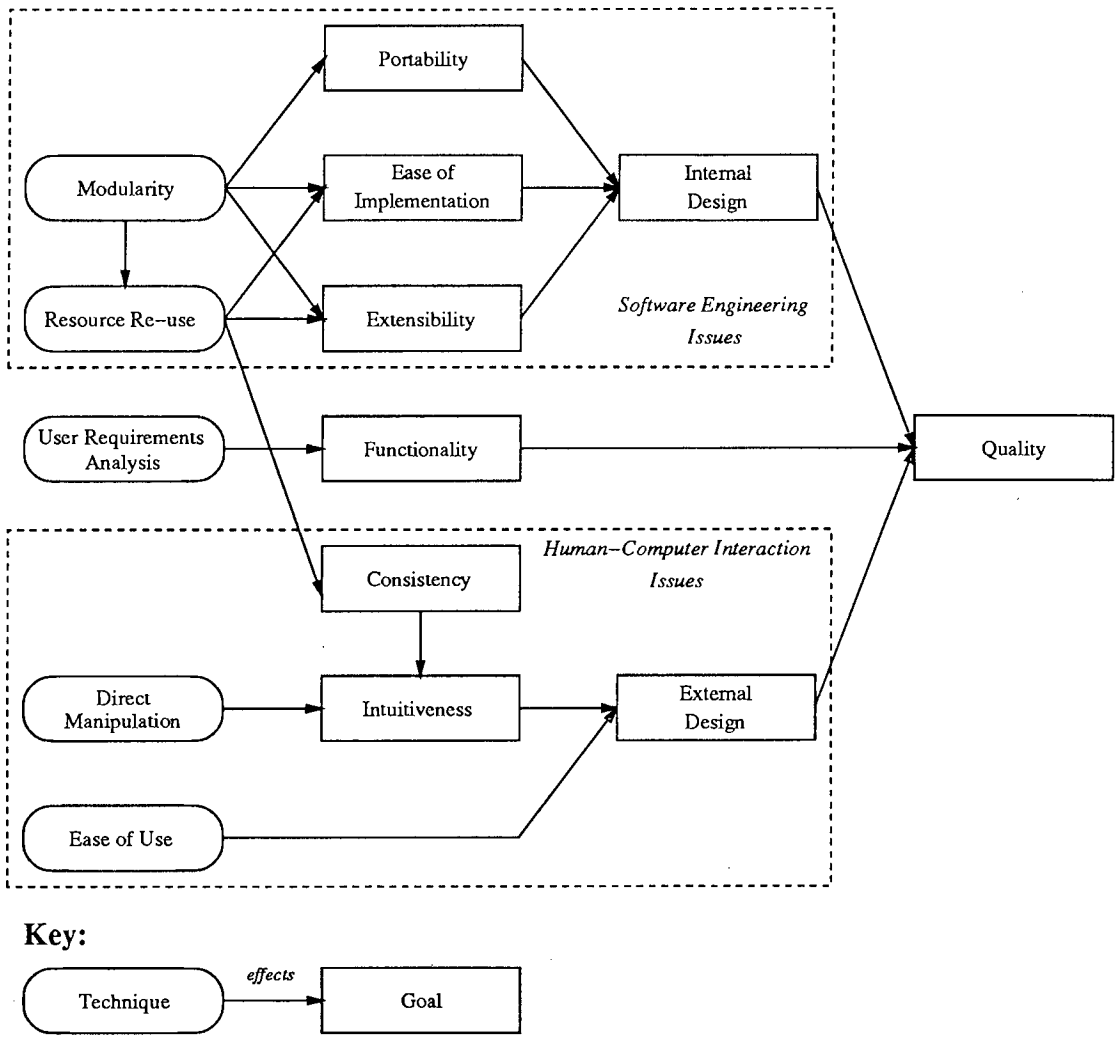


Figure 3–1: The interaction of the goals and associated techniques in both internal and external design considered relevant to this project.

Goal

o *Functionality*

The interface to a piece of software can be of the utmost importance, to the extent of affecting the degree to which a program is actually used. However, no matter how good its interface, a program with no underlying functionality will still not be used. It has been proposed [33] that a major limiting factor on the growth in the use of home computers is not that they have a particularly bad interface, but rather they have limited useful application.

> *"A primary consideration in the design of a building or of an interface to a computer system is that it* works, *that it fulfils the purpose for which it was intended. . . . No matter how beautiful a screen display is, an interface is ineffective unless the functionality of a system is revealed."*
>
> K. Hooper [34]

Contributors to functionality include:

User Requirements Analysis:

The designer can identify exactly which areas could benefit from software support by carrying out a user requirements analysis before implementing any tools. Armed with the information thus accrued, the designed may then produce tools which may be of real utility to the user community.

3.2.2 Software Engineering Issues for Effective Internal Design

The term *software engineering* was first introduced in the late 1960s [35]. It arose out of the software crisis that developed from then to the mid 1970s [36] because of the introduction of micro-chip technology. This produced computers which were orders of magnitude more powerful than those available previously. Applications which had been previously impracticable became feasible on available hardware, but their implementation required more sophisticated and much larger software systems. It was discovered that the building of larger software systems was not simply a case of scaling up existing software techniques, much as the industrial production of chemicals is generally considerably more involved than simply scaling up a laboratory synthesis procedure. The new larger applications required *engineering* rather than merely *building*.

To illustrate this distinction, consider the following analogy between *software engineering* and another form of engineering, *civil engineering*. In the early stages of society someone encountering the problem of crossing a narrow stream would perhaps look around for a fallen branch with which to build a bridge, or rocks to form stepping stones. The problem would then be solved. Nowadays much greater channels of water may be crossed. This problem is more complex and

the solutions have become more systematic and less *ad hoc*. The bridge must be properly designed by architects and engineers. The stresses and strains it may encounter must be modelled. The manner in which the bridge deals with them must be evaluated.

Similarly, a more formal approach is appropriate for the design and evaluation of software now that more complex programs are being designed and written. Programs must be written to stand the test of time, to be maintainable and extensible to cope with advances in technology.

Those techniques and goals considered to be software engineering factors in Figure 3–1 and their mutual interdependencies are discussed, in turn, below:

Techniques

- *Modularity*

Modularity is the decomposition of a program into logical units. Each unit contains code relevant to a specific function of the program — a classic example of "divide and conquer". For example in the molecular graphics tool to be described, the code relating to the geometric rotation of the molecule was implemented in one source file, the code for drawing atoms and bonds in another, *etc*. As detailed in later relevant sections, modularity aids implementation portability, extensibility and code re-use by adding logically distinct *tasks* in separately implemented *modules*.

- *Resource Re-use*

The re-use of existing software should be encouraged whenever possible. This reduces the amount of code which must be written, tested and documented [35]. Code may be designed to be reusable by offering routines which are self-contained, with well defined input and output. Data object specifications may be shared between programs to enable the sharing of data in addition to code.

Contributors to resource re-use include:

Modularity:

Splitting a program into logically self-contained modules with well-defined Input/Output (*I/O*) enables those modules to be used within other systems without interference to the original module and with minimum inconvenience to the designer.

The advantages provided by modularity to extensibility and consistency are detailed later in this section.

Goals

o *Portability*

The rate of change in computer hardware technology is such that the computing machinery often becomes obsolete long before the programs running thereon [35]. It is therefore important that programs should be constructed in such a manner that they may be implemented under more than one computer/operating system environment.

Contributors to portability include:

Modularity:

Portability is enhanced if a program is modular. Only those modules which contain the code incompatible with the new system need be updated. A program which runs on a direct manipulation interface may be ported to a system without the facilities required to run such an interface such as a bitmapped screen and pointing device. This simply requires the modules containing the interface code to be replaced with new modules appropriate to the available resources. Provided the data passed between the application code and the new interface remains unchanged, the function of the program will remain unchanged.

High-Level Languages:

Low-level languages are more hardware dependent than high-level languages. High-level languages are further abstracted from machine code. A program should therefore be implemented in a high-level language if designing for portability.

Self Containment:

A program should be self contained. Ideally, any libraries used should be widely available standards.

A completely portable item of software of any substantial size can still prove difficult to design despite these guidelines. Programs may remain machine dependent owing to the design of the machines themselves. The storage mechanism for floating point numbers may differ. Integers may be stored with the most significant bit rightmost or leftmost. Even the character set may vary from machine to machine. Programs may also depend on features of the operating system running on the host machine.

- *Ease of Implementation*

In this context, "ease of implementation" refers to the ease with which the program was developed, with which new features could be added, with which old ones may be improved and with which bugs may be fixed. Ease of implementation can be considered to be related to that of extensibility in that, in the former case the program is being "extended" from scratch instead of from some established level. Thus the advantages afforded by *modularity* and *resource re-use* to the ease of implementation are considered to be the same as for extensibility, and are therefore not discussed further here.

- *Extensibility*

The extensibility of software refers to the ease with which new functionality may be incorporated, in such forms as new features in a program or new programs in a package. Extensibility is particularly important in a research environment.

It is in such environments at the forefront of science and technology that the users tend to require new features to keep pace with current technology and discoveries.

Contributors to extensibility include:

Modularity:

The extensibility of a program is enhanced by modularity. The extension of a program will simply involve appending a new module to the original code together with the definition of the data to be passed between the new module and the original program. The addition of a new module should not interfere with any of the original functionality if the program is appropriately modularised. The program architecture should be headed by an easily extended controlling module which invokes all top-level modules.

Resource Re-use:

New programs may be added to a system without interfering with existing programs if data objects are clearly defined — in this case data files. Moreover, resource re-use should ensure that the data-structures and code for the storage and I/O of data objects are already available.

3.2.3 Human-Computer Interaction Issues for Effective External Design

Why should the human-computer interface be important? Taking this question literally, one could reply that without an interface the user could not communicate with the application. The application would then not be able to respond to the user, rendering the tool useless. A less extreme response would be that the quality of an interface can determine the extent to which the tool is used. An otherwise useful tool with a poorly designed interface could be rejected in favour of a less functional tool with a better or more familiar interface, especially by novice users. A "real-life" example is an article published in the proceedings

of a conference, which the author read during this research. The poor presentation of the article rendered it virtually useless — letters were smudged and unevenly spaced, words ran in to each other and were misspelt. It would take an exceptionally interested and dedicated reader to plough through to the end of this article.

Which specific goals are appropriate to the design of a new user interface? A major requirement of an effective interface is the provision of a means of communication with the application which allows the user to address problems in the task domain without having to cope with problems presented by the interface itself. The user should not have to learn about the computer and "computing techniques" [37]. It would be unreasonable to expect a chemist who wishes to use a computer to aid research to spend time becoming a computer expert, before being able to continue what would be considered to be the real work — chemistry research.

Ironically, the user will be unaware of an ideal user interface. An effective interaction design should virtually disappear from consciousness so that the user may concentrate on the work in hand with little conscious attention to the tools and maximal effectiveness at the intended work [19]. A poorly designed interface can provoke negative and undesirable feelings — such as panic, boredom, frustration, confusion or discomfort in the user. In contrast, a well designed interface promotes more desirable feelings of naturalness, interaction, ease of use and satisfaction.

> *"The promise of interactive graphics is to provide a user with a medium for communication with a computer which is at once benign, responsive and graphic."* J. D. Foley *et. al.* [19]

There has been much research into interface design. Alty *et. al.* identified attributes desirable in user interfaces by reviewing several articles with regard to constructing a set of guidelines for interface design [17]. Some of the guidelines are not relevant in this instance; others are covered in other chapters. Some, however, are applicable and are considered below.

Techniques

◦ *Direct Manipulation*

The key concept underlying a direct manipulation interface is that the user should interact with the computer by directly manipulating objects on the screen. This promotes a feeling of performing actions by physically manipulating the objects on the screen [36] and is achieved by the use of buttons, pull-down menus *etc.* rather than typing commands. Direct manipulation interfaces are often termed *WIMP* interfaces — Windows Icons Menus Pointing or Windows Icons Mouse Pull-down-menus. As stated above, a good interface should disappear from the consciousness of the user. The user should feel him or herself to be carrying out operations directly rather than instructing a machine to perform those operations, "cutting out the middle man". Direct manipulation interfaces claim to do just this. Hutchins *et. al.* [20] wrote a paper advocating direct manipulation interfaces. Much of the information in this section is taken from that paper. Hutchins states therein that the user should be unaware of the presence of any interface when using a tool with direct manipulation interfaces.

> *"To produce a feeling of direct engagement the system needs the interface to be unobtrusive, not interfering or intruding. If the interface itself is noticed, then it stands in a third person relationship to the objects of interest, and detracts from the directness of the engagement.*
>
> *The members of a well-staged play willfully suspend their beliefs that the players are actors and become directly engaged in the content of the drama. In a similar way, the user of a well-designed model world interface can willfully suspend belief that the objects depicted are artifacts of some program and can thereby directly engage the world of objects."*
>
> E. L. Hutchins *et. al.* [20]

The term "direct manipulation" was first coined by Shneiderman [21]. Such interfaces afford a continuous representation of the object of interest. Physical actions or labelled button presses are employed instead of complex syntax. Operations are rapid, incremental and reversible. The effect of operations on the object of interest is immediately visible [20].

Benefits of direct manipulation interfaces include:

- Several classes of syntax errors are eliminated. There are no hidden operations, no syntax or command names to learn. The user can not attempt to act on an object which does not exist or perform an operation which does not exist.

- Typing errors are almost eliminated since the user is rarely required to type.

- The user does have to remember obscure command names for infrequent operations since all operations are apparent to the user.

- The user requires expertise in the task domain, but only minimal knowledge of the computer or of computing.

◦ *Ease of Use*

Many features of interface design have been recommended in the literature, as listed and discussed below.

Menu Design:

A hierarchical menu structure is generally considered to be preferable to a single "flat" menu where there are many menu items. However, this may be unnecessary for fewer items [19]. Organisation can be alphabetical, logical, or by frequency of use .

Continuously Observable State:

Disorientation and confusion are reduced by making the current state of dialogue continuously observable to the user [17, 37–39].

Input Flexibility:

Maximum flexibility in a system's ability to handle input makes interaction with that system easier [17, 37, 40].

User Control:

The user should control the dialogue exchange rate. Computer operations should be a clear consequence of the user's actions [17,39,40].

Flexibility:

The system should be able to adapt to all levels of user expertise [17, 40]. The system should be designed in such a way that both novice and experienced users should be able to perform the task in hand with minimal inconvenience, whilst imposing the minimum of tedium on experienced users.

Familiarity:

Familiar icons or metaphors may be employed to prompt the user of the application. They convey their meaning more concisely and immediately than text and contribute to a feeling of naturalness [19]. Where icons are inappropriate, self-explanatory names should be adopted.

Goals

o *Consistency*

Consistency is an important consideration in interface design [17, 19, 20]. However, it has proved somewhat difficult to define consistency rigourously — or perhaps more significantly, to define its absence [41–44]. Young gave an informal description of the latter:

> *"One possibility is that a machine should be regarded as inconsistent if the user, exposed to certain aspects of its behaviour, is led to form expectations about other aspects which the machine then violates"* R. M. Young [45]

The advantages of consistency and the disadvantages of inconsistency are more clear. Consistency causes an interface to be more intuitive as detailed later in this section. Inconsistency may cause the user to become confused and frustrated. Interface consistency is recommended both by Shneiderman [46] in [36]

and Alty [17] in his review on HCI literature — being the second most frequent recommendation with ten citations [32, 38–40, 47] from the fifteen papers reviewed.

Factors contributing to consistency include:

Resource Re-use:

Code sharing promotes consistency between programs since the function of that code will be common to those programs. Any modification to that code will propagate to all relevant programs on recompilation. Similarly, re-use of I/O objects, data structures and widgets such as buttons and menus encourage the designer to maintain the same "look and feel" of a program when introducing a new tool to the tool-kit.

◦ *Intuitiveness*

As acknowledged above, a feature of an interface is more easily assimilated if it fits into a pattern which the user already recognises. The user may apply existing knowledge about the function of a previous application to a new application and be confident in its use more quickly. This proposition applies even if the "previous application" is not some other program but rather some aspect of the outside world. As Foley says "If the information fits into categories or concepts we already understand then the learning can proceed rapidly." [19] A command such as a button press or a menu choice can imply its function simply by its appearance.

Contributors to intuitiveness include:

Consistency:

An interface should not cause any surprise to the user [18]. A consistent system facilitates generalisations by the user, who may apply knowledge gained from one part of a system to infer the function of another [42]. An inconsistent system may confuse and frustrate the user if the knowledge gained from one part of the system causes an error or an unexpected result

when applied to another part. Phyllis Reisner gave an example at the HCI summer school at Heriot-Watt University in 1989. She described an early drawing tool comprising a drawing area and buttons [41]. This allowed the user to draw rectangles, circles and other graphic "primitives". On evaluation it was discovered that people were unable to use the text facility since there was no button for that purpose. In fact, text could be generated simply by typing. However, this was inconsistent with the other functions. The users were not able to generalise knowledge accrued from drawing other graphics primitives, resulting in an inability to use the text function.

Direct Manipulation:

A direct manipulation interface increases intuitiveness since no complicated syntax need be learned. All available commands are at hand. Icons may illustrate the function behind a command.

3.3 Objectives

The user-driven requirements discussed in the previous chapter are translated in this section into design objectives. In addition, some objectives have been included which were not explicitly requested in the user requirements analysis. Some of these relate to human computer interaction issues, but mostly they refer to software engineering issues, which unsurprisingly did not emerge in the questionnaire responses. Software engineering issues are important commercially in ensuring that software is cheap to maintain once issued. Although the tool-kit is not being designed as part of a commercial project, it is intended to be made available to the research community, where people are likely to want to adapt it to their own use and/or pass it on to others. For these reasons, software engineering issues are felt to be just as important to this tool-kit as to a commercial package. The objectives thought to be relevant to the tool-kit as a whole are detailed below. However, the evaluation of the tools with regard to

these objectives is postponed until Chapter 4 since no implementation has yet been discussed.

1. *Direct Manipulation:*

   That the programs, where appropriate, have graphical, direct manipulation interfaces; that the user can see the immediate effect of any commands executed; that all commands are readily available *via* menus, buttons and text input windows; that complicated syntax need not be remembered.

2. *Intuitiveness:*

   That the commands within programs such as buttons or menu choices give the user an indication of their purpose; that a command which appears likely to initiate one procedure does not, in fact, initiate another; that the programs do not cause the user any unwanted surprises.

3. *Consistency:*

   (a) *Resource Re-use:*

   That code, data-structures and data definitions may be shared between programs. This has the added benefit that users need not edit files between programs and that programs have common features.

   (b) *Look and Feel:*

   That tools will have the same *look and feel* and, where appropriate the same interface; that tools are both mutually and internally consistent; that in each program similar commands initiate similar procedures.

4. *Integration of Programs:*

   That programs should be fully integrated. That it should be possible to use the output object from one program as the input object of another, without the need for manual editing. That programs exploit each other where appropriate. That it is not necessary, for example, to terminate a theoretical or analytical tool before the results may be viewed with a visualisation tool.

5. *Input Flexibility:*

   That input objects may come in a variety of formats as long as the necessary information is included.

6. *Expertise Flexibility:*

   The program should be usable by novices, whilst not causing frustration to the more experienced.

7. *Structure:*

   That different forms of data are distinct, *i.e.* that the information necessary to perform an experiment is not expected to be held in the same object as the model that experiment is to be performed on.

8. *Defaults:*

   That it is not necessary for the user to enter established *facts* such as atomic radii or number of valence electrons as data unless those held internally are to be overridden; that the required data be kept to a minimum.

9. *Appropriate Presentation of Data:*

   That results are presented in such a way as to *convey* the maximum amount of information.

10. *Portability:*

    That the program may be ported from one system to another with the minimum of adaptation.

11. *Extensibility:*

    That other programs may be added to the tool-kit or features added to the individual programs in the tool-kit. This should incur the minimum disruption to the existing tools and the minimum effort from the person amending the software.

3.4 The Tool-Kit

As explained in the previous chapter two tools were chosen to illustrate the design and implementation of the tool-kit. These were named *the Visualisor* — a molecular graphics tool — and *the Locator* — a tool for locating hydride ligands in transition metal clusters. During the implementation it became necessary to implement a third tool — *the Orthogonalisor* — to convert models between coordinate systems.

Figure 3–2 is a diagrammatic representation of the tool-kit as it stands. This shows the relationships between the tools, the types of molecular models and other data objects each may process. The data which constitutes a molecular model is discussed in more detail on page 85. This diagram shows the tool-kit contained in the lightly shaded area. Only those data transactions which cross or occur outside this area require any user interaction. Those completely inside this area are performed automatically by the appropriate tool.

The implemented tools directly provide support for two of the activities included in the user model in Figure 2–3: the Visualisor is a tool to aid *"Interpretation"* and the Locator is a theoretical modelling tool; *"Theory"*. The Visualisor may provide indirect support for most of the other activities. It has been identified earlier that *"Synthetic Chemistry"* leaves little room for software support. However, the Visualisor could provide support for analytical chemistry, the documentation of results, literature surveying, and other facets of theoretical chemistry. The Visualisor could provide an interface to any existing or future analytical or theoretical tools which involve molecular models. For example, the crystal structure solution program, SHELX, could benefit greatly from such an interface. Instead of the current method of running SHELX, which is considered to be crude, laborious and error-prone as discussed in Chapter 2, page 42, the user could be presented with a three-dimensional manipulatable representation of the molecule. The atoms to be retained or rejected could be selected with a

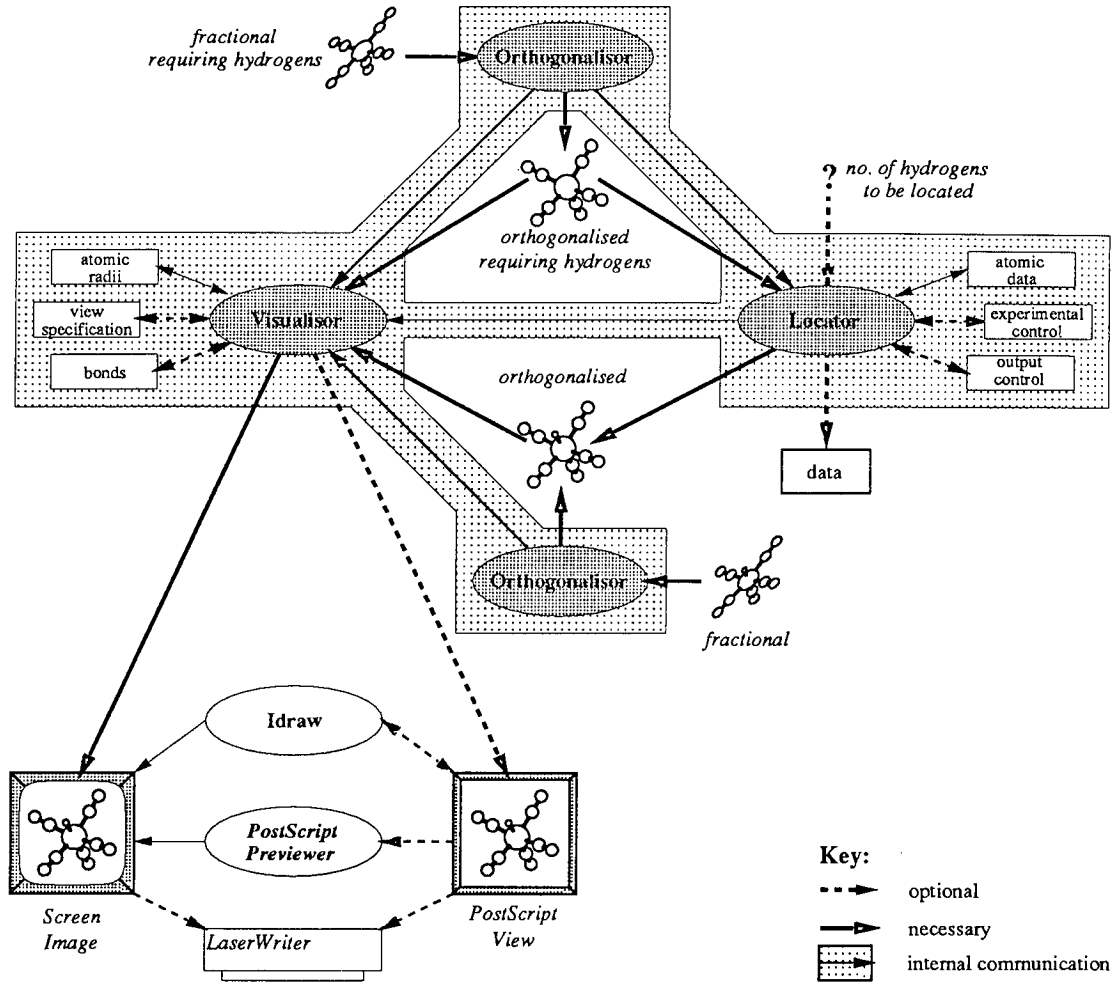


Figure 3–2: The interaction of the tools within the toolkit and with the outside environment.

pointing device and a button depressed to initiate the next iteration. Similarly databases may be interactively browsed and molecular models extracted, and views of models saved in PostScript form for inclusion in documents.

3.4.1 The Orthogonalisor

Molecular models tend to be published in *fractional* coordinates as opposed to *orthogonalised Ångström* coordinates. Fractional coordinates refer to the position of the atoms within the molecule with respect to the *unit cell*. The unit cell is a parallelepiped region of space which encloses exactly one "molecular unit"

within the crystal structure. Depending on the molecular symmetry of the crystal, the edges of the unit cell may not necessarily have equal dimensions in the x, y and z directions and need not be mutually orthogonal. The parameters which describe the shape of the unit cell are known as the *cell parameters*. Orthogonalised Ångström coordinates, as the name would suggest, refer to a standardised unit cell with orthogonal unit length edges.

Each form of representation has its particular merits and drawbacks. Certain symmetry operations may be carried out more readily on the model when it is described by fractional coordinates, such as regeneration of a dimer when only the monomer is available. Orthogonalised Ångström coordinates may be more appropriate for the substitution and generation of functional groups.

Thus the Orthogonalisor was written to convert molecular models from fractional coordinates to orthogonalised Ångström coordinates. As a tool in its own right, the Orthogonalisor plays only a minor rôle in the tool-kit. It was used occasionally to convert model coordinates into orthogonalised Ångström space so that small functional groups could be substituted for larger ones when so required by the Locator. The Locator could only deal with a limited number of atoms. This was for technical reasons since the program was only allocated limited storage space, but perhaps more significantly also for pragmatic reasons because execution time increased with the square of the atom count in a molecule — $O(n^2)$ in the usual computer science notation.

The importance of the Orthogonalisor lies not in that it is used as a tool, but rather in that its underlying code is shared by the two larger tools in the environment — the Locator and the Visualisor. Each of these tools requires the model to be in orthogonal Ångström coordinates. On recognising a model in fractional coordinates with non unit cube cell parameters[1] — either tool will invoke this code to orthogonalise the model before proceeding.

---

[1] that is, parameters other than unit side lengths, a=b=c=1, and right angles, α=β=γ=90°

3.4.2 The Visualisor

The Visualisor may accept an orthogonalised model directly or a model in fractional coordinates *via* the Orthogonalisor, as Figure 3–2 shows. These routes will appear identical from the user's point of view. In either case a model is loaded and a picture of that model will be displayed. It is not even necessary for the user to know the type of model since, as described above, it is automatically recognised.

The molecular model is the only data object required by the Visualisor from the user. All other objects — denoted *bonds, view specification* and *atomic radii* in the diagram — are provided by the tool-kit.

The user may select any view of the molecule for output in PostScript, a device independent standard for representing the printed page. This may be sent directly to a PostScript printer or browsed in real time with preview tools such as *GhostScript*. Alternatively, the output may be edited for purposes such as annotation with the *Idraw* InterViews tool. The Visualisor's PostScript output is specifically designed to be compatible with that of Idraw, allowing any PostScript pictures to be shaded and annotated, as the user requires, in a post-processing stage. The Visualisor is described in more detail in Chapter 4.

The underlying code of the Visualisor is itself shared with the Locator, providing a consistent direct manipulation interface between the two tools.

3.4.3 The Locator

The location of metal bound hydride ligands in transition metal clusters by conventional analytical techniques is difficult and sometimes impracticable. This tool seeks to locate hydrogen numerically in such compounds by employing a Simplex routine to optimise the calculated potential energy of the model.

It can be seen from Figure 3–2 that the Locator requires the same data as the other tools — the molecular model. The number of hydride ligands to be located is

required in addition if more than the default of one. As already mentioned, the Locator uses the underlying code of both the other tools — the Orthogonalisor to convert models from fractional coordinates and the Visualisor to provide an interface and a visualisation mechanism for the model to display the most recently calculated hydride positions. The use of the Visualisor as an interface provides consistency between the tools. Moreover, the Locator inherits all the facilities such as molecular rotation and the ability to output a view in PostScript. The problem of hydrogen location and the solution adopted by the Locator are discussed in more detail in Chapters 5 and 6.

3.5 Implementation

Established software engineering and human-computer interaction issues for the implementation of software were discussed above, including the mutual effects that one may have on another as shown in Figure 3–1. This section demonstrates how these issues are applied to the implementation of the tool set.

3.5.1 General Issues

User Requirements Analysis, Functionality

The implementation of the user requirements analysis and its contribution to the identification of appropriate functionality were discussed in the previous chapter.

3.5.2 Software Engineering Issues

Modularity

The program has a modular design. All code which is connected to one particular function such as the the rotation of the molecule, or the identification of

molecular bonds is collected together in a self-contained file with a well defined procedural interface. Wherever possible variables and functions are *local, i.e.* they are invisible from other files and therefore not accessible from inappropriate routines.

Resource Re-use

◦ *Common Code*

The consistency between the programs in consideration of the HCI issues provided ample opportunity for code to be shared in consideration of software engineering issues. Each program accepts the same format of molecular model, so the routines which read and interpret that model could be shared between programs. Other shared routines read in atomic data from the atomic data file, determine which atoms are bonded to each other and give the user warning or error messages.

Sharing code between the programs wherever possible had various ramifications. Providing a consistent interface for the user (see HCI issues) was somewhat easier. Whenever a routine contributing to the interface was modified, such as the error notification routine, this change was propagated through all programs. The programmer did not have to worry about updating the equivalent routines in each program, or whether updating one program left another behind.

◦ *Common Data Structure*

Each program deals with the entity of a molecular model. This has certain attributes such as a specified atom count, charge *etc.* It was therefore considered appropriate to have a common data structure to store the information about the model. Figure 3–3 shows a representation of that data-structure. It can be seen that a structure called *molecule* contains entities to store information about the molecule as a whole. These include a structure called *cell* for storing the unit cell information and structures called *atoms* for storing atomic information.

Where necessary, the information for filling this structure comes partially from the molecular model, partially from the common data file and partially from the calculations done in the respective programs.

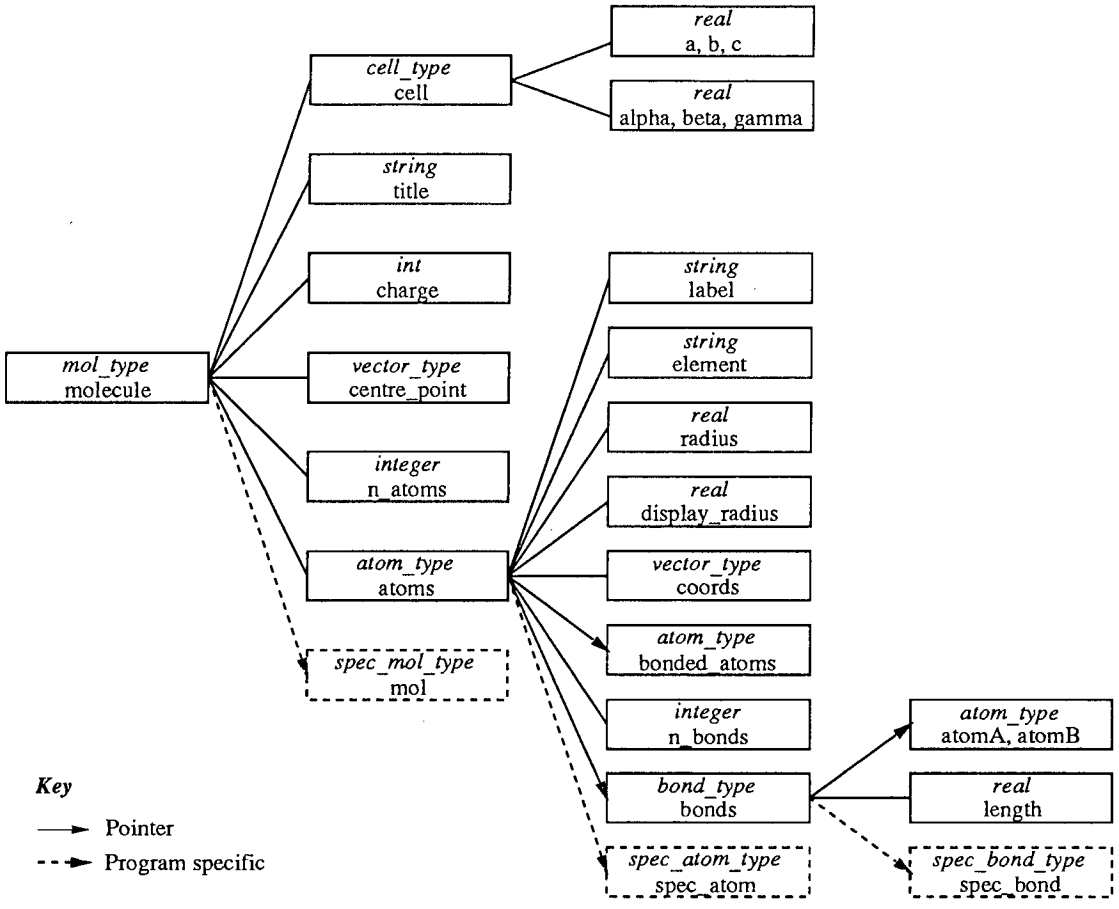


Figure 3–3: The molecular data structure.

During implementation it was observed that certain pieces of information, which logically should be stored as part of the molecular data structure, were particular to the program in question. Examples of such data are the screen coordinates and the display radius of the atoms for the molecular graphics program, and the valence electron count of each atom for the hydrogen location program. This issue for the sharing of data-structures was addressed by leaving spaces in the data-structure to accommodate information particular to the program. Referring to Figure 3–3 it can be seen that within the structure *molecule*

there is a pointer to a structure *spec_mol* (specific molecule). Similarly within the structure *atom* there is a pointer to a structure *spec_atom* (specific atom). These structures may be appropriately specified by a particular program to store local information.

Portability

In the interest of portability an effort was made to design the tool-kit to use only software and hardware which are common in industry and academe and as such are "de-facto" standards.

- *Hardware*

The minimal hardware requirement necessary to achieve the direct manipulation interface objective (see objective 1, section 3.3) was a machine with a bitmapped screen and a pointing device. Such machines are becoming the standard in all areas of scientific application. Thus it was decided that any software should be designed to run on such machines. The only other requirement placed on hardware was that the machine should be capable of supporting the software detailed in the rest of this section.

- *Operating System*

Unix is a standard for the intermediate sized machines such as the workstations mentioned previously.

- *Windowing System*

The most common windowing system is the X Window system [48–53]. Several tool-kits are available to program on the X Window system. Of these, *Xlib* [54, 55], *Xt (the X Toolkit Intrinsics)* [56–59] and the *Athena Widget Set* [58–60] were readily available at the time of beginning this project and fully supported locally. All three are provided the X Window system system software distribution by MIT. There should therefore be no problems with potential users having to

obtain extra software before being able to compile or link programs using these packages.

- *Language*

The programs were largely written in C. The single exception to this was those parts of the hydrogen location tool which were developed from an existing Fortran implementation.

- *Libraries*

In the Fortran part of the hydrogen location tool it was necessary to use routines from the standard Nag library which is widely available.

Ease of Implementation

The modular design of the implementation enabled the designer to address one problem and concentrate on just one part of the program at a time, with no need to consider implications on or from any other. New modules could be introduced without worrying about adverse effects on others, and could be fully implemented and tested before moving on to the next.

Extensibility

Much as for the ease of implementation, the extensibility of a program is aided by a modular design. If future users/designers wish to add new features such as the ability to add and delete functional groups and atoms, the module could be written independently. It is acknowledged that total independence is not feasible — there would have to be interaction with some other parts of the main program such as entries in the main structure or the use of the control panel. However, apart from the necessary data transfer, modules may remain independent and need not adversely effect any of the rest of the program. The user/designer may wish to adopt a different input format, for example,

extracting models directly from a structural database. The implementation of this would be a simple case of replacing the module which reads in the molecular data structure with an equivalent accessing the database.

3.5.3 HCI Issues

Direct Manipulation, User Control, Continuously Observable State

Both the Locator and the Visualisor are operated by buttons, pull-down menus and scroll-bars. The keyboard need only be used on occasions such as editing prospective file names for PostScript files or specifying printers. Such methods of interaction give the user full control over the actions of the program. Furthermore, on depressing a button or selecting a menu item the effect of that action is immediately viewable. For example, on selecting "rotate" from the menu a rotation panel immediately appears. On depressing a displayed rotation button the user will see the view of the molecule rotate. Each user action has a corresponding computer reaction.

Ease of Use

o *Menu Design*

Commands are collected together under logical headings in the menu structure (see Figure 4–2). Thus users may quickly locate appropriate commands.

o *Input Flexibility:*

HCI literature [17,37,40] recommends that the user should be allowed as much input flexibility as is possible. The user may face some disadvantages if he or she is restricted to a certain format:

- The user must know or find out the accepted format before being able to continue using the program.

- There can be a greater margin for error if only one format may be accepted.

- If the data comes directly from another program as opposed to being typed by the user, the format of this data obtained may not concur with that for the program to be used.

Before the flexibility of the input procedure can be discussed, the data required by the respective tools must be detailed. The user object required as input for each tool is the *molecular model* as Figure 3–2 shows. This comprises the data describing the molecule, which is to be visualised, orthogonalised, or have hydride ligands located within it. For the purposes of this project a molecular model consists of the following:

- For all tools:

  Atomic Coordinates

  For each atom, coordinates identifying the position of the atom and text identifying the atom type are required.

  Cell Parameters

  If the data is presented in fractional coordinates, cell parameters are required for the interpretation of these atom coordinates.

- For the Locator:

  Charge

  If the molecule has a charge, *i.e.* it is an ion, this charge must be specified to the Locator since the electron count is essential to the calculations therein.

- Optional:

  Title

  The title is purely for the user's convenience. Any title given will be displayed along the top of the display, above the menu bar as shown in Figure 4–1. The name of the molecular model file is also shown in this panel.

It appeared that the most appropriate manner to promote flexibility in allowing files produced by other programs to be valid input files was to adopt *keyword identifiers*. The reasons behind this were twofold. Firstly, the use of keywords would enable the respective program to extract the relevant data whilst ignoring the rest. Secondly, there are several existing tools in the Chemistry Department (*e.g.* CCCP, SHELX) which use a similar system and some of the same keywords. Thus, provided the required data is present in the input file object — in whatever order — with appropriate key identifiers, the input file will be valid. Furthermore, the use of keywords allows user comments to be placed anywhere in the file, perhaps for identifying functional groups or unusual atom positions. Figure 3–4 shows an annotated example molecular model object.

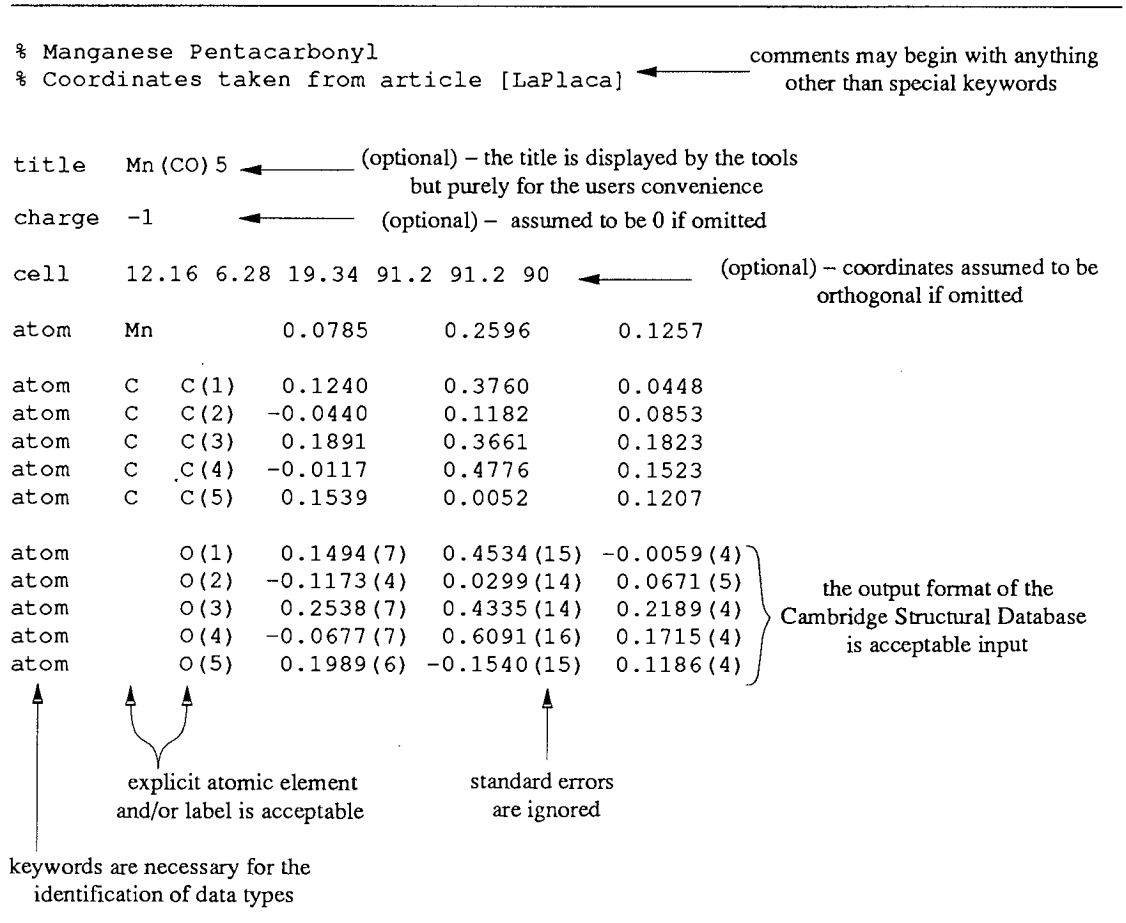


Figure 3–4: An example molecular model.

Regarding the data within the file, the user may identify the atom by one of three methods, illustrated by the different blocks of Figure 3–4. Referring back to Figure 3–3, it can be seen that atom types have both element labels and atom labels. The first identifies the type of atom such as *H, C, Mn* whilst the second is a unique label to distinguish between atoms of the same type. Either or both of these labels may be specified. Where only one is specified the other is inferred from that provided. The cell data is optional, but if omitted the model is assumed to be in orthogonalised Ångström coordinates. In some crystallographic programs a seventh parameter is given in the cell record. This is the wavelength of the incident X-radiation for the X-ray diffraction. If included, this parameter is recognised by the file reading routine and simply ignored. The file format available from the Cambridge Structural Database[2] may be read in directly to the tool-kit programs after the addition of appropriate keywords.

This input flexibility is in sharp contrast to the style of input files briefly discussed in Chapter 1, where rigourous format conventions are enforced.

◦ *Flexibility:*

The benefits of direct manipulation interfaces to novice users have already been discussed. However, such interfaces are not entirely unsuitable for experienced users either. The main risk with experienced users using programs designed for inexperienced users is boredom and frustration. However, this is not thought to be a great problem with the tools in the tool-kit. Constant evaluation by the designer has identified some features causing such negative reactions in experienced users — here the designer is in the rôle of an experienced user. To address this issue, a set of command line arguments have been made available to the experienced user. These flags include:

---

[2] the Cambridge Structural Database is also known as the Cambridge *Crystallographic* Database

`-f`	*file name*:	Molecular model to be loaded.
`-H`	*number*:	Number of hydride ligands to be located.
`-o`	*optimisation*:	Optimisation routine to be applied (Simplex or Newtonian).
`-p`	*paper width*:	Paper format for the output of matrices (80 or 120, narrow or wide).
`-r`		Restrict the hydride position to within a certain distance of all atoms in the molecule.
`-sp`	*type*:	Starting point method to be used (Projection or Biggest Appropriate Hole) — these are discussed in Chapter 6.
`-spf`	*filename*:	File name containing user starting coordinates.

- *Familiarity:*

Wherever possible familiar icons have been adopted to prompt the user. The buttons to rotate the molecule are labelled with circular arrows, and examples of the shading techniques are given — see Figure 4–5.

Consistency

The use of the X Windows tool-kit and Athena Widget set promoted consistency across the interface of each tool and between tools. For example, every time a button was required for some part of the interface an instance of the Athena command widget was created. The user need only recognise a single button form and infer its action. Similarly the user should quickly become familiar with the consistent use of scroll-bars, text input windows and menus. A panel is provided down the side of the visualisation tool — and therefore the hydrogen location tool since they share an interface — for any user commands which require further user interaction. Each command has the same *done* and *cancel* buttons, provided by a single piece of code.

As mentioned earlier and described in greater detail later, the molecular visualisation tool provides an interface for the hydrogen location tool, thereby maintaining even greater consistency.

Intuitiveness

The internal and external consistency of the program, and the use of a direct manipulation interface, contribute to the intuitiveness of the tools. Consistency contributes to intuitiveness by providing analogies. Once a user has learned how to use one command, that knowledge can be applied to other commands. The use of a direct manipulation interface ensures that all the commands are to hand, displayed through buttons, menus and text input windows. This eliminates the need for the user to remember the syntax of complicated commands.

Visual aids are used wherever possible. For instance, the rotation of the molecules is performed *via* widgets displaying circular arrows pointing in the direction of movement. The help facility is denoted by a question mark. The menu entries are collected together under logical headings.

Chapter 4

The Visualisor

4.1 Introduction

The study of molecules is perhaps the most important part of chemistry; indeed, chemistry could be considered to be the science of molecules, with regard to their structure and interactions. In the course of chemistry molecules are synthesised and fragmented, and the atoms contained therein, together with their intramolecular relationships, are investigated. The internal structure of a molecule cannot itself be viewed. However, sufficient information about the spatial relationships between atoms within the molecule may be gleaned from various analytical techniques to build a theoretical picture of the molecule, resulting in, amongst other formulations, a molecular model. In itself, this model constitutes molecular data. It must be interpreted to convert the raw data into useful information. Full interpretation may be promoted by visualisation.

Scientific visualisation and molecular graphics literature is reviewed in the remainder of this chapter. A list of objectives for the implementation of a molecular visualisation tool — *the Visualisor* — is presented in addition to those detailed in the previous chapter. The implementation of the Visualisor is discussed in terms of its user interface and the facilities presented. Finally, the Visualisor is evaluated in terms of the degree to which the objectives detailed in this and

the previous chapter have been achieved. Conclusions are drawn regarding the usefulness of the tool both in its own right and as an interface for other tools.

4.2 Literature Review

4.2.1 Scientific Visualisation

The term *scientific visualisation* was first introduced by McCormick *et. al.* [61,62]. Scientific visualisation is a term describing the enhancement of the communication and the interpretation of scientific data through visual methods. The concept of visualisation is not new — graphs, diagrams and charts have long aided interpretation and communication [63,64]. Scientific visualisation has been exploited since the advent of science. It is not restricted to the use of computers; hand drawn graphs and diagrams may aid interpretation quite adequately. However, graphical aids to visualisation can be generated more quickly and more reliably by computer than manually.

The amount of data produced by scientific software is often huge [62]. Much scientific data is still presented numerically, only a limited volume of which may be readily assimilated at a time. It can be difficult to identify both global trends and local critical points in large amounts of numerical data. The ICON8 tool discussed briefly in Chapter 2 of this thesis is an example of such a program. ICON8 may produce several large matrices as part of its output, describing such properties as the overlap between atomic orbitals. Since these matrices may constitute up to $200 \times 200 = 40000$ data items, interpreting anything other than a small part of these data is practically impossible without resorting to graphical methods. Data presented textually fails to convey all but a small part of its meaning. When data is presented graphically however, the user may employ other parts of the brain in the assimilation of information and for a more effective interpretation. As Hamming [65] said "The purpose of computing is insight, not numbers". An estimated 50% of the brain's neurons are associated

with vision [61]. This extra neurological machinery may be put to work to make more effective use of human capabilities. Scientific visualisation aims to do just that. Those mechanisms in humans which allow them to perceive, use and communicate information are studied to make maximum use of human interpretation capabilities in the field of scientific visualisation.

Visualisation may be of benefit to the user at three different stages [64]:

- *post-processing* — visualisation of the results from a program after the program has terminated,

- *tracking* — visualisation of the results from a program while the program is running and

- *steering* — similar to tracking, but with the user able to interact with the program, entering new data mid-stream.

Much research effort has addressed post-processing, although some programs do exist which allow tracking or steering. Of the latter, tracking is perhaps the more common; the program described in Chapter 6 of this thesis employs visualisation by tracking. An important benefit of scientific visualisation is the insight gained by spotting anomalies visually [61].

The representation of three or more dimensional data on a two dimensional surface (screen or paper) necessarily reduces the information content [66]. However there are many visualisation "techniques" which may be employed to minimise the impact of this reduction. Simple techniques such as use of colour, plotting of subsets of data (slicing) [61] and wire-frame or shaded pictures with or without perspective may all give a greater impression of three dimensions. Other, more advanced techniques include stereopsis — two images, each slightly rotated with respect to the other and presented independently to each eye to give an impression of a truly three dimensional object — cine sequences, head-mounted displays and varifocal mirrors — mirrors which provide a "true"

three-dimensional perception of a distribution of glowing points of light. Visualisation may be enhanced if the viewing position can vary in real time.

Scientific visualisation finds many applications throughout science [64, 66–75]. Interactive computer graphics have provided insight into chemical complexity for some time, generally to support the activities of synthesis planning, analysis and communication [61].

4.2.2 Molecular Graphics

Many molecular graphics packages are available to the practising chemist. These range from the large non-interactive Fortran programs widely used in university chemistry departments to more recent programs which offer interactive facilities and produce beautiful colour ray-traced shaded images of molecules. Several of these packages are reviewed in this section.

Connolly and Olson [14]

GRANNY uses the language GRAMPS[1], a graphics language designed to enable development of real-time interactive computer graphics applications. Portability of the former depends on the availability of the latter. GRANNY is interactive, but does not have a direct manipulation interface. The user is required to type commands interactively and may need to communicate directly with GRAMPS for tasks unanticipated by GRANNY. The command syntax is "English-like" but rigid and sometimes complex; atoms are specified by molecule name, segment name, residue number, residue name and atom name in that order. Molecular models may be represented by solvent accessible surfaces; however, this surface must be preprocessed by another program. Similarly the input files for bonds and chain (virtual bonds) are generated by a preliminary program. Bond rotation is performed with a dial. Six colours are available.

---

[1]GRAMPS is discussed later in this review

Dayringer et. al. [76]

PROTEUS is a graphics system for proteins. The program is written in C and runs on a VAX and an Evans and Sutherland PS300. PROTEUS facilities include the ability to rotate, translate and scale objects, to mutate, insert and remove residues and change the distance, angle and dihedral angles between atoms and molecules. A collection of objects may be treated as a single object. Addition and removal of bonds, stereopsis and parameter measurement are also possible. Furthermore, "best matches" may be searched for on a database and there are no restrictions on the atom numbers. Regarding the interface, the user has a choice of an (easier) menu system on the PS300 or a (more powerful) command-line interface directly on the VAX.

Diamond [77]

Bilder is designed to run on a PDP11/50 in conjunction with an Evans and Sutherland Picture System 1. Bilder occupies the whole of this machine, necessitating several software overlays or "swaps" during its operation. Bilder must interact with a more powerful IBM 370/165 for more numerically intensive mathematical operations. Interaction is *via* pages of menus offering addition and removal of residues, bond rotation and chain severing.

Hubbard [78]

HYDRA (*H*arvard *Y*ork *Dra*wing program) runs on the Evans and Sutherland PS300. A binary data format is adopted for rapid data handling. This enables surface accessibility, immunogenecity or crystallographic thermal parameters to be represented effectively and easily. Hydrogen bonds are calculated, bond rotation is possible and manipulation sequences may be stored as movies.

Jones [15]

FRODO runs on many machines and on many different graphics peripherals. Interaction is through a menu system. However, the order in which commands are activated is important; YES, SAVE has an entirely different effect to SAVE, YES. The molecular data set must follow a particular sequence.

O'Donnell and Olson [16]

GRAMPS is a graphics language interpreter. It runs on the Evans and Sutherland Multi-Picture-System, an interactive vector display list processor. GRAMPS was written in Fortran 77 with the aim of providing a graphical standard for the multitude of differing chemical applications. Until that point, such applications had tended to be written *ad hoc* with a single graphics task in mind. There are about 50 commands available, which roughly fall into the three categories of picture creation, picture transformation and system utility commands. The syntax of a GRAMPS command is of the form *verb keywords arguments*, although joysticks and dials and a tablet may be used for interaction. GRAMPS saves transformational values instead of actual vector positions for speed and memory efficiency.

Sayle [79]

A novel algorithm for rendering molecular representations on a parallel computer or distributed system is presented. Current thinking on accelerated ray-tracing methods is reviewed with respect to the degree of parallelism in each algorithm and the speedup achieved utilising several processors concurrently. An algorithm is presented for rendering molecular representations on a parallel computer or distributed system, together with a demonstration of the implemented system. The results of this experiment are compared with the existing methods in the review.

Sayle and Bissell [80]

This paper details the development of an interactive program, RasMol, for the visualisation of proteins and nucleic acids. The rôle of shadows in creating a perception of depth, particularly when moving across a surface under rotation, is identified. Current techniques for displaying the three-dimensional structures of molecules and their methods for the determination of their cast shadows are reviewed. An efficient hybrid ray-tracing algorithm for molecular graphics based upon a uniform spatial subdivision acceleration scheme is described. Results are presented for the implementation of this algorithm on both Transputer based multiprocessors and UNIX workstations under the the X Window system. Both versions are claimed to have the fastest rendering times for shadowed union-of-spheres surfaces on multi-processor and single processor architectures respectively published to date.

Stewart [81]

Seven molecular graphics packages are reviewed, these being DISCOVER, INSIGHT, GRAMPS, GRANNY, MOGLI, SYBYL and FRODO. The packages are compared in tabular form with respect to several different attributes, including usability, host hardware and required environment.

The January 1985 issue of *Chemistry in Britain* featured several articles relating to then state-of-the-art molecular graphics. These include applications to medicinal chemistry [82], drug design [83] and surface chemistry [84]. Furthermore, a comprehensive history and review of molecular graphics were presented at the SIGGRAPH 1992 conference on Computer Graphics and Interactive Techniques [85]

4.3 Objectives

All the requirements discussed in Chapter 3 apply to the Visualisor. In addition to these, the following objectives are relevant to molecular graphics:

12. *Default Representation*

    The default display should be of a ball-and-stick representation with hidden surface removal.

13. *User Controlled Atom and Bond Sizes*

    The user should be able to alter the sizes of atoms and bonds either dependently together or independently. In the extremes of atom and bond size this would effectively give the user a choice of different representations of display — line-drawing with bonds only, space filling with atoms only and ball-and-stick with both.

14. *Rotation*

    It should be possible to investigate spatial relationships with the molecule by interactively rotating the image in three dimensions, in whichever style or representation is chosen.

15. *Zooming*

    The user should be able to zoom in on any part of the image for a more detailed view.

16. *View Saving*

    The user should be able to save a desired view of the molecule. This should be loaded automatically whenever the corresponding molecular model is loaded.

17. *Graphic Output*

    It should be possible to output a view of the molecule to a file. This should

be in a standard image format which may be loaded into other programs for subsequent viewing, editing or incorporation into text formatters or word-processors.

18. *Printing*

    The user should be able to print a current view of the molecular representation directly on an appropriate printer.

19. *Consistency*

    The graphic output and direct printing above should directly mirror the view of the molecule on the screen.

4.4 Implementation

The program originated from an *Interactive Molecular Graphics Package* [86] written by Irvine as a final year project at the University of Edinburgh. Irvine's program was implemented on the SunView window system and displayed molecules with hidden surfaces removed by using the "painters' " or "list of priority" algorithm [87]. The initial intention was simply to port the program to the X Window system since SunView was no longer to be supported by the manufacturer. The port should have allowed the molecular graphics program to become part of, and be consistent with, the rest of the tool-kit. Once the procedure was started, however, it was soon discovered that both the interface, and the underlying application could be developed further with the Xt and Xlib facilities and the Athena widget set. The porting exercise rapidly became a total re-write in which some of the ideas from the original program were adopted but the code was not. The resulting tool was subsequently named "The Visualisor". An earlier version of this tool was presented at *The European X Window System Conference* in London, November 1990 and the published article is reproduced in Appendix G of this thesis.

In accordance with the design of the tool-kit detailed in the previous chapter, the program was written to run on the X Window system using Xlib, Xt and the Athena widget set. Thus the program is driven by direct manipulation *via* menus, buttons, toggles and scroll bars. Figure 4–1 shows an annotated screen dump of the Visualisor illustrating the constituent parts of the display.

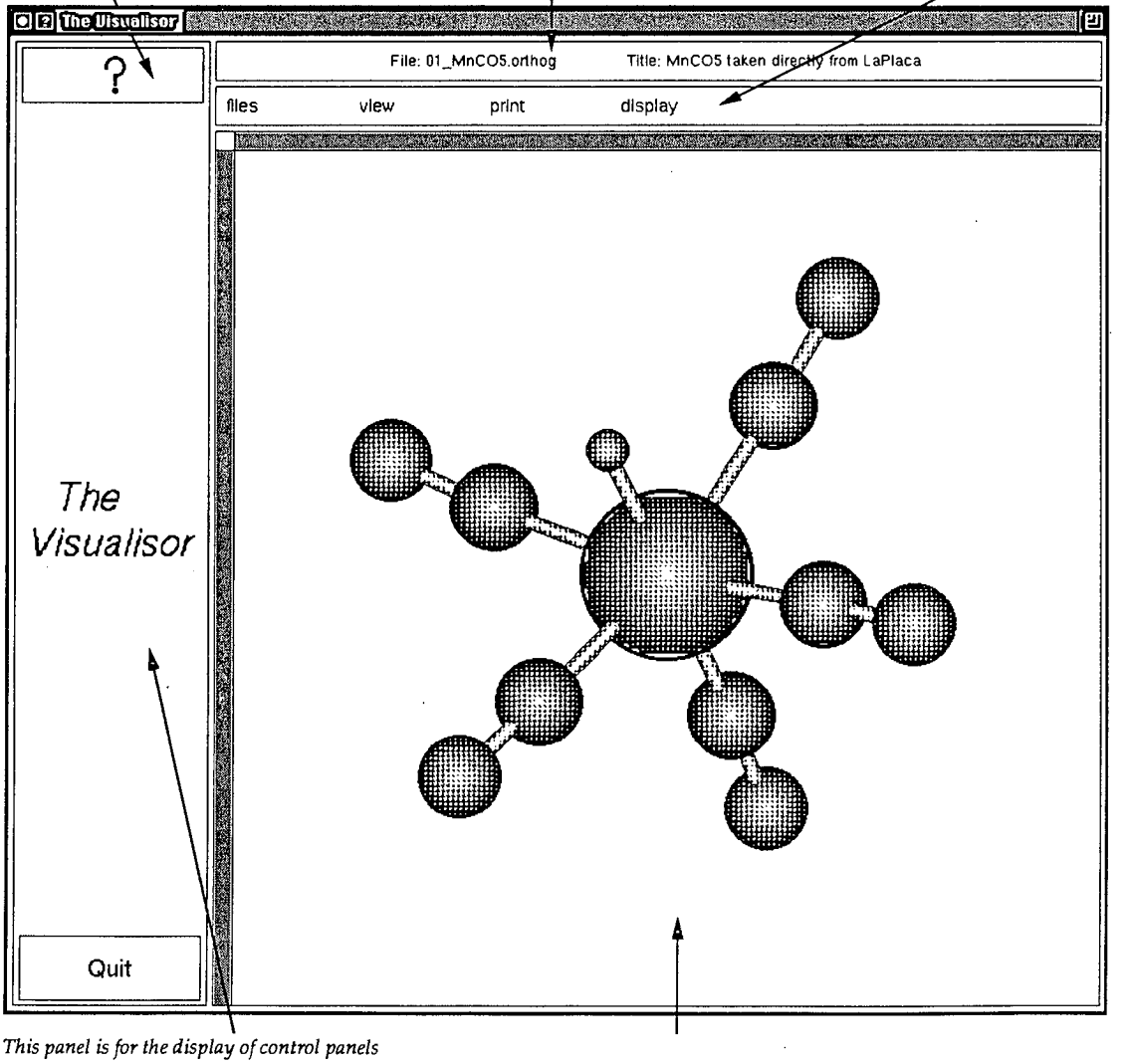


Figure 4–1: A screen dump of the Visualisor.

With the exception of the help facility, all user operations are collected into appropriate categories listed in the menu panel and are initiated *via* pull down menus. Where further actions are required a control panel appears over the side panel — as indicated in the diagram — allowing the user to carry out those further actions. On completion the user may select the "done" button at the bottom of the panel and the panel will be deleted. The panels for rotation and shading respectively are shown in Figure 4–5.

The facilities provided by the program are organised into a menu structure for intuitiveness and ease-of-use. Figure 4–2 shows a representation of the menu structure of the Visualisor — the facilities that are available to the user are discussed in more detail under the appropriate menu headings.

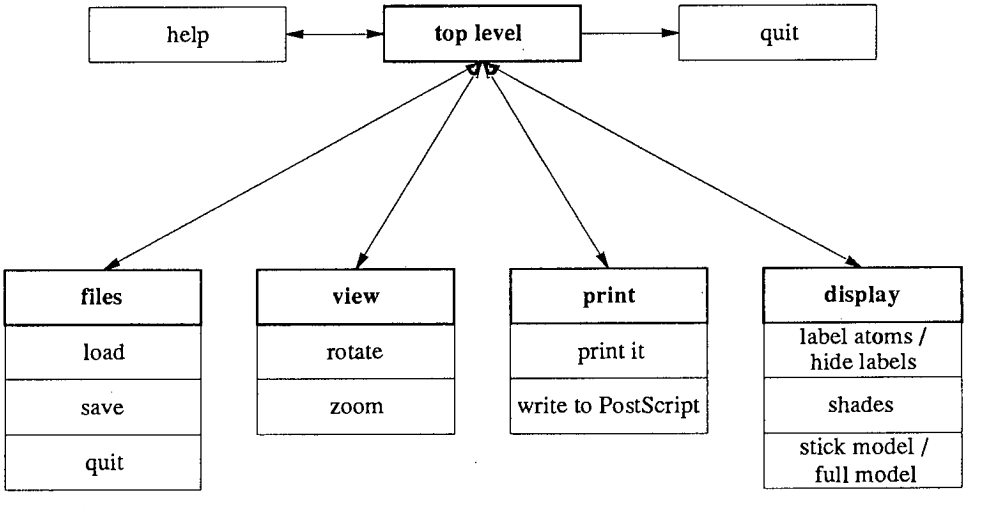


Figure 4–2: Menu Structure of "The Visualisor".

4.4.1 Files Menu

As identified in the list of objectives earlier in this chapter, molecular models may be loaded, and subsequent operations such as rotations or alterations to the display performed thereon may be saved. In order to retain the standard nature of the molecular model file the operations are stored in a new file. This is

automatically generated with the name *file_name*.sav, and is referred to in this text as the *view file*. This file is subsequently recognised and loaded automatically whenever the original model file itself is loaded. Recognition is based on the fact that the two files have the same root name excluding the .sav extension. If the molecular model object 01_MnCO5 is loaded for instance, the Visualisor will search the directories described below for the view file 01_MnCO5.sav. Figure 4–3 shows an example of such a file, containing the rotation matrix and key for selected shade.

```
rotation_matrix.x  -0.681998  -0.000000  -0.731354
rotation_matrix.y   0.687248  -0.342020  -0.640869
rotation_matrix.z  -0.250138  -0.939693   0.233257
shade  -2
```

Figure 4–3: An example .sav *file.*

On attempting to load a file, the user is presented with a list of all the files in the current directory by default, controlled by a scroll bar. The user need only select the appropriate file with the mouse and click the load button as shown in Figure 4–4.

When loading a file, the user is offered the *current directory* by default, that directory from which the Visualisor is invoked. However, the program may be directed to look elsewhere in the directory structure by appropriately setting *environment variables, e.g.* "AJW_MODEL_PATH" "AJW_SAV_PATH". Environment variables are Unix facilities which are adopted by the tool-kit specifically for this purpose. In the *bash* shell for example, the command

```
export AJW_MODEL_PATH=$HOME/models
```

would direct the Visualisor to search for the molecular model object in the "models" sub-directory of the user's home directory. Similarly, the command

```
export AJW_SAV_PATH=$HOME/models:$HOME/models/save_files
```

would direct the Visualisor to first search for the view file in the "models" sub-directory of the user's home directory, and if the view file is not found there,

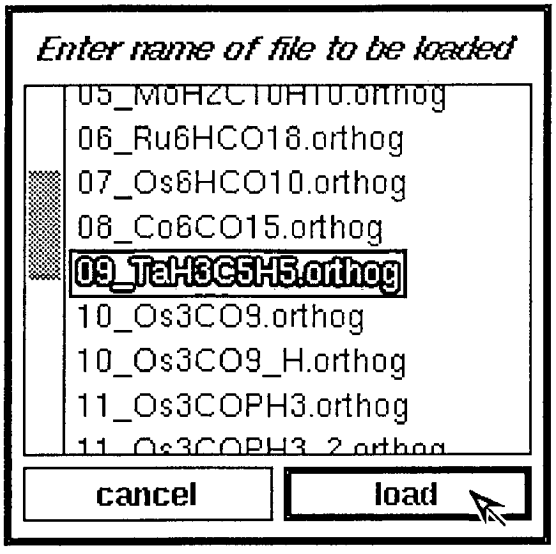


Figure 4–4: The file loading mechanism.

then to look in the directory "`models/save_files`" sub-directory of the user's home directory. The environment variables appropriate to this application, and the information they convey, are:

`AJW_DATA_PATH:`	source directory for atomic data files.
`AJW_XBM_PATH:`	source directory for bitmaps used by the application.
`AJW_MODEL_PATH:`	source directory for molecular models.
`AJW_SAV_PATH:`	source/target directory for view files.
`AJW_PS_PATH:`	target directory for PostScript files.

4.4.2 View Menu

Models may be rotated in three dimensions by discrete amounts of 1°, 5°, 10°, 45°or 90°. The current incremental angle is selected by means of a series of "radio buttons" for each axis and direction as shown in Figure 4–5. The user need only click on the button with the appropriate number of degrees — one of the *increment* buttons. Logically, only one number may be chosen at one time since the selection of a new button will automatically deselect the previous. To

actually rotate the molecule, the user need only click on the appropriate *direction* until the desired rotation has been achieved. The view may be fine-tuned by selecting the small granularity increment buttons. Although this process seem complex, it must be remembered that this is an interactive process performed by the depressing of buttons rather than the error-prone and unintuitive manual editing of files. Moreover, the results of the rotation may be seen immediately. The rotation buttons are marked with graphical rather than textual representations of the rotation axis for greater intuitiveness.

It is possible for the user to zoom in or out of the graphics display area to view the model in more or less detail. Scroll-bars are provided in zoom mode to enable movement across the zoomed model. The "zoom in" and "zoom out" buttons are labelled graphically in a similar way to the rotation buttons for intuitiveness.

4.4.3 Print Menu

A view of the model may be printed directly or a PostScript version written to a file for editing or printing at a later date. The destination filename offered to the user is the original root file name with a "`.ps`" extension. This may be edited to whichever name the user chooses. In the example shown in Figure 4–6 the original molecular model file is `01_MnCO5.orthog`. The offered PostScript filename was `01_MnCO5.orthog.ps` which has been subsequently edited to `01_MnCO5.ps` by the user. Similarly, for direct printing, the user may specify any desired printer options simply by editing the text in the dialogue box.

The PostScript output file is written to be compatible with the format adopted by the Idraw drawing package. The output from the Visualisor may be loaded and edited within this program. This editing may include many operations such as annotation, highlighting and shading. Idraw provides many features such as pattern filling and brush and font selection for enhancing the picture of the molecule. The objects within the PostScript file are grouped to facilitate editing in Idraw. For example, the graphics primitives which constitute a bond

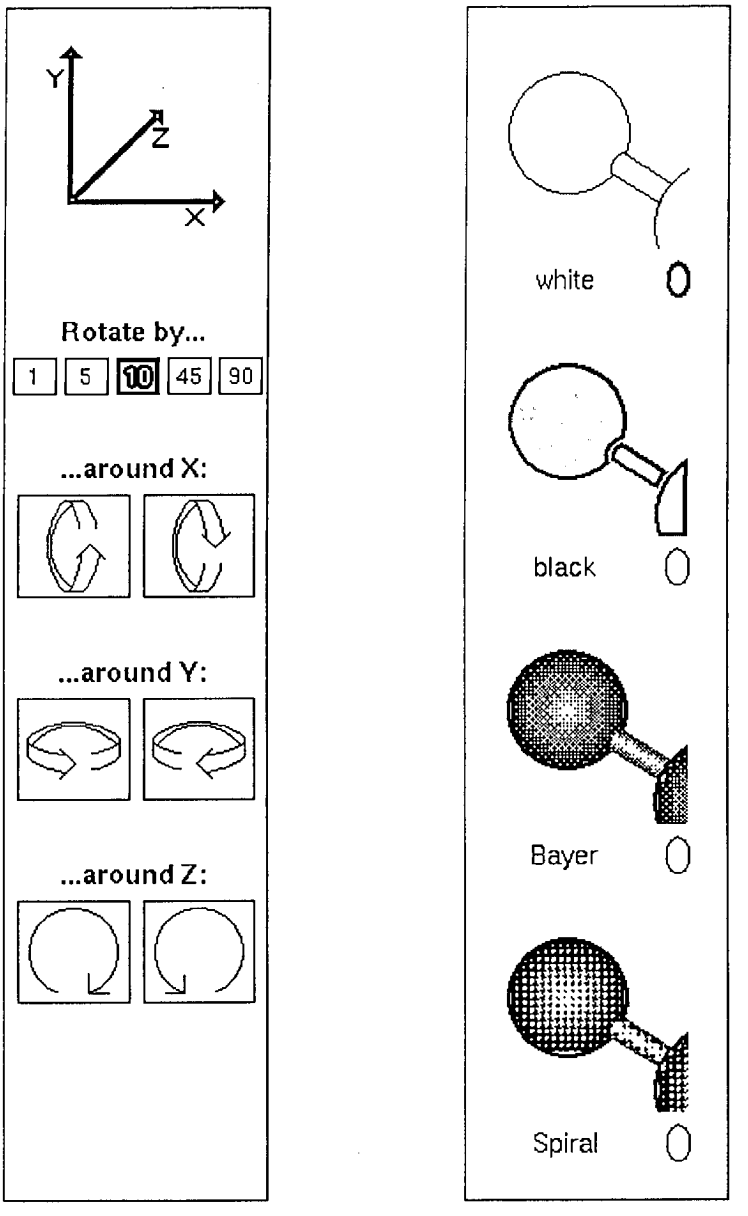


Figure 4–5: The panels for rotating and shading the view of the molecule.

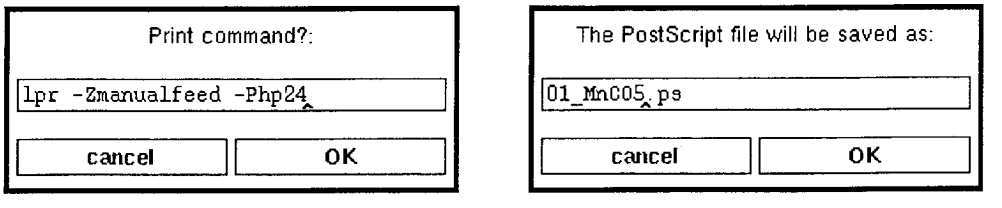


Figure 4–6: The dialogue box for directly printing or saving a file.

are grouped together so that individual bonds may be shaded, whilst bonds and atoms are grouped together to constitute a molecule. Atom labels are grouped together, but separately from the graphics constituting the image of the molecule itself. The former may therefore be moved without any risk of altering the latter.

4.4.4 Display Menu

The model may be displayed on the screen in four different shading modes, as shown in Figure 4–5. Atom labels may be selected or deselected. Molecules may be displayed as a simple line-drawing as shown in Figure 4–7 by selecting an option from the *display* menu.

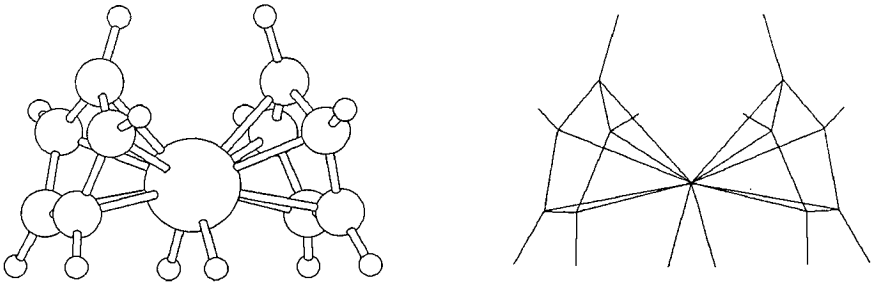


***Figure 4–7:** Examples of possible display.*

4.4.5 Chemical Knowledge

Atomic data are read in from the atomic data file as described in Chapter 3. In this case only the atomic radii are required, to determine both the respective display sizes of the atoms within the molecule and the bonded atom pairs. The intra-molecular bonding is determined by the comparison of the distance between the two atoms potentially linked by a proposed bond and the sum of their atomic radii. It was observed that the rule

$$\text{bond exists} \Leftrightarrow \text{interatomic distance} \leq \text{sum of atomic radii} + 0.3\text{Å}$$

proved to be successful in determining bonding, with the program correctly finding all of the bonds in each of Mitchell's thirteen models, as listed in Chapter 5.

4.5 Evaluation

Ideally, the Visualisor would be evaluated by thorough exposure to prospective users who are competent in the field of the application, but are new to the program. In order to identify errors, inconsistencies or undesirable features which are manifested only intermittently, a thorough evaluation would have to involve extensive use of the program.

Unfortunately this is not an ideal world. The designer, in this case, did not have access to a large user community willing to spend a substantial amount of time evaluating programs. Thus she was required to adopt a secondary rôle, that of a user. Although the software should preferably be evaluated by prospective users who are new to the program. It was considered that evaluation by the designer has some validity. Long-term constant use by someone familiar with the package is likely to expose different problems to a short evaluation session; the two methods may complement each other.

The programs discussed in this thesis have been used by the designer almost constantly since they were first implemented. Evaluation by the designer is more likely to pick up those errors which only occur intermittently. For example, it was discovered after a few months use that the routine for rotating the molecule was erroneous. However, because the symptoms were only apparent under certain conditions this error may not have be discovered by an explicit user-evaluation procedure. An example of a design issue rather than an error also arose from the molecule rotation procedure. The original design allowed a separate angle increment selection mechanism for each axis. After several months use it was discovered that this was unnecessary and, at times, counter-intuitive. Experience of the tool indicated that a single angle increment applied

to all axes and directions was more natural and convenient. As soon as such bugs or features became apparent, the corresponding sections of the Visualisor code could could be corrected or redesigned.

The Locator was demonstrated at the International Conference on the Chemistry of the Copper and Zinc Triads at Edinburgh in July 1992 where the interface received favourable feedback. As discussed in Chapter 6, the Locator adopts the Visualisor for an interface.

4.5.1 Achievement of Objectives

The Visualisor was designed according to both the generic objectives detailed in the previous chapter and specific objectives for a molecular visualisation tool. The degree to which the objectives detailed in the previous chapter have been met is considered first. The application of relevant software engineering and HCI issues to the implementation of the tools within the tool-kit has already been discussed in the previous chapter. The objectives detailed in Chapter 3 overlap with these software engineering and HCI issues in some cases and so have already been discussed — 1. Direct Manipulation, 2. Intuitiveness, 3a. Consistency (Resource Re-use), 5. Input Flexibility, 10. Portability and 11. Extensibility. Thus, discussion in this chapter is confined to the achievement of the objectives not covered in the previous chapter.

3b. *Consistency (Look and Feel):*

Since the Locator adopts an interface based on the Visualisor, the two major tools show an almost identical interface. Identical commands have identical functions in each tool. The only difference between the two is that the Locator has additional commands. Internally the interface is consistent, commands are initiated from the pull-down menu to be followed up, if necessary, on the side control panel. An exception is the main controls for the Locator, which are continually displayed on the main panel for convenience. It is considered that other tools which involve

molecular models and would benefit from an interactive graphical display feature could adopt the Visualisor as an interface. Thus, this objective is considered to have been achieved.

4. *Integration of Programs:*

   The programs are integrated in that the file specifications for each are identical. The output file from one may form direct input to the other. In fact, the code which reads and writes data is shared between the tools. Any appropriate change to the file format may be propagated throughout both tools by a single modification to this code. The integration is taken one step further by the use of one tool as an interface to the other. Thus the user does not have to wait for the hydride location procedure to converge and terminate and may view intermediate results.

6. *Expertise Flexibility:*

   It is considered that the direct manipulation interface would allow a novice user to make full use of the Visualisor without the need to consult a manual. If confronted with any problems there is a clearly labelled button to initiate a context-sensitive help system. The Visualisor also addresses the needs of the more expert user. Run-time arguments are available to enable shortcuts, for example to load molecular model files. The Visualisor has been in constant use by the designer, since its completion. This has given ample time for the evaluation of the use of the Visualisor by an expert user — the designer. Any features which were found to be unnecessarily time-consuming or cause frustration in any way during this time were re-designed to be less so, whilst maintaining intuitiveness for the novice user. Although the use of the menu-system was not found to cause frustration to the designer, it is feasible that some experienced users may find its constant use irritating. The provision of alternative menu selection by interactive keystrokes would reduce this. This is discussed in greater detail later in this section under "further work".

7. *Structure:*

   The only input object required from the user is a molecular model file. Manipulations are automatically stored in a another file. All other data are either stored internally or set interactively.

8. *Defaults:*

   All data "facts", such as atomic radii, are stored internally. Bonding interactions are automatically inferred but may be overruled by the user if so desired.

9. *Appropriate Presentation of Data:*

   Being a visualisation tool, the appropriate presentation of data by the Visualisor is paramount. The form chosen to represent a molecular model is the popular ball-and-stick representation. This representation is considered adequate to convey and aid assimilation of the molecular model for most cases envisaged. Depending on the application of the tool, however, other representations may be more appropriate. For instance, space-filling or solvent accessible surface representations may be more appropriate for reaction planning. This is discussed in greater detail later in this section, and again under further work. It is considered that this objective has largely been met.

Secondly, the degree to which the specific objectives for molecular visualisation detailed in this chapter have been met is considered. Several of the objectives were fully achieved and need little further discussion. The default representation is of a ball-and-stick type (objective 12), the molecular representation may be rotated in three dimensions (objective 14), printed directly to a selected printer (objective 18) and the user may zoom in (and subsequently zoom out again) on any part of the image (objective 15). The degree to which the other objectives have been met requires more extended discussion.

13. *User Controlled Atom and Bond Sizes*

    This objective was not achieved in the allotted time. The user does have

some control over atom and bond sizes, however, this is minimal, being a simple choice between the standard ball-and-stick and "stick" representations shown in Figure 4–7. It is possible to change the atom and bond sizes by simply manually editing two parameters in the source code. However, since this is not interactive and would furthermore require recompilation of the program it is considered unsatisfactory.

Some ideas regarding the implementation of this facility are discussed later in this chapter under further work.

16. *View Saving*

As discussed earlier, a particular view of the molecule may be saved by storing the current rotation matrix and shade parameter to a file. The data from this file are automatically restored if the associated molecular model is subsequently reloaded. However, it is considered that the method of recognition of the associated view file could be improved. Currently the two files are only matched if they have the same root name, disregarding the "`.sav`" extension. This would appear to be unsatisfactory for various reasons. Firstly the user is required to remember to update the names of both files if the naming convention is modified. Secondly, it is feasible for the user to have differing version of the same molecular model for which the same view file would be valid. Currently, the user must maintain multiple copies of the same view file in such a situation, one for each model. Thirdly, the user may wish to store several views of the same model. In conclusion, it is considered that this objective has been met, but that the "tagging" method could be improved.

17. *Graphic Output*

When the user selects the "write to PostScript" option the resultant file is a view of the current image in, as the command would suggest, PostScript. PostScript is a generally accepted standard page description language. Many laser printers can print PostScript objects directly and many tools exist to preview and edit PostScript objects. The PostScript generated

by the Visualisor is compatible with the graphics editor, Idraw. This enables images output by the Visualisor to be directly loaded into Idraw for enhancement or highlighting of certain areas. Thus, it is considered that this objective has been fully achieved.

19. *Consistency*

Currently when the user selects the *"write to PostScript"* or the *"print it"* option the subsequent PostScript output is of a model represented in a certain display style, not necessarily matching the style of the model displayed on the screen. The resultant view will be of a "full model" representation rather than a "stick model" representation, in white with black borders and with the atom labels selected. Figure 4–8 shows an example of such a PostScript image. In this respect, the objective of consistency was not fully achieved.

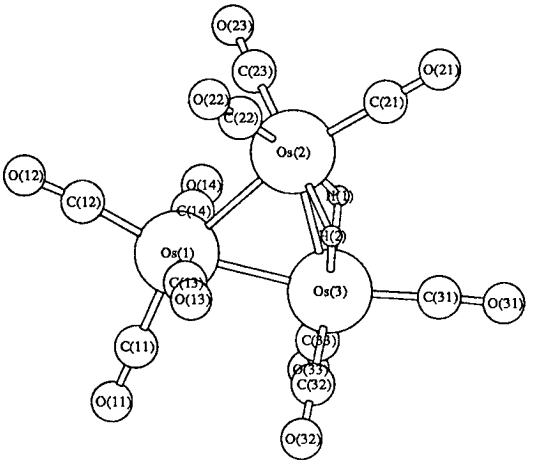


Figure 4–8: An example of a raw PostScript picture produced by the Locator.

4.6 Discussion

Earlier sections have discussed how the graphical display of molecular models can convey much more information than a simple atom coordinate list. Graphic display takes advantage of many more of the user's faculties. The molecular models to be visualised can come from a variety of sources. Perhaps the most common are:

Publications:

Molecular models can often be obtained from articles published in academic journals. These generally result from either X-ray diffraction or neutron diffraction.

Databases:

Models may be extracted from databases such as the Cambridge Structural Database. By agreement of the publishers, atomic coordinates for any model published in any British journal is required to be sent to the Cambridge Structural Database.

Experimental:

Compounds which are synthesised in the laboratory are often characterised by X-ray or sometimes neutron diffraction, resulting in a molecular model.

Theoretical:

Theoretical modelling tools such as that described in the following chapter can produce either a new molecular model or a revised version of a previous model.

Models obtained from any of the above sources may be loaded directly into the Visualisor, provided they contain the appropriate data and keywords. However,

this is considered to be a somewhat limited use of the Visualisor. Admittedly, little else is appropriate for models obtained from publications. However, the Visualisor could play a more integral part in obtaining models from the other sources — databases, analytical experiments and theoretical modelling. The Visualisor could provide an appropriate interface for the program producing or dealing with the molecular models in such cases.

A molecular database would be a good candidate for an application which may benefit from such an interface. With some adaptation, the interface would facilitate browsing of the database for required models. The user could graphically specify groups, atoms, or any other key in the model. The specification would thereby become more and more (or if appropriate less and less) refined until the appropriate models are identified. Similarly, the X-ray crystallographic structure solution program SHELX could benefit from a molecular visualisation interface. New, potential, atoms could be displayed after each refinement iteration, along with but distinguished from atoms of established location if appropriate. Each proposed atom site may be accepted or rejected before proceeding with the next iteration. The potential of the molecular visualisation tool as an interface for another application is demonstrated by the hydrogen location tool described in the next chapter. The "kernel" of the hydrogen tool itself remained essentially unchanged after adopting the Visualisor as an interface. Only the surrounding "husk" dealing with the input and output of data is affected by the incorporation of this interface.

4.7 Conclusions

It is acknowledged that other molecular graphics tools (see literature review) can provide more sophisticated pictures of molecules. These may render features such as colour coordinated functional groups, solvent accessible surfaces and high quality ray-traced displays. However, two of the strengths of this tool are considered to be its simplicity and the fact that it is an integral part of a

larger environment. Perhaps the greatest strength of this tool is not the tool in itself rather its potential as a common user interface for any tool which deals with molecular models. Referring back to the user model in Chapter 2, molecular models may be employed in all of the aspects of chemistry, such as for publications and databases in *data acquisition*, theoretical modelling in *theoretical chemistry* and crystal structure solution in *analytical chemistry*.

4.8 Further Work

Several of the suggestions for further work discussed in this section stem from the shortfalls in fully realising the objectives detailed earlier in this chapter. In such cases the original objective is referred to.

4.8.1 Atom Number Restrictions

The maximum number of atoms for a molecular model which can be loaded into the Visualisor is currently limited to 200, purely because of the data structure employed. It would be a simple matter to redesign the data structure to remove this restriction on atom numbers.

4.8.2 Loading and Saving Files

On attempting to load a model, the user is offered files from the current directory by default, unless another directory is specified by the appropriate environment variable as described earlier in this chapter. Although lower or higher directories in the file structure may be selected, it is considered desirable that a more *interactive* control should be provided over the directory initially presented. Such control would be desirable for all the file loading and saving operations, *e.g.*, the saving of manipulations and writing of PostScript, which is currently

always written to the current directory. Possible approaches for addressing this issue include:

- The provision of an option in the files menu to enable the selection of a directory to become the default directory.
- The provision of a feature to remember the most recently accessed directory and use this as the default directory.

In either of the above cases, there is a further issue as to whether a single default directory is maintained, or several separate directories are maintained, one for each type of file saved or loaded, such as model or PostScript files.

4.8.3 Display

It is considered that the display of a molecule could be improved in several ways:

Atomic Labels

It may be desirable for a feature to be available which allows an atom's label to be moved interactively to a desired location, to be stored along with the other "manipulations" in the *.sav* file. An appropriate implementation would seem to be the well-used method of allowing the user to "pick up" a label with the mouse button to be moved to a new location. This is generally achieved by positioning the cursor over the object to be moved, depressing the mouse button, moving the cursor to the new location and releasing the button. The movement of the label may be indicated by;

- simply redrawing the object at the new position when the mouse button is released,
- mapping the movement of the cursor with an outline of the object or

- mapping the movement of the cursor with a representation of the object itself.

HCI ideals must be set against execution time in determining which of these options is the most appropriate.

Variable Atom and Bond Sizes

Failure to fully achieve objective 13.

The atom and bond display diameters are currently "hard-wired" into the program. They may not be changed by the user under normal circumstances, that is, without re-compilation. A more satisfactory situation would allow both the atom and bond display diameters to be adjusted from a stick representation at one extreme, to a space-filling representation at the other. A possible implementation would be the provision of sliders to control the atom and bond diameters, with the molecule being redrawn at every new setting to enable the user to fine-tune the display. Furthermore, it would be desirable for the user to be able to select groups of atoms, particularly by element type, for group resizing.

Perspective

Incorporating perspective into the display may make the model look more "realistic" and possibly give the user a greater perception of the depth relationships of the atoms in the molecule. However, absolute distances can be more difficult to judge from perspective than orthographic images, thus the user should be able to select or deselect a perspective view as desired.

Alternative Representations

The ball-and-stick is just one way of representing molecules, and is not necessarily the most appropriate for all situations. Some other programs adopt other

representations. The molecular graphics program Ortep [88] allows users to represent atoms as *thermal ellipsoids*, that volume described by the movement of individual atoms with respect to the rest of the molecule. Connolly [1,2] describes a method for representing the molecule by solvent accessible surfaces. Each of these representations conveys different information and may be the most appropriate representation for certain users in certain situations. Thus it is considered desirable that both of these representations, and any others if appropriate, should be available to the user of the Visualisor.

4.8.4 PostScript Output

Failure to fully achieve objective 19.

The printed and PostScript views of the molecular model do not necessarily match that on the screen, as described earlier. In the interest of consistency it is considered that the PostScript output *should* exactly reflect that on the screen. Each option chosen by the user for the screen display should apply also to the PostScript output.

4.8.5 Keyboard Shortcuts

As described earlier it is considered that the provision of keyboard shortcuts — the use of associated keystrokes instead of menu selection — may be beneficial to experienced users. Figure 4–9 shows suggested shortcuts which could be implemented.

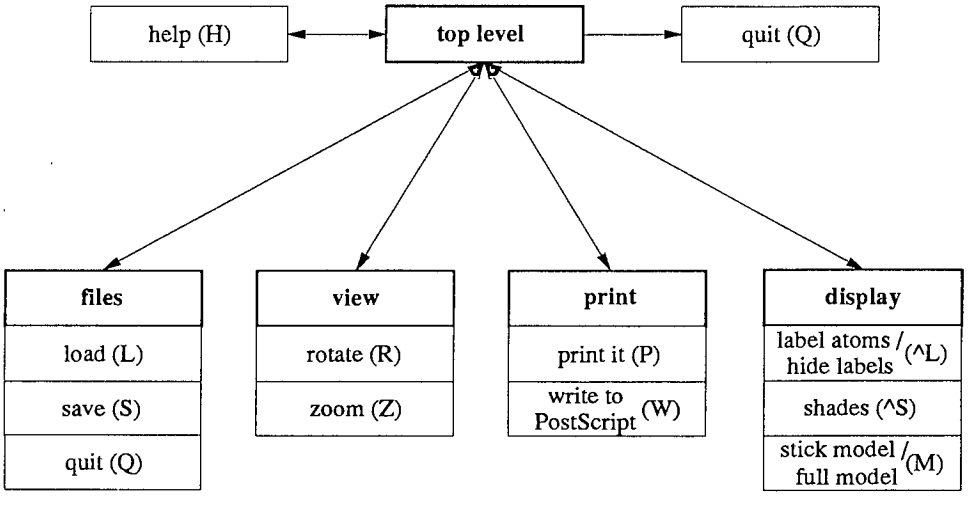


Figure 4–9: Suggested Keyboard Shortcuts.

Chapter 5

Hydride Ligand Location in Transition Metal Complexes

5.1 Introduction

Transition metal cluster compounds continue to be the subject of considerable academic and industrial interest [89–92]. Bulk metals may be modelled with transition metal cluster compounds, since the packing in the two is similar. Industrially it is important to understand how small molecules such as CO, H_2 and C_2H_4 *etc.* are chemisorbed onto metal surfaces, but precise studies of this are difficult because the bulk metals are generally insoluble. By studying soluble molecular species it is possible to extrapolate from the metal clusters to metal surfaces.

5.1.1 Characterisation

The physical, chemical and even biological properties of a molecule are determined not only by the atoms contained but also by the relationships between these atoms [93]. The number and types of bonds between the atoms, the length of those bonds, the presence of functional groups, the arrangement of the atoms

in space and the way in which these atoms are linked by the bonds can all contribute to the molecular characteristics thus:

- physical properties such as boiling point and colour,
- chemical properties such as reactivity and acidity and
- biological properties such as toxicity and biodegradability

can all be affected by the afore-mentioned structural characteristics of the molecule. Therefore, accurate knowledge of the spatial distribution of all the atoms is crucial to any understanding of the properties and chemical reactivity of compounds and the development of bonding theories. Unfortunately, the proper characterisation of a transition-metal hydride complex is one of the more difficult problems facing a structural chemist [94].

5.1.2 X-ray Diffraction

The determination of structure can be accomplished by a single crystal *X-ray diffraction* experiment [95, 96] in most cases. Unfortunately, however, the nature of X-ray diffraction renders it effectively useless in locating metal-bound hydrogen atoms in a important class of cluster compounds, the hydride metal clusters [91]. X-rays are scattered by electrons; consequently the magnitude of the diffraction by any atom in the molecule is proportional to the electron count of that atom. Only one electron orbits a hydrogen nucleus; consequently a hydrogen atom produces an insignificant diffraction when compared with that produced by the much larger transition metal atom. The problem is exacerbated, not only as the metal atom gets larger — titanium has only 22 electrons whereas mercury has 80 — but also as the hydrogen atom is bonded to more metal atoms [94, 97–99], as in the case of edge-bridging (bonded to two metal atoms) and face-bridging (bonded to three metal atoms). Therefore, the hydride atom contribution to X-ray scattering is a small fraction of the total scattering [13] and the electron density owing to hydride ligands is often swamped by spurious ripples in the vicinity of the metal atoms, even if the experiment is carefully performed. In addition, the

distance between the hydride atom and the metal atom is small (of the order of 2Å). To resolve this requires data of reasonable resolution and hydrogen atoms only scatter weakly at such angles. Moreover, in transition-metal hydrides the hydride ligands are often protonic. As the bond tends to M^-–H^+, — the hydride ligand losing its electron — the hydride ligand would become transparent to X-rays [100]. X-ray diffraction is currently employed, and is improving, but even when the hydride ligands are directly located their positions are still of low precision [94]. M–H distances measured by X-ray methods are often 0.1Å – 0.2Å shorter than their true values because the hydrogen atom electron density is somewhat perturbed from the nucleus toward the M–H bond.

5.1.3 Neutron Diffraction

A solution to this problem is the complementary technique, *neutron diffraction.* Although the techniques and theory of the two scattering experiments are similar, neutron diffraction has an advantage over X-ray diffraction. Hydrogen atoms scatter neutrons with similar efficiency to most other elements since the size of the signal is governed only by the existence of a nucleus. Thus the technique does not suffer from bias against hydrogen atoms. Unfortunately, however, this technique is not routine since:

- large crystals (approximately $10mm^3$ [94]) are required because of the low flux of neutron beams [101],
- the source of neutrons is prohibitively expensive (requiring a nuclear reactor or a high energy accelerator) and
- the overall results are typically of lower precision [94].

5.2 Current Methods of Hydride Ligand Location

Accordingly, several methods have been developed in which likely sites for hydride ligands are suggested by empirical considerations, given a conventionally determined non-hydrogen molecular skeleton.

5.2.1 Steric Methods

Often it is possible to infer hydride ligand positions by recognising their stereochemical effect on the rest of the cluster. The bending back of adjacent ligands may suggest [102, 103] or confirm [104] the presence of a hydride ligand. Similarly, the lengthening of the bonds between metal atoms may suggest [102, 105, 106] or confirm [101, 103] the presence of an edge or face-bridging hydride ligand. Conversely, interstitial hydride ligands have been inferred by the lack of any of the above mentioned effects [107, 108]. By observing such structural irregularities the topological position of the hydride ligand may be conjectured.

Limitations

To use such an approach is a crude way of solving an important problem since the method is flawed by the assumption that there are no other origins for a small change from the mean bond length [13], and chemistry being an inexact science it is impossible to consider all the possible reasons for such a lengthening. Even at their most successful these methods can only assign the topology of hydride ligands — not their coordinates with respect to the other atoms [13]. At their worst steric methods could predict incorrect hydride ligand positions [12]. For example, an angle of 159°was predicted for W–H–W in $W_2H(CO)_9(NO)$ [94] which neutron diffraction subsequently revealed to be 125°. This indicates an entirely different bonding interaction. In $[Co_6H(CO)_{15}]^-$, the X-ray analysis

revealed that one Co–Co bond was significantly longer than the other and the carbonyl groups surrounding the long bond seemed to be splayed out, indicating an edge-bridging hydride ligand. Neutron diffraction revealed the hydride ligand to be interstitial [94].

Using steric methods for hydride ligand location frequently leads to ambiguous or even erroneous results and such approaches are more appropriate for confirmation rather than determination.

5.2.2 Potential Energy Calculations

Theory

One of the better empirical methods is the potential energy approach developed independently by Ciani [11] and Orpen [12], using the method of LaPlaca and Ibers [109]. In this approach the position of the hydrogen atom is located by minimising a *modified repulsive potential energy function*, described by empirical parameters, of the form

$$M = \sum_i [ae^{-br}r_i^{-d} - cr_i^{-6}] + \sum_k \left|\frac{(R_k - R)}{S}\right|^2$$

The first term, $\sum_i [ae^{-br}r^{-d} - cr^{-6}]$, is the potential energy (kcal mol^{-1}) attributable to the interaction between the metal-bound hydride atom and all the nearby ligands/atoms to which the hydride ligand is *not* bonded. The parameters in this part of the equation are as follows:

- i is the number of atoms *not bonded* to H,
- a–d are numerical constants dependent on the atom type in the H$\cdots$X contact [110] and
- r is the distance between H and the ith of these atoms.

The second term, $\sum_k |(R_k - R)/S|^2$, is the potential energy for the interaction with those atoms which *are* proposed to be bonded to the hydride atom. The parameters in this part of the equation are:

k is the number of atoms *bonded* to H,

R_k is the interatomic distance to the kth of these atoms,

R is the specified M–H distance and

S is the permissible error (usually ≈ 0.05 Å).

Hydex

The potential energy method was implemented by Orpen in the computer program *Hydex*. Hydex works by scanning a candidate area established by the sphere, circle or points defined by user-defined information described below. All bridges are assumed to be symmetric, and an atom/ligand is deemed to be "nearby" if it is within approximately 3.5 Ångströms.

To locate hydride ligands with this program, the user must provide:

1. an (X-ray determined) skeleton of the model, supplemented with calculated coordinates of the non-hydride ligands when necessary,

2. parameters describing the non-bonding interactions,

3. the nature (τ, μ_2, μ_3 or ι) of the site(s) to be investigated — in this thesis the terms τ, μ_2, μ_3 and ι denote *terminal, edge-bridging, face-bridging* and *interstitial* hydride ligands respectively,

4. the atoms to which the hydride ligands are bonded and

5. the M–H bond lengths appropriate to the site, with permissible variations in these bond lengths.

○ *Limitations*

The program assumes knowledge of the topological position of the hydride ligand — the "nature" of the hydride ligand and the atom(s) to which the hydride ligand(s) is/are bonded. If this information is unknown then the most likely position for the hydride ligand can only be found by trial and error.

Here, the user is obliged to try candidate topologies and manually compare the calculated energy of the molecule in each case.

The method often works well, but is based on the *artificial* premise that the hydrogen atoms occupy positions of least steric repulsion in the specified area of the molecule and not on the *correct* premise that hydrogen atom positions reflect the molecular stereochemistry at which stability of the molecule as a whole is maximised. Thus, the potential energy method simply minimises repulsions between hydrogen atoms and nearby non-bonded ligands whereas a more appropriate method would be one which also maximises attractions between the metal atoms and the hydrogen atoms.

The imperfection of the potential energy approach is highlighted by the growing number of examples for which this method fails to correctly predict hydrogen atom sites. Notable amongst these are cases where the individual metal co-ordination geometry is not close packed, for example, $Rh_3H_3\{P(OMe)_3\}_6$ [106], cases where the hydrogen atom bridges a metal–metal bond asymmetrically (*e.g.* $Os_3H_2(CO)_{10}(\mu_2\text{–}CH_2)$) and several mononuclear polyhydride species [111]. It is acknowledged [12] that the nature of the potential energy method means that in must fail for interstitial hydride ligands, yet such species are amongst the most relevant in establishing the often suggested analogy between discrete transition metal cluster compounds and metal surfaces [112]. Methods based on molecular orbital (MO) approaches are therefore appropriate since, in this case, the attractions between the metal atom(s) and hydride ligand(s) present are *maximised*, as opposed to the *minimisation* of repulsive forces employed by potential energy methods.

5.2.3 Molecular Orbital Calculations

Theory

∘ *Atomic Orbitals*

Solution of the Schrödinger wave equation ($H\psi = E\psi$) leads to a series of eigenvalues (E_n) and eigenvectors (ψ_{nlm}). The eigenvalues correspond to energy levels whereas the eigenvectors correspond to the atomic orbitals:

$$\psi_{nlm} = \Upsilon_{lm}(\theta, \phi)R_{nl}(r) \tag{5.1}$$

$\Upsilon_{lm}(\theta, \phi)$ is the angular component, which is easily evaluated and will not be discussed further. The radial term $R_{nl}(r)$ is less easy to evaluate, engendering the need for approximate methods.

∘ *Molecular Orbitals*

In describing a molecule the atomic orbitals (ψ) combine to produce *molecular orbitals* (Ψ), thus it is necessary to deduce the molecular orbitals and their associated energies for a multi-nuclear system, and then consider how they are occupied by the electrons available. Molecular orbitals must satisfy the equation

$$H_F\Psi = E\Psi \tag{5.2}$$

where:

Ψ	is the eigenfunction (molecular orbital),
E	is the molecular orbital energy,
H_F	is the Fock operator, a combination of the core Hamiltonion operator and two two-electron operators.

It is necessary to solve Equation 5.2 iteratively — this is known as the Self Consistent Field (*SCF*) approach. One method of evaluating the molecular orbitals is the Linear Combination of Atomic Orbitals (*LCAO*) approach, and is represented by the following equation:

$$\Psi = \sum_i c_i \psi_i$$

where:

c_i is the coefficient of ψ_i.

o *Linear Combination of Atomic Orbitals*

From simple LCAO methodology, it is possible to derive a set of *secular* equations such that the determinant can be calculated to be:

$$\begin{vmatrix} H_{ii} - ES_{ii} & \cdots & H_{ij} - ES_{ij} \\ \vdots & & \vdots \\ H_{ij} - ES_{ij} & \cdots & H_{jj} - ES_{jj} \end{vmatrix} = 0 \qquad (5.3)$$

where:

H_{ii} $= \int \psi_i H \psi_i$ — the Coulomb integral,
H_{ij} $= \int \psi_i H \psi_j = \int \psi_j H \psi_i$ — the resonance integral,
S_{ij} $= \int \psi_i \psi_j$ — the overlap integral,
S_{ii} is unity if the atomic orbitals are normalised.

As stated earlier, it is the radial part which presents the problem in the evaluation of atomic orbitals (Equation 5.1). One approximate solution is to use Slater Type Orbitals (*STOs*), of the form:

$$R_{nl}(r) = r^{n^*-1} e^{-\zeta r}$$
$$\text{where } \zeta = \frac{Z - \sum S}{n^*}$$

n^* is related to the principle quantum number, a constant,
S is the screening number, a constant,
Z is the nuclear charge.

Thus, assuming the atomic orbitals are normalised, it is possible to solve the secular determinant in Equation 5.3 for two values of E, the lower one being the the energy of the Bonding Molecular Orbital (*BMO*), the higher one that of the Anti-Bonding Molecular Orbital (*ABMO*). The total, or configuration energy is the sum of the occupied molecular orbital energies and is written W.

Solving Ψ in such a way gives good results as it includes all atomic orbitals (a full basis set) in the calculations. However, because of the large number of calculations involved, this method is impractical for anything other than small molecules.

Empirical or semi-empirical methods which employ a restricted basis set tend to be adopted in practice. The remainder of this section describes such methods, and the assumptions which are made by them in the evaluation of Ψ.

○ *Hückel Method*

Hückel Molecular Orbital (*HMO*) theory makes the following assumptions:

1. σ and π molecular orbitals do not mix.
2. Atomic Orbitals are normalised ($S_{ii} = 1$).
3. Overlap between atomic orbitals is ignored ($S_{ij} = 0$).
4. All long-range interactions are ignored ($H_{ij} = 0$ unless the atoms are adjacent).
5. All atoms are equal.
6. Hydrogen atoms make no contribution

Furthermore H_{ii} is parameterised to be a constant α, whilst H_{ij} is parameterised to be β, transferable between molecules.

Thus the determinant in Equation 5.3 becomes

$$\begin{vmatrix} \alpha - E & \beta \\ \beta & \alpha - E \end{vmatrix} = 0 \qquad \begin{vmatrix} \alpha - E & 0 \\ 0 & \alpha - E \end{vmatrix} = 0$$

for adjacent pairs of atoms *for all other pairs of atoms*

and can easily be solved for the energies of the BMO and ABMO respectively.

The HMO method is a fast and easy method, which gives reasonably good results for organic π systems. However, since all atoms are assumed to be the same the HMO method is of little use for other systems.

∘ *Extended Hückel Method*

The Extended Hückel Molecular Orbital (*EHMO*) method [113–115] does not take such a simplistic view:

1. Every atom is considered.
2. The S_{ij}s are evaluated using single STOs.
3. Different atoms are recognised as such.
4. Long-range interactions are considered.
5. H_{ii}s are set equal to VSIEs.
6. H_{ij}s are calculated.

The extended Hückel method affords molecular energies that are of doubtful absolute value but which are reliable in a relative manner. Since the best sites for hydrogen ligands within one molecule are sought rather than any comparison between molecules, the EH method which is particularly appropriate for hydrogen ligand location is often acceptable. Moreover, the EH method is advantageous in terms of computing time and simplicity.

∘ *Modified Extended Hückel Method*

However, attempts to geometrically optimise structures using EHMO calculations generally fail because simple EH theory neglects two-body atomic repulsions. Thus the EHMO method was modified [116] such that account is taken of such repulsions. The MEHMO configuration energy is given by

$$W*(R) \quad = \quad W_{EH}(R) + W_{\beta}(R)$$

where

$W_{\beta}(R)$ is the two-body repulsion contribution.

Anderson has demonstrated a modified EH (MEH) approach which adequately reproduces not only the correct bond angles but also correct bond distances in optimised structures.

Complete Coordinate Convergence Program

An MEHMO approach has been developed locally [13] and implemented in the Complete Coordinate Convergence Program (*CCCP*). In this method all the atoms in the molecule are considered — not just those within a 3.5Å radius as in Hydex — and the bonding interactions between metal atoms and hydride ligands are included to produce an overall energy calculation. The hydrogen atom sites at which the consequent molecular structure is the most stable are then predicted with optimisation algorithms.

The program simultaneously varies the positions of some or all the hydride ligands using a *simplex* optimisation routine, and performs a MEHMO calculation on the molecule at each stage. The hydrogen atoms are freely refined and not restrained to predetermined metal-hydrogen distances. The only necessary prior knowledge is the number of hydride ligands to be located. For clusters with unknown hydrogen atom positions the combination of spectroscopic data and electron counting rules is generally sufficient to infer the number of hydride ligands. Other than that the program only requires the user to suggest a starting position for the simplex search.

○ *Limitations*

It had been observed through general use that the resultant hydrogen atom positions produced by CCCP are highly dependent on the initial starting positions — indeed to the extent that success or failure could depend on the estimated location given by the user. Furthermore, it was observed that this *implementation* of the MEHMO method appeared not to be parameterised for locating more than three hydride ligands, although it should be noted that the MEHMO *method* has no corresponding limitations.

Ab Initio Calculations

The *ab initio* method of molecular orbital calculations, as its English translation *"from the start"* would suggest, makes no assumptions in the LCAO calculation and considers all the electrons in the molecule, *i.e.* a full basis set of atomic orbitals is addressed. A terminally bound hydride ligand was located in $[HFe(CO)_4Mo(CO)_5]^-$ using this technique [97]. *Ab initio* molecular orbital calculations were performed on the X-ray determined skeleton together with a selection of possible locations in the vicinity of edge-bridging, semi-bridging and terminal sites. An apparently global minimum was detected near the terminal site

○ *Limitations*

By its nature of considering each electron in each atom, this method is computationally intensive. The previously discussed methods, Hydex and CCCP, are concerned with only the valence electrons, whereas in *ab initio* calculations the electrons in the sub-shells must be considered too. As an example, in the simple transition metal hydride $MnH(CO)_5$ an *ab initio* program would have 96 electrons to consider whereas CCCP would have only 58, a difference of 38 (see Table 5–1). A more indicative example would be $[Os_6H(CO)_{18}]^-$which produces the greater difference of 480 electrons, with CCCP considering 230 as opposed to 710 for *ab initio*. In addition to the greater numbers of electrons to be considered, *ab initio* calculations take into account the interactions between the electrons in the same orbitals (configuration interaction). In conclusion, it is acknowledged that *ab initio* MO calculations produce more accurate evaluations, and working at this level of calculation on real molecules is clearly the ultimate objective. However, for the larger transition metal clusters — those which contain multiple and/or heavier transition metal atoms, the calculations are, at this moment in time, too prohibitively complex to be practised routinely.

Atoms	CCCP	*ab initio*
H	1 × 1 = 1	1 × 1 = 1
C	5 × 4 = 20	5 × 6 = 30
O	5 × 6 = 30	5 × 8 = 40
Mn	1 × 7 = 7	1 × 25 = 25
Total	58	96

$MnH(CO)_5$

Atoms	CCCP	*ab initio*
H	1 × 1 = 1	1 × 1 = 1
C	18 × 4 = 72	18 × 6 = 108
O	18 × 6 = 108	18 × 8 = 144
Os	6 × 8 = 48	6 × 76 = 456
–	1	1
Total	230	710

$[Os_6H(CO)_{18}]^-$

Table 5–1: *Two examples illustrating the difference in numbers of electrons to be considered by two molecular orbital methods.*

5.3 Analysis of CCCP

Each of the hydride ligand location methods discussed above has merits and drawbacks. Steric methods may be somewhat unreliable, but if the main research interests of the chemist lie elsewhere then steric methods may suffice. At the other end of the scale, where the researcher is particularly interested in the position of the hydride ligand, a neutron diffraction study may be performed, but at great expense and requiring access to scarce resources. From the remaining methods, potential energy methods seem to work well for certain types of models, but not at all for others. *Ab initio* methods show great potential, but as was demonstrated earlier, are still computationally expensive. The MEHMO methods, specifically as implemented in CCCP, would seem to be a reasonable compromise, one which produces reasonable results for a range of molecules, but is not so computationally intensive as to prohibit casual use.

On this basis it was decided to investigate the program further, particularly the importance and extent of the apparent dependency of the results on the user-given starting position and the "ease-of-use" of the program for inexperienced or casual users. Thus a series of experiments was designed and performed to identify causes of and subsequently find a solution to this problem. These

experiments and their results are discussed below, together with a critique of the human-computer-interaction aspects of this program.

The results of the experiments were judged (where appropriate) on three criteria:

Success:

A qualitative assessment of the predicted hydrogen atom location. CCCP is deemed to have succeeded if the resultant model agrees with the topological description of the hydride ligand.

Accuracy:

The accuracy (given in Å) is the distance from the position found by neutron diffraction (or otherwise) to that found by CCCP. This quantity is henceforth known as the *distance error*

Speed:

The speed of the program is assessed by the iteration count. This is a major consideration since, for larger, more complex, models, CCCP can take up to three hours of computing time on a shared Sun 4/690 Workstation.

The models investigated in these experiments are listed in Table 5–2. The number alongside each model is a key to a larger list of molecular models in Appendix C which details references to source papers and Cambridge Structural Database keys. Further, each of the models listed in Table 5–2 is shown in picture form in Appendix D. These are models with which Mitchell [13] evaluated CCCP. Each model was specifically chosen by Mitchell as one for which Hydex failed, Hydex being the standard control for the evaluation. It should be noted that it is common practice to "reduce" molecular models by converting large functional groups such as $P(OCH_3)_3$ to smaller groups, in this case to PH_3, before performing EHMO calculations on them. It can be seen that such a reduction has been performed on some of the models listed in Table 5–2. The orbital parameters describing the atoms are shown in Table 5–3, although the values

Chemical Formula	*No.*	*Site*	*Comments*
Neutron diffraction test cases			
$MnH(CO)_5$	(5)	τ	Used by Mitchell [13] to tune the orbital exponent and the potential H_{ii}.
$Os_3H_2(CO)_{10}$	(76)	$2\mu_2$	This model has two, essentially symmetric, bridging hydride ligands bound to the same two osmium atoms.
$Rh_3H_3\{P(OMe)_3\}_6$	(86)	$3\mu_2$	Modelled[a] by $Rh_3H_3(PH_3)_6$. Hydex fails for this model as the metal atoms are not fully coordinated and it has a sparse ligand distribution.
$Os_3H_2(CO)_{10}(\mu_2\text{–}CH_2)$	(77)	$2\mu_2$	The hydride bridges are asymmetric. Hydex mistakenly assumes that hydrogen bridges are symmetric.
$MoH_2(\eta^5\text{–}C_5H_5)_2$[b]	(6)	2τ	This model has a sparse ligand distribution; Hydex fails for such models.
$[Ru_6H(CO)_{18}]^-$	(162)	ι	This model and the next have the same metal carbonyl skeleton but different hydride ligands; CCCP finds the hydrogen in both cases.
$[Os_6H(CO)_{18}]^-$	(127)	μ_3	(see above)
$[Co_6H(CO)_{15}]^-$	(161)	ι	An interstitial hydride; Hydex fails for interstitial hydrides.
$TaH_3(\eta^5\text{–}C_5H_5)_2$	(14)	3τ	Sparse ligands (see model 6).
X-ray diffraction test cases			
$Os_3H_2(CO)_9(PPh_3)_3$	(111)	$2\mu_2$	Modelled[a] by $Os_3H_2(CO)_9(PH_3)$.
$Os_3H_2(CO)_{10}(PPh_3)_3$	(174)	τ,μ_2	Modelled[a] by $Os_3H_2(CO)_{10}(PH_3)$. Hydrides located by X-ray diffraction. This model has an asymmetrically bridging (see model 77) and a terminal hydride.
$FeH_2\{(CCH_3)_2B_4H_4\}_2$[b]	(20)	2τ	Sparse ligands (see model 6).
$Fe_3H(CO)_9\{S(C_6H_{11})\}$[c]	(100)	μ_2	Asymmetric bridging hydride ligands located by X-ray diffraction, (see model 77).

a Adapted by substituting hydrogens for the appropriate groups and adjusting the bond lengths accordingly

b Dimer generated from the monomer using crystallographic symmetry.

c Cyclohexyl hydrogens added using CALC — this program is discussed in Chapter 2.

Table 5–2: A list of the models used.

of 0.85au^{-1} and -12.6eV respectively were taken for the exponent and energy of the metal-bound hydride ligand(s) [13].

			N			Exponent				Energy				
		e-	s	p	d	s	p	d1	d2	s	p	d	c1	c2
1	H	1	1			1.3				-13.6				
2	He													
3	Li	1	2	2		0.650	0.650			-5.4	-3.5			
4	Be	2	2	2		0.975	0.975			-10.0	-6.0			
5	B	3	2	2		1.300	1.300			-15.2	-8.5			
6	C	4	2	2		1.625	1.625			-21.4	-11.4			
7	N	5	2	2		1.950	1.950			-26.0	-13.4			
8	O	6	2	2		2.275	2.275			-32.3	-14.8			
9	F	7	2	2		2.425	2.425			-40.0	-18.1			
10	Ne													
11	Na	1	3	3		0.733	0.733			-5.1	-3.0			
12	Mg	2	3	3		0.950	0.950			-9.0	-4.5			
13	Al	3	3	3		1.167	1.167			-12.3	-6.5			
14	Si	4	3	3	3	1.383	1.383	1.383		-17.3	-9.2	-6.0		
15	P	5	3	3	3	1.600	1.600	1.400		-18.6	-14.0	-7.0		
16	S	6	3	3	3	1.817	1.817	1.500		-20.0	-13.3	-8.0		
17	Cl	7	3	3	3	2.033	2.033	2.033		-30.0	-15.0	-9.0		
18	Ar													
19–21														
22	Ti	4	4	4	3	1.175	0.800	4.55	1.60				0.4391	0.7397
23	V	5	4	4	3	1.30	0.875	4.70	1.70	-8.81	-5.52	-11.00	0.4755	0.7052
24	Cr	6	4	4	3	1.70	1.70	4.95	1.80	-8.66	-5.24	-11.20	0.5060	0.6750
25	Mn	7	4	4	3	1.80	1.80	5.15	1.90	-9.75	-5.89	-11.67	0.5320	0.6490
26	Fe	8	4	4	3	1.90	1.90	5.35	2.00	-9.10	-5.32	-12.60	0.5505	0.6260
27	Co	9	4	4	3	2.00	2.00	5.55	2.10	-9.21	-5.29	-13.18	0.5679	0.6059
28	Ni	10	4	4	3	2.10	2.10	5.79	2.00	-8.86	-4.90	-12.99	0.5683	0.6292
29	Cu	11	4	4	3	2.2	2.2	5.95	2.30	-11.4	-6.06	-14.0	0.5933	0.5744
30–36														
37–40														
41	Nb	5	5	5	4	1.89	1.85	4.08	1.64	-10.10	-6.86	-12.10	0.6401	0.5516
42	Mo	6	5	5	4	1.96	1.92	4.54	1.90	-8.34	-5.24	-10.50	0.6097	0.6097
43														
44	Ru	8	5	5	4	2.08	2.040	5.378	2.303	-7.73	-4.44	-11.23	0.5340	0.6365
45	Rh	9	5	5	4	2.135	2.10	4.29	1.97	-8.09	-4.57	-12.50	0.5807	0.5685
46	Pd	10	5	5	4	2.19	2.152	5.983	2.613	-7.32	-3.75	-12.02	0.5535	0.6701
47–54														
55–72														
73	Ta	5	6	6	5	2.28	2.241	4.762	1.938	-10.10	-6.86	-2.10	0.6815	0.5774
74	W	6	6	6	5	2.341	2.309	4.982	2.068	-8.26	-5.17	-10.37	0.6940	0.5631
75	Re	7	6	6	5	2.398	2.372	5.343	2.277	-9.36	-5.96	-12.66	0.6662	0.5910
76	Os	8	6	6	5	2.450	2.286	5.650	2.417	-10.36	-5.23	-12.42	0.6680	0.5885
77	Ir	9	6	6	5	2.50	2.20	5.796	2.557	-11.36	-4.50	-12.17	0.6698	0.5860
78	Pt	10	6	6	5	2.554	2.554	6.013	2.696	-9.077	-5.475	-12.59	0.6334	0.5513
79	Au	11	6	6	5	2.602	2.584	6.153	2.794	-10.92	-5.55	-15.07	0.6851	0.5696
80–														

All parameters are from Hoffmann *et. al.* [117] except those for *V* (Hoffmann *et. al.* [118, 119]) and *Ru* ([120])

Table 5–3: The atomic orbital parameters used by CCCP.

The experiments were conducted in three stages:

1. To prove that (for all models) initial starting conditions could be chosen which would force CCCP to fail, *i.e.* predict a hydrogen atom location which is unacceptable.

2. An investigation of the area around the known hydrogen atom location. This was undertaken to establish how close to the true position the initial starting position needs to be for CCCP to successfully locate the hydride ligand to a given tolerance. Moreover, this investigation was designed to provide an estimate of the effect of the initial starting point on the speed and efficiency of CCCP.

3. Assuming no knowledge of the true position, attempts were made at finding ultimately successful initial starting positions from the available atom coordinates and chemical knowledge.

5.3.1 Stage 1: Failure of CCCP

The aim of this section is to show that the program is fallible in that it depends on the starting positions given by the user. Moreover, that there are regions of most models from which a starting position therein will cause the program to find a local minimum and fail.

The experiments performed to show this are listed below.

1.1 Arbitrary coordinates were chosen as an initial estimate at starting points which may cause CCCP to fail.

1.2 – *1.4* For those models for which CCCP successfully located the hydrogen atom in stage *1.1*, the initial starting point was specifically calculated to be in regions of the model where it appeared that there may theoretically be local minima which would cause the program to fail.

Results

Table 5–4 shows the results of stage 1, the series of experiments designed to establish the failure of CCCP. The iteration counts are not published here, since execution time is not an issue at this stage.

Model		nH	Experiment	Distance Error (Å)			Success
$MnH(CO)_5$	(5)	1	1.1	3.1605			fail
$Os_3H_2(CO)_{10}$	(76)	2	1.1	0.1720	0.1770		pass
			1.2	0.2108	9.7168		fail
			1.3	0.2402	4.4244		fail
			1.4	0.2205	4.9075		fail
$Rh_3H_3(PH_3)_6$	(86)	3	1.1	0.1430	8.2087	10.3887	fail
$Os_3H_2(CO)_{10}(\mu_2\text{–}CH_2)$	(77)	2	1.1	0.1361	0.2308		pass
			1.2	0.2830	3.2198		fail
			1.3	0.2055	3.0101		fail
$MoH_2(\eta^5\text{–}C_5H_5)_2$	(6)	2	1.1	0.4719	9.0845		fail
$[Ru_6H(CO)_{18}]^-$	(162)	1	1.1	29.1800			fail
			1.2	3.6633			fail
$[Os_6H(CO)_{18}]^-$	(127)	1	1.1	16.5773			fail
$[Co_6H(CO)_{15}]^-$	(161)	1	1.1	14.8679			fail
$TaH_3(\eta^5\text{–}C_5H_5)_2$	(14)	3	1.1	0.1538	0.2201	0.2248	pass
			1.2	0.5191	4.7867	200.3642	fail
$Os_3H_2(CO)_9(PH_3)$	(111)	2	1.1	no	no		fail
$Os_3H_2(CO)_{10}(PH_3)$	(174)	2	1.1	0.1450	2.4685		fail
			1.2	0.1490	5.2651		fail
			1.3	0.4961	4.3295		fail
$FeH_2\{(CCH_3)_2B_4H_4\}_2$	(20)	2	1.1	no	no		fail
$Fe_3H(CO)_9\{S(C_6H_{11})\}$	(100)	1	1.1	15.4088			fail

Table 5–4: *Stage 1 — Attempts to establish that the choice of starting point(s) can cause CCCP to fail.*

Discussion

It can be seen for these models that starting with the arbitrary coordinates of *1.1* CCCP was only successful for three of the eleven case studies. Moreover, for those three models for which experiment *1.1* succeeded it was not a difficult task to find starting points within the space of the molecule which would cause CCCP to fail to find a reasonable hydride ligand position.

It has been shown that for all the models in stage 1, user provided starting points can cause the program to fail. The low success rate when using arbitrary coordinates shows that this method is not satisfactory and that some means of making an "initial estimate" is necessary to produce a reasonable success rate. The required accuracy of the starting point to ensure settling on a *global* rather than *local* minimum is investigated in the next two sections.

5.3.2 Stage 2: Investigation of the area around the known "correct" position

The true positions of the hydrogens in most of the models are already known. For these models it is possible to try different starting positions at and around the true position and compare the effects on the success, accuracy and efficiency of the program. The methods of calculating the starting points taken for the stage two experiments are listed below

2.1 Coordinates taken from the literature. Such coordinates are usually published to 4–6 significant figures. This starting point is, of course, the ideal and is adopted as a standard. However, this information would not be available in the normal use of this tool — that is to find these positions.

2.2 Literature values with each coordinate rounded to one decimal place in Ångström space (approximately 2 significant figures).

2.3 Literature values with each coordinate rounded to the nearest integer in Ångström space (approximately 1 significant figure).

Results

The results of the experiments performed in stage 2 are shown in Table 5–5. Models 20, $FeH_2\{(CCH_3)_2B_4H_4\}_2$, and 111, $Os_3H_2(CO)_9(PH_3)$, were not appropriate to this stage of the evaluation as the true positions were un-

known. It was decided to include models 100, $Fe_3H(CO)_9\{S(C_6H_{11})\}$, and 174, $Os_3H_2(CO)_{10}(PH_3)$, (the hydride ligands in these models were located by X-ray diffraction), despite the published hydride atom coordinates not being as accurate as those located by neutron diffraction.

The iteration count before convergence (denoted "#" in the table) is also given since speed and efficiency are being considered.

Model		nH	Experiment	Accuracy (Å)			#	%Δ	Success
$MnH(CO)_5$	(5)	1	2.1	0.0237			43	—	pass
			2.2	0.0386			45	4.65	pass
			2.3	0.0371			51	18.60	pass
$Os_3H_2(CO)_{10}$	(76)	2	2.1	0.1481	0.1594		115	—	pass
			2.2	0.1653	0.1749		114	-0.87	pass
			2.3	0.1582	0.1667		168	46.09	pass
$Rh_3H_3(PH_3)_6$	(86)	3	2.1	0.1125	0.1446	0.2646	337	—	pass
			2.2	0.1163	0.1512	0.2523	228	-32.34	pass
			2.3	0.1253	0.1556	0.2651	356	5.64	pass
$Os_3H_2(CO)_{10}(\mu_2\text{–}CH_2)$	(77)	2	2.1	0.1330	0.2352		129	—	pass
			2.2	0.1208	0.2311		180	39.53	pass
			2.3	0.1305	0.2228		168	30.23	pass
$MoH_2(\eta^5\text{–}C_5H_5)_2$	(6)	2	2.1	0.1337	0.1347		115	—	pass
			2.2	0.1360	0.1410		79	-31.30	pass
			2.3	0.1338	0.1414		242	110.43	pass
$[Ru_6H(CO)_{18}]^-$	(162)	1	2.1	0.0250			36	—	pass
			2.2	0.0160			35	-2.78	pass
			2.3	0.0193			47	30.56	pass
$[Os_6H(CO)_{18}]^-$	(127)	1	2.1	0.7430			65	—	pass
			2.2	0.7568			61	-6.15	pass
			2.3	0.7407			68	4.62	pass
$[Co_6H(CO)_{15}]^-$	(161)	1	2.1	0.1503			48	—	pass
			2.2	0.1521			59	22.92	pass
			2.3	0.1476			58	20.83	pass
$TaH_3(\eta^5\text{–}C_5H_5)_2$	(14)	3	2.1	0.1457	0.2196	0.2259	225	—	pass
			2.2	0.1516	0.2103	0.2313	179	-20.44	pass
			2.3	0.1438	0.2120	0.2343	948	321.33	pass
$Os_3H_2(CO)_{10}(PH_3)$	(174)	2	2.1	0.1071	0.3531		172	—	pass
			2.2	0.1116	0.3573		182	5.81	pass
			2.3	0.1186	0.3513		215	25.00	pass
$Fe_3H(CO)_9\{S(C_6H_{11})\}$	(100)	1	2.1	0.3206			53	—	pass
			2.2	0.3223			64	20.75	pass
			2.3	0.3257			53	0.00	pass

Table 5–5: Stage 2 — The aims of this stage were to establish how inaccurate the starting point may be without causing CCCP to fail, and the effects of the starting point on accuracy of results and rate of convergence.

	Success	Accuracy	%Δ Efficiency
2.1	100%	0.199 ±0.036	
2.2	100%	0.202 ±0.036	-0.02 ±6.79
2.3	100%	0.202 ±0.036	55.76 ±28.07

Table 5–6: The success, accuracy and efficiency of CCCP averaged over all methods.

Summing the results in terms of the afore-mentioned criteria:

Success:

All experiments were successful for all models.

Accuracy:

Table 5–6 shows that there was a slight decrease in the accuracy of experiments 2.2 and 2.3 over 2.1, but with no difference between the former two.

Efficiency:

Table 5–6 also gives a measure of the average efficiency of the program. This measure is calculated as the percentage change in the iteration count needed for the program to converge in comparison to the count required for convergence in experiment 2.1. For example, the value for stage 2.2, $MnH(CO)_5$, would be a $100 \times (45 - 43)/43 = 4.65\%$ increase in iteration count, whereas for stage 2.2, $MoH_2(\eta^5\text{–}C_5H_5)_2$, the value would be $100 \times (79 - 115)/115 = -31.30\%$, a negative number indicating a decrease.

It can be seen that the efficiency of the program actually improves marginally for stage 2.2 with an average -0.02% decrease in iteration count as the accuracy in the starting position decreases. However, the efficiency degrades significantly for stage 2.3 with an average 55.76% increase in iteration count, as may be expected.

Discussion

It was expected that this investigation would show that the closer the starting point(s) to the true position(s) the more accurate the final result and the less iterations the program would take to reach that result. Unfortunately, the experiment described above did not include enough molecular models to illustrate the trends and the results were inconclusive. Having said that, Table 5–6 does show two characteristics:

- that starting points accurate to 1 or 2 significant figures (experiments *2.2* and *2.3*) produce less correct results than those to in experiment *2.1*, which are accurate to 4 or 5 significant figures and

- starting points accurate to 1 significant figure cause CCCP to take longer to converge than starting points accurate to 4 or 5 significant figures.

5.3.3 Stage 3: Estimates using Chemical Knowledge

Stage 3 is similar to stage 2 in that an investigation is made regarding the required accuracy of the starting point to the true position for CCCP to successfully locate the hydrogen atom and the effects of the choice of starting point on accuracy and efficiency in finding the final result. However, it differs in that in the selection of a starting point any pre-knowledge of the true location is ignored. General knowledge of chemistry is applied instead. The purpose of this stage is to establish how closely the chosen molecular models follow chemical rules and whether the combination of knowledge of these rules and the CCCP program can produce a reasonably successful, accurate and efficient estimate at hydrogen atom location. The methods investigated are listed below:

3.1 The centre (of volume) of the molecule.

3.2 The centre (of volume) of the transition metal atoms in the molecule.

3.3 A coarse estimate was made of the position of the individual hydrogen(s) by applying chemical knowledge, but not knowledge of the true hydrogen atom position(s).

3.4 On the failure of *3.3* a more refined estimate of the hydride ligand location was made.

Sometimes, when the coordinates of the estimate coincide with an already existing atom or when there are two or more atoms, it is necessary to adjust the starting point coordinates[1]. This is done by (a) truncating the coordinates for the first coincident atom and (b) rounding up the estimate by 1 dp (usually to 5 dp) for the second.

Results

Table 5–7 shows the results from the stage 3 experiments

Discussion

1. The success, accuracy and efficiency of CCCP are all highly dependent on the user defined starting point to the extent that no two (different) starting positions (however close) produce identical final positions.

2. Within each model investigated there are regions where, if the user defined starting point falls within, CCCP will fail to produce a reasonable location for the hydrogen atom and, therefore, fail.

3. The closer the original starting point is to the true position;

   (a) the more likely the program is to succeed,

---

[1] An initial function of CCCP is to calculate all atom-atom distances. The reciprocal distance is applied in later calculations, resulting in a division by zero if any two atoms are coincident

Model		nH	Experiment	Accuracy (Å)			#	%Δ	Success
$MnH(CO)_5$	(5)	1	3.1	0.0397			105	—	pass
			3.2	0.0330			105	0.00	pass
			3.3	0.0406			43	-59.05	pass
$Os_3H_2(CO)_{10}$	(76)	2	3.1	0.1603	0.1716		179	—	pass
			3.2	0.1737	0.1795		165	-7.82	pass
			3.3	0.1629	0.1664		481	168.72	pass
$Rh_3H_3(PH_3)_6$	(86)	3	3.1	0.1115	0.1139	0.2953	495	—	pass
			3.2	0.1210	0.1765	0.2067	760	53.54	pass
			3.3	0.1566	0.1839	0.2112	889	79.60	pass
			3.4	0.1447	0.1961	0.2381	603	21.82	pass
$Os_3H_2(CO)_{10}(\mu_2\text{-}CH_2)$	(77)	2	3.1	1.5715	2.1566		364	—	fail
			3.2	1.5690	2.1493		203	—	fail
			3.3	1.6016	3.7348		297	—	fail
			3.4	0.1335	0.2162		187	—	pass
$MoH_2(\eta^5\text{-}C_5H_5)_2$	(6)	2	3.1	0.1462	0.1478		440	—	pass
			3.2	6.9364	8.5415		41	—	fail
			3.3	0.1105	0.1352		174	-60.45	pass
$[Ru_6H(CO)_{18}]^-$	(162)	1	3.1	0.0250			36	—	pass
			3.2	0.0308			34	-5.56	pass
$[Os_6H(CO)_{18}]^-$	(127)	1	3.1	0.7437			84	—	pass
			3.2	0.7250			50	-40.48	pass
$[Co_6H(CO)_{15}]^-$	(161)	1	3.1	0.1535			43	—	pass
			3.2	0.1558			48	11.63	pass
$TaH_3(\eta^5\text{-}C_5H_5)_2$	(14)	3	3.1	0.2819	8.2979	12.8868	253	—	fail
			3.2	0.6294	1.1270	4.4818	683	—	fail
			3.3	0.1780	0.1817	0.3091	938	—	pass
$Os_3H_2(CO)_9(PH_3)$	(111)	2	3.1	yes	yes		978	—	pass
			3.2	yes	yes		154	-84.25	pass
$Os_3H_2(CO)_{10}(PH_3)$	(174)	2	3.1	0.0844	0.3116		406	—	pass
			3.2	0.1112	0.3525		331	-18.47	pass
			3.3	0.1113	0.3537		482	18.72	pass
$FeH_2\{(CCH_3)_2B_4H_4\}_2$	(20)	2	3.1	no	no		45	—	fail
			3.2	yes	no		168	—	fail
			3.3	yes	no		553	—	fail
$Fe_3H(CO)_9\{S(C_6H_{11})\}$	(100)	1	3.1	3.1511			57	—	fail
			3.2	0.3240			131	—	pass

Table 5–7: *Stage 3: An attempt to show that, with the aid of chemical knowledge, starting points could be calculated which, when subsequently optimised, produce a correct hydride location.*

(b) the more accurate the result will be and

(c) the shorter the time (or the least amount of iterations) CCCP will take to find a (correct) result.

4. If the true hydride ligand position is unknown, a combination of chemical knowledge and CCCP may be applied for most models to find reasonable hydride atom positions.

Summary: the initial starting point need not be accurate for most models if CCCP is to succeed, although it is more likely to succeed if an initial estimate is made, and likely to succeed with greater accuracy and efficiency if the initial estimate is a good one.

5.3.4 User Interface

Input

o *Description*

```
@ANGLE <-------- "Print angles"
@OPT <-------- "Optimisation required"   Output control for printer   Output control for another program
@SIMPLEX <-------- "Use Simplex Method"
MnCO5 [LaP1aca] <-------- Title card
  1 11   0   0 -2   0FFFFT 1.75 0.750-13.60TTTTFFFTFFFFFFFFFFFFFTFFFFFFFTFFFFFFFFFFF
atom   H       0.44131 0.97488 3.75710
atom   C(1)    1.50751 2.36010 3.75710
atom   C(2)   -0.53492 0.74237 1.64536
atom   C(3)    2.29895 2.29759 3.42938
atom   C(4)   -0.14224 2.99873 2.88565
atom   C(5)    1.87101 0.03183 2.29446
atom   O(1)    1.81631 2.84593-0.21178   } Atom Coordinate cards
atom   O(2)   -1.42606 0.18836 1.32365
atom   O(3)    3.08553 2.72043 4.11188
atom   O(4)   -0.82305 3.82467 3.25394
atom   O(5)    2.41809-0.96797 2.26333
atom   Mn      0.95435 1.62951 2.37690
 C C C C C O O O O O * <-------- Heavy atom card
MN  7   4 1.80 -9.75    4 1.80 -5.89    3 5.15 -11.67 .5320 1.90   .6490 <-- Atom definition card
3 <-------- Number of coordinates to be optimised
1 2 3 <-------- Those coordinates to be optimised
0.302335 0.801626 3.921696 <-- Starting Point
00 <-------- Iterated atoms card
END
```

Figure 5–1: An example input file of CCCP.

Figure 5–1 shows an example of a CCCP input file. Because the program has been developed over a period of years the different styles are evident in the input file. A major difference between the two parts is that whilst the original part is in fixed format, the part added later is in free format.

The original input file (that section shown in normal type) originates from ICON8 [121], a program which calculates the energy of the molecule by the

EHMO method amongst other quantities. This part of the input file must contain:

Title Card:

The user may enter a title here to identify, for example, the model and/or experiment. The line may be left blank but may not be omitted.

Data Card:

This line allows the user to enter several data about the model, the experiment and output control. Looking at the entries in order:

- `1` and `11` are the numbers of hydrogen and non-hydrogen atoms in the model respectively, `0` is the molecular charge.
- `0` tells the program which method to use (from a range of 0→3) and `-2` and `0` (from a range of -4→2) control the output of the program.
- `FFFFT` sets various experiment and output control parameters.
- `1.75` is an constant in the EHMO calculations, `0.750` and `-13.60` are the hydrogen orbital exponent and energy respectively.
- The remaining "T"s and "F"s control both output to hardcopy and output to another program (termed "punching" in the CCCP manual).

Atom Coordinate list:

Next is a list of hydrogen atoms and non-hydrogen atoms. Hydrogen atoms must be listed *first* for this program. Each atom is specified by a label and position coordinates. A combined maximum of 50 atoms is allowed, of which no more than 40 may be non-hydrogen.

Heavy Atom List:

Following: a list of element types to identify the non-hydrogen atoms in sequence. The atomic orbital information is internally stored for some of the more common atoms so these may be identified by atomic symbol. The others, however, must be defined by the user identifying that atom as

"*" in the atom list and entering appropriate data on the following lines. This must be done for each *new* atom type. However, once one has been defined subsequent atoms of that type must be referred by the previously defined name. In the example in Figure 5–1 manganese is unknown to CCCP, thus has to be defined by the user on the next line.

That part of the input file shown in italicised type is added for CCCP. This specifies further experimental and output control in a selection of optional "@" statements. There is also a facility which allows the user to specify the positions to be optimised and from which point the optimisation method should start. The 00 is charge iteration data.

o *Discussion*

It would seem that there are five main problems with the format of the input file as it stands:

1. *Obscurity:*

   It is not at all obvious what the strings of "T"s and "F"s, or the various numbers in the data card refer to. Without checking the manual even the most experienced user would have difficulty in successfully controlling the experiment, and the output. To further complicate matters, whilst "T" (for true) will suppress the action for 20 of the Boolean values, it will enable the action for the remaining twenty-five — a clear case of inconsistency. In practice a template of the input file is typically obtained from another user. This "card" is barely modified. Only the necessary changes to the parameters defining the model are made, without the user ever understanding the function of these switches.

2. *Lack of logical structure:*

   Output control, experimental control, model data and atomic parameters are intermingled with nothing to separate or distinguish them.

3. *Inconsistency:*

   Some of the output control and experimental control parameters are set in the data card and others by the "@" statements. A further complication is that there are two different methods of controlling output and "punching". The user has a choice of explicitly enabling or repressing the individual output resources as well as setting defaults. It is unclear which of the two has priority.

4. *Superfluity:*

   Much of the data which must be included is not specific to either the experiment or the model and is simply copied by the user from a table held elsewhere. There is no need for this data to be user-defined as it may easily be stored internally or read from a table by the program.

5. *Inflexibility:*

   Most of the input file is in fixed format, that is data may only be entered in certain positions on certain lines. Failure to comply with this excessively rigid structure could cause the program to crash (to fail completely with a run-time error reported to the user) or, much more seriously, fail (to read the data incorrectly and therefore produce incorrect results, with no warning to the user). Paradoxically, complying with the rigid structure means that data items may run into each other, making them indistinguishable to the user (counter-intuitiveness). In addition restrictions may be placed on the position of a datum within its allotted space. Whilst floating-point numbers and text entries may appear anywhere in their slots, standard Fortran formatting requires that integers must be right-justified, and CCCP itself requires that entries on the heavy atom card must also be right-justified — a further manifestation of inconsistency.

The user has so much to remember in setting up an input file, so many rules and regulations to follow, that even the more experienced user may find it difficult to create a file for the simplest model, and may not be confident that, in fact, the

file is correctly created. The program generally warns the user if an error in the input file is detected, but since not all errors are detected this is not satisfactory.

Output

∘ *Description*

The method of output for CCCP is to print whichever items have been requested by the user in the output control section once in every iteration. Some of the items which may be output are large matrices. The user may be left with an output file which is pages long if the printing of a few of these matrices is selected and there are several iterations before the solution is found. It can be difficult for the user to locate useful information amongst such verbose output data.

∘ *Discussion*

As discussed in Chapter 4 rows and columns of numbers are not the best way to convey information. The purpose of this program is to locate hydride ligands and then convey this information to the user. All other information produced during the optimisation should be incidental. Currently the the program generates superfluous information, some of which may be suppressed by the user, the rest of which may not. Unless the output file is edited manually, a user requiring a hard copy of the location must print out sheets of paper for at most a few lines of data. Although the information other than the final location may occasionally be of some use to some of the chemists using the tool, it should take second place to the final located hydride ligand coordinates, perhaps being presented only on specific request from the user.

Having established that focus should be concentrated on the located hydride ligand position(s), the presentation of that data should now be considered. The current method does not show the spatial relationship(s) between the located hydride ligand(s) and the rest of the model, giving only the optimised coordinates. Unless the refined coordinates are obviously incorrect, the user will often

wish to verify and interpret this output. The optimised coordinates must then be appended to the original model and the updated model fed to program which indicates the bonding interactions of the located hydride ligands — perhaps by providing visualisation or calculating bond lengths and angles. Although this process may work it is unsatisfactory because of the unnecessary amount of work the user is obliged to perform before the output may be fully interpreted. The time spent in undertaking all these operations could be spent more productively and each stage, however simple, may be open to error.

5.4 Conclusions

There is clearly room for improvement of this program in two distinct areas; the human-computer interaction properties of the program and its performance.

HCI Properties

As it stands, the interface makes it difficult to set up the input file either:

- correctly so that the located hydride ligand position(s) can be trusted,
- to instruct CCCP to perform the exact experiment the user wishes,
- to control the output to obtain the exact information required.

Assuming the user has circumvented the problems of setting up the input file, the output remains difficult to interpret and several other operations must be performed before the information sought may be fully understood.

Performance

The program depends highly on the starting position provided by the user, but gives no clues or recommendations, either visual or textual, as to what that position should be. The user must either calculate likely starting positions using

visualisation of the molecule, geometrical and chemical knowledge, or give entirely random coordinates. Each method would seem to be unsatisfactory, one requiring unnecessary preparation from the user, the other laying the program open to failure by locating local rather than global minima.

Chapter 6

The Locator

6.1 Introduction

CCCP optimises the position of hydride ligands more accurately, more efficiently and with greater success from appropriate initial position(s) than from user supplied or random estimates, as shown in the previous chapter. The initial position(s) for energy optimisation should be reasonably accurate. An accurate initial estimate may prevent settling at a local rather than global energy minimum and ensure an acceptable rate of convergence. However, some discrepancy between the initial estimate(s) and true position(s) may be tolerated, since the program may succeed even when the topology of the initial estimate is quite incorrect. Indeed, the exact position(s) of the hydride ligand(s) is/are unknown *a priori*, this being the very information sought. However, sufficient evidence may be deduced from known constraints on the hydrogen ligands themselves and the molecule of which they are part to form reasonably accurate initial estimates. Given knowledge of properties of the transition metals themselves, specifically the number of valence electrons each has and the number received by each of the ligands bonded thereto, the metal atoms to which each of the hydride ligands are likely to be bonded may be determined.

Once the respective M–H bonds are determined, the knowledge of typical M–H bond lengths ties the location of the hydride ligand to certain regions in the vicinity of the hydride-bound metal atoms, depending on the site type. Examples of the different site types are shown in Figure 6–1. A terminally bound (τ) hydride ligand is constrained to the surface of a sphere surrounding the metal atom to which it is bonded. An edge-bridging (μ_2) hydride ligand is constrained to the circumference of the circle defined by intersection of the two spheres surrounding each of the two metal atoms to which the hydride ligand is bonded. Similarly a face-bridging (μ_3) hydride ligand is constrained to intersection of three spheres (two points — one above and one below the plane of the hydride-bound metal atoms) and an interstitial (ι) to the centre of the hydride-bound metal atoms. With the above knowledge, and the knowledge of the spatial distribution of the other atoms in the molecule, an initial estimate of the location of the hydride ligand(s) can be made.

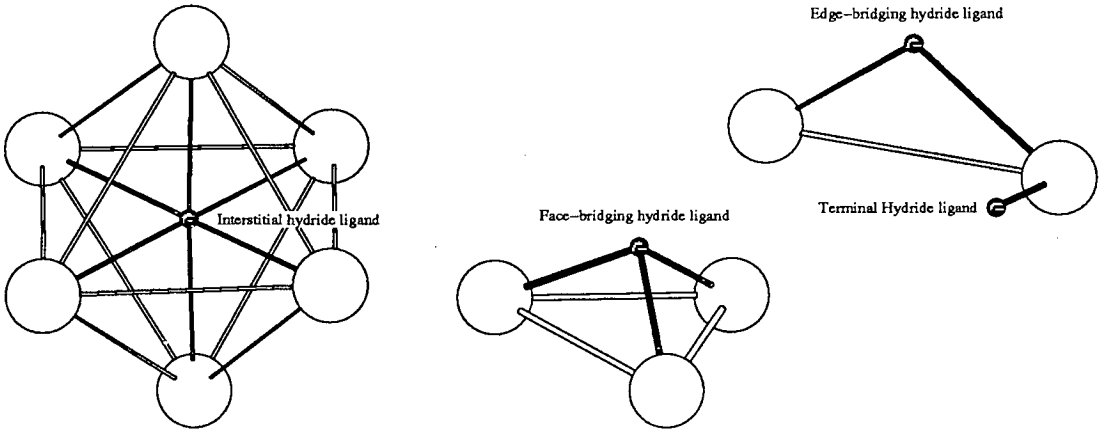


Figure 6–1: Examples of possible hydride ligand site types.

This chapter describes the implementation of a program — *the Locator* — which exploits the chemical knowledge and heuristics described above to *estimate* hydride ligand location(s) in transition metal clusters for subsequent potential energy optimisation by CCCP. The program allows the user full experimental control through a direct manipulation interface almost identical to that described

in Chapter 4. The whole molecule may be observed and manipulated with the updated position of the hydride ligand at any stage of the optimisation.

It should be noted that when the implementation described in this chapter was embarked upon, the author was under the arguably justified impression that the hydride ligand location program, CCCP, did in fact work. Problems discussed in this and the previous chapters arose during the course of implementation.

6.2 Objectives

Being part of the tool-kit, all of the objectives outlined in Chapter 3 apply directly to the Locator. The objectives outlined here are in addition to these. The topic of hydrogen location was not specifically raised in any of the responses to the questionnaire. However, several requirements have become apparent through personal use of CCCP. A list of design objectives have been constructed from these requirements. The tool may be evaluated in the light of these objectives. The additional design objectives to those detailed in Chapter 3 are:

1. *Consistency:*

   Results from the tool should not be dependent on an initial point given by the user.

2. *Constraints:*

   The program should not be restricted to molecules with an unreasonably limited atom count. CCCP is restricted to atom counts of no more than fifty atoms, of which no more than forty may be non-hydrogen atoms.

3. *Success:*

   The program should converge to reasonable hydride ligand locations within transition metal clusters for a high proportion of the experiments performed.

4. *Accuracy:*

   These positions should be located to a reasonable degree of accuracy.

5. *Speed and efficiency:*

   The program should take the minimum possible time to execute.

These objectives should ensure that hydride ligands may be located in transition metal clusters with the minimum of necessary effort. The user should thereby be able to concentrate on the work in hand.

6.3 Principles of the Initial Estimate Method

6.3.1 Rationalisation of the model

The metal atom(s) to which the hydrogen atom(s) are bonded must be ascertained from the information provided by the user — the number of hydrogen atoms to be located and a list of atomic coordinates. This procedure will be henceforth be referred to the as the *rationalisation* of the model. Figure 6–2 gives a schematic view of the way in which this information is applied to rationalise the model and establish the bonding of the hydrogen atom(s) to be located. An explanation of the stages in this procedure is given below.

1. Bonding interactions within the molecule are established by comparing the distances between all atom pairs with the sum of their atomic radii, the latter being taken from the atomic data file. Two atoms are deemed to be bonded if the distance between them is less than the sum of their atomic radii. [1]

[1] within a tolerance of 0.3 Å as described in Chapter 4

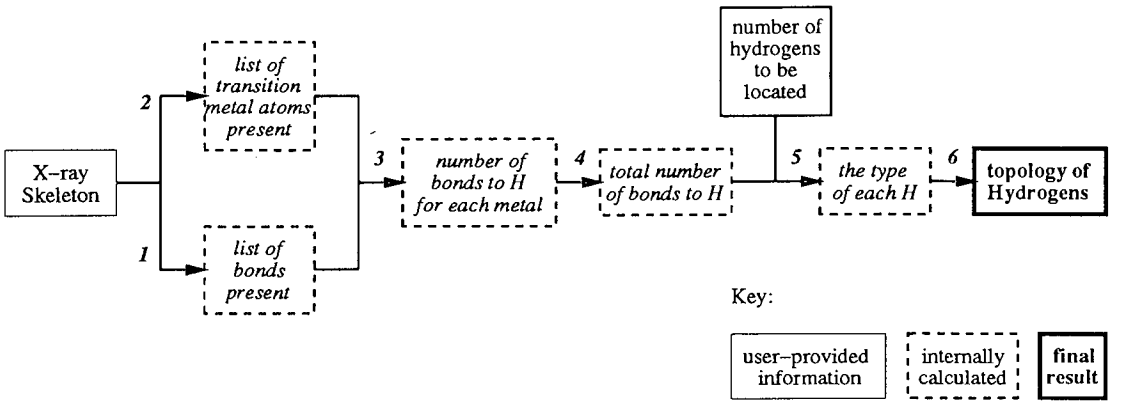


Figure 6–2: The rationalisation of the molecular model.

2. Each atom is considered in turn and checked against an internally stored table of transition metals. If the atom is present in the table a pointer to it is added to a list. Thus a list of all transition metal atoms within the model is maintained.

3. The number of valence electrons of each metal atom in the list of transition metal atoms is counted. The number provided by the metal atom itself as taken from the data file is added to those provided by the groups bonded to the metal atom. For example, two electrons are accumulated for every $-CO$ or $-PH_3$ and one for every $-CR_3$. Generally, chemistry dictates that every transition metal atom should have eighteen electrons [89], except under some circumstances when it may have sixteen. It should be noted that examples of compounds *are* known where transition metal atoms do not have either sixteen or eighteen electrons, but they are rare. The number of hydrogen atoms that will be bonded thereto, each donating one electron, can easily be calculated as the difference between the number of electrons the metal atom requires (eighteen) and the number it already has (the accumulated electron count).

4. By summing the number of hydrogen atoms required by each transition metal atom it is possible to calculate the total number of hydrogen ligands

— each contributing a single electron — with which each transition metal atom in the molecule needs to form a bond.

5. Once the number of available hydrogen atom sites has been established it can be compared with the number of actual hydride ligands to be located in order to determine what *type* (ι, τ, μ_2 or μ_3) each hydride ligand is — in other words, to how many metal atoms it is bonded.

   Under some circumstances the number of hydrogen atoms indicated by the user may fail to satisfy the electron requirements of the metal atoms in the molecule, or may not be accommodated following normal chemical rules. In this case the control would come back to the user, requesting information, specifically the metal atom(s) to which each hydrogen atom is bonded. To aid the user in determining this information, the current electron count is provided for each transition metal atom in the model. Although such deferring to the user may seem to be "cheating", it is considered justifiable as the subject of chemistry is so complex one that *no* software system could ever hope to encapsulate all of the knowledge and "intuition" of a working chemist. For this reason it is considered important that this software system should bow to the superior experience and knowledge of the chemist user if a problem that cannot be solved automatically is ever encountered.

   Returning to the problem in hand, the number of possible combinations of hydride ligands and metal atoms is large, but many may be either identified immediately or eliminated:

   (a) If there is only one hydrogen atom and four or more equivalent metal atoms, the hydrogen atom is likely to be interstitial.

   (b) Otherwise, there are the three other possible types of hydride ligand; terminal, edge-bridging and face-bridging, bonding to one, two or three metal atoms respectively. If the hydrogen atom is known not to be interstitial, any of the combinations which would necessitate

a hydrogen atom bonding to more than three metal atoms can be eliminated.

(c) It is possible to have at most two face-bridging hydrogen atoms across the same three metal atoms, although even two is chemically unlikely.

After eliminating all impossible combinations it is generally possible to find a distribution of the hydrogen atoms throughout the molecule whereby all metal atoms are satisfied.

6. It is relatively simple to match the hydrogen atoms with the transition metal atoms once the type of each hydride ligand, and the number of hydride ligands each metal atom requires is known.

6.3.2 Initial Estimate of Hydride Ligand Position

Three methods were developed for estimating the most likely hydride ligand position(s) within transition metal clusters once the model has been rationalised. One method — the *centre of the metal cage* — is appropriate for transition metal clusters with a potential interstitial hydride, where the others, *projection* and *biggest appropriate hole* are suitable for all others.

Centre of Transition Metal Cage (Interstitials)

Molecules which require only one hydride ligand, but which have four or more *equivalent* — *i.e.* each have the same number of valence electrons — transition metal atoms are assumed to contain an interstitial hydrogen atom. The centre of the group of metal atoms within the molecule is taken as the initial point for CCCP in this case.

Projection Method

In this method, the centre of the molecule is projected through the centre of the group of metal atoms to which the hydrogen atom is bonded. This latter point is calculated as the vector mean of the coordinates of the hydride-bound metal atoms and is referred to subsequently as the *bonding point* or *bonding centre*. The calculation of the former is described in detail later. The heuristic behind this is that projection *away* from the centre will take the hydride ligand *to* an area of the molecule of relatively low atom density.

Figure 6–3 shows examples of this method on some theoretical models. As indicated by this diagram, if the hydrogen atom is thought to be *terminal* the initial position is estimated by projection through the single metal atom to which the hydrogen atom is bonded. However, if the hydrogen is though to be *edge-bridging*, the projection bisects the line connecting the two metal atoms to which the hydrogen atom is bonded. If the arrangement is thought to be *face-bridging*, the projection goes through the centre of the triangle formed by the three metal atoms to which the hydrogen atom is bonded. It is acknowledged that different types of hydride ligands have different bond lengths, with the general trend "terminal < edge-bridging < face-bridging". In accordance with this, the values of 1.7Å, 1.84Å and 1.87Å are applied to calculate the initial estimate of the location of terminal, edge-bridging and face-bridging hydrogen ligands respectively.

At first the centre of the molecule was calculated on an atom centre basis, being taken as the atom centre mean. This proved unsatisfactory for some models examined. It was noted that the outer, distant atoms of the molecule would tend to exert less steric influence on the hydrogen ligand than those nearer the hydrogen ligand. A method of calculating a "molecular centre" on which those distant atoms from the *bonding* centre would have less effect seemed more appropriated in such cases. A possible method would be to take a *weighted* mean of the known atomic centres. A method was proposed whereby such weights were based on the distance from the bonding centre. This method proved more

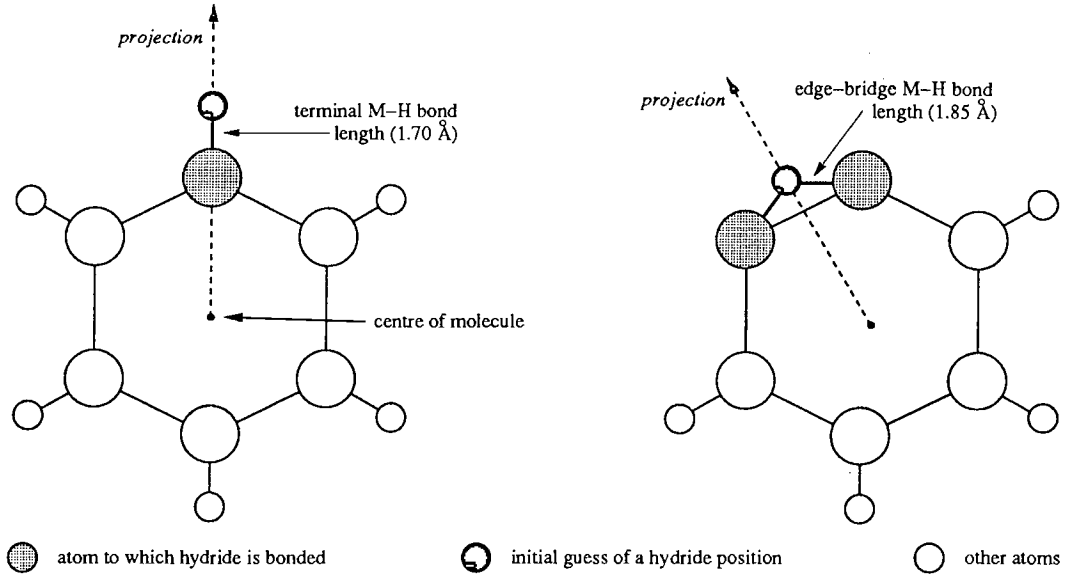


Figure 6–3: The projection method.

successful but still failed for one model. A further modification in considering not the atom centre to atom centre distances but rather the atom edge to atom edge distances remedied this.

Unfortunately the projection method breaks down for those models with equivalent hydrogen atoms — those bonded to the same metal atoms. In these situations it is unclear whether the position of the first hydride ligand should be included when calculating the second. If the first hydrogen atom position is not taken into account an identical position is predicted for the second hydride ligand. If it is taken into account the second hydride ligand is either predicted to be at the same position as the first, or is projected back through the molecule centre if the first hydrogen has "dragged" the weighted molecular centre to the other side of the bonding centre. The first site is meaningful being, in essence, an average between the equivalent hydrogen atom positions, but the second site has no real meaning. Thus, in the calculation of hydride ligand position by the projection method the position(s) of previous hydride ligands are not taken into account and identical positions are predicted. As explained in Chapter 5, CCCP will fail if any of the coordinates exactly coincide. Any duplicated initial estim-

ates are therefore separated by successive addition of a nominal perturbation to all but the first hydride ligand position estimate.

Biggest Appropriate Hole Method

An alternative to the projection method was developed for the prediction of more reasonable ligand positions for models with equivalent hydrogen atoms. The "Biggest Appropriate Hole" method, as its name would indicate, attempts to locate the largest hole within the model whose centre lies one metal–hydrogen (M–H) bond length distance from the metal atom. The M–H bond length is relatively short (less than 2Å). The investigation of spatial relationships between many of the outer atoms of the molecular model which may lie several Ångströms away from the hydride ligand bound metal atoms would therefore be inappropriate to the search for the biggest *appropriate* hole.

There may be many other atoms between these outer atoms and the metal atoms to which the ligand is bonded. Hence the first stage in the biggest appropriate hole method is to identify an *inner core set* of atoms, within which the biggest appropriate hole is to be located. Provided that both the inner core and biggest hole therein can be determined, this observation provides a heuristic approach to finding an appropriate initial point for subsequent optimisation.

∘ *Stage 1: Identification of the inner core of a group of atoms*

There are several possible approaches to selecting an inner core of a set of points relative to a known centre position:

1. One approach would be to take all of those atoms bonded to the same metal atom(s) as the hydride ligand, the *inner coordination sphere.* On first consideration it would seem that such an approach is reasonable, being both simple and returning an appropriate inner core set of atoms. On further consideration, however, it must be realised that this method will fail if incorrect calculations of interactions within the model result in an un-

derestimation of the number of bonds. Bonding interactions could be corrected by the user before proceeding with the rationalisation of the model. However, any such *obligation* on the user is unsatisfactory. For example, in the calculation of bonding interactions of $Re_3H_3(CO)_8\{(PH_2)_2O\}_2$ and $Mn_3H_3(CO)_{12}$ (illustrated in Appendix D as models 83 and 105 respectively) the metals are too separated to be considered bonded with respect to tabulated atomic radii. Thus, although the research chemist would concede that they are in fact bonded, they would be rejected from the automatically calculated coordination sphere and so would excluded from the inner core set. This method of selecting the inner core set of atoms is therefore considered unsatisfactory.

2. Another approach would be to take the nearest points — either those within some tolerance distance of that centre, or a certain number of nearest points. However, either of these criteria may result in an inner core set which *does not* contain the bonding centre — in that that centre is outside the *convex hull* of the set[2] — due to local clusters of atoms. Intuitively, the centre point should be within the convex hull of the inner core set. Figure 6–4 shows such an example where the bonding centre actually lies outside of the inner core if judged by either of the above criteria. Atoms are named in order of increasing distance from the bonding centre in this diagram. If the six nearest atoms were selected, atoms *A–F* would constitute the inner core. Similarly, if those atoms within the area shown by a dashed circle and line in the diagram were selected, atoms *A–E* would constitute the inner core. In either case, the inner core would not contain the bonding centre since atom *G* would be rejected. Again, such methods are rejected as unsatisfactory.

---

[2]the *convex hull* of a set of points is the intersection of all the convex sets containing these points. A set is convex if given any two points in the set the line segment between them lies entirely within the set; more intuitively the convex hull would be the shape assumed by cling-film wrapped around the points.

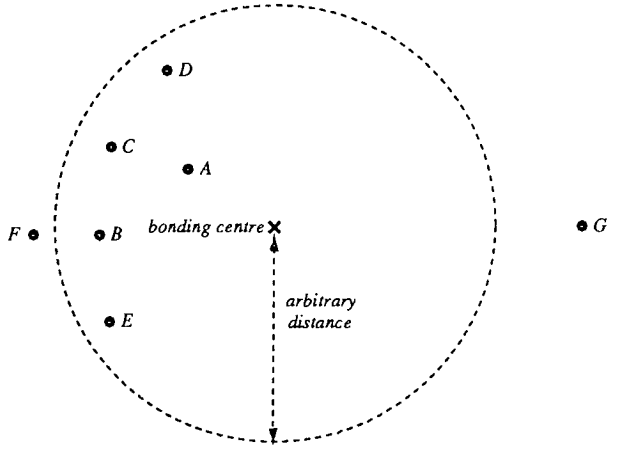


Figure 6–4: A situation in which selection of a certain number of the nearest atoms, or those within a certain distance would be inappropriate.

3. A third method would be to select an inner core from a given set of points as comprising the inner points in much the same way as the convex hull represents the "outer" points — perhaps thought of in terms of a "balloon" being inflated from the bonding centre whilst retaining a convex shape. This method neither relies on the success of any previous bonding calculations nor results in an inner core set which does not contain the bonding centre. It was decided to implement an inner core atom set selection method with these desirable properties.

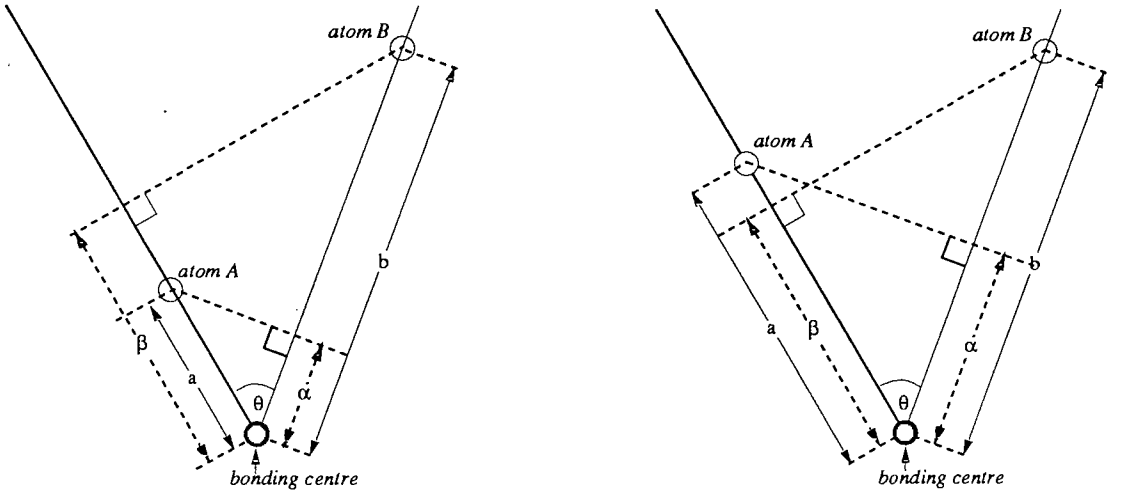
One method of defining the inner core might be as the set of points corresponding to the convex hull in an "inverted" space. The mapping to the inverted space would transform each point to lie in the same direction relative to the centre, but at the reciprocal length; in vector notation

$$T(r\underline{u}) = \frac{1}{r}\underline{u}$$

where r is scalar length and $\underline{u}$ is a unit direction vector. However this would require the transformation of every point and the subsequent application of an algorithm to determine the convex hull [122]. Such algorithms tend to incur a high computational load in finding the *precise* convex hull. The required

properties of an inner core set are perhaps more heuristic, allowing a more efficient determination.

An alternative selection of the inner core is to start with the entire set and discard any points distant from the centre to leave a set of nearby points. The rejection criteria should ensure that the centre lies within the remaining set. To this end, the *direction* of each point relative to the centre should be taken into account as well as the distance. Points may be considered in pairs, represented by the centre-relative vectors $\underline{a}$, $\underline{b}$ with respective lengths a, b and mutually subtending an angle θ about that centre. The "dot-product" of these vectors is denoted as $\underline{a} \bullet \underline{b}$ and is defined as the sum of their component-wise products; this is known to evaluate to $ab\cos\theta$. If one point is clearly "further" than the other from the centre it may be rejected from the inner core. Here, "further' refers not to absolute distance, but rather to distance in *specific* directions — along the two vectors in question. Let $\alpha = a\cos\theta$ be the distance of $\underline{a}$ along $\underline{b}$ — that is, the length of the perpendicular projection of $\underline{a}$ onto $\underline{b}$. Similarly, let $\beta = b\cos\theta$ be the distance of $\underline{b}$ along $\underline{a}$ — that is the length of the perpendicular projection of $\underline{b}$ onto $\underline{a}$, as shown in Figure 6–5.



(a) atom B is rejected. *(b) neither atom is rejected.*

Figure 6–5: The elimination procedure to establish the inner core of atoms.

An atom is rejected from the inner core if it lies further from the centre along *both* vector directions, that is:

$$\begin{array}{ll} \underline{b} & \text{is rejected if both} \quad a < \beta \text{ whilst } \alpha < b \quad \text{and} \\ \underline{a} & \text{is rejected if both} \quad b < \alpha \text{ whilst } \beta < a \end{array}$$

Multiplying these conditions through selectively by a (which is > 0) or b (which is > 0), gives:

$$\begin{array}{ll} \underline{b} & \text{is rejected if} \quad aa < a\beta \text{ whilst } b\alpha < bb \quad \text{and} \\ \underline{a} & \text{is rejected if} \quad bb < b\alpha \text{ whilst } a\beta < aa \end{array}$$

Observing that

$$\begin{array}{rclcll} \alpha b & = & ab\cos\theta & = & \underline{a} \bullet \underline{b}, & \\ \beta a & = & ab\cos\theta & = & \underline{a} \bullet \underline{b}, & \\ bb & = & bb\cos 0 & = & \underline{b} \bullet \underline{b} & \text{and} \\ aa & = & aa\cos 0 & = & \underline{a} \bullet \underline{a}, & \end{array}$$

these criteria take on the more readily evaluated forms

$$\underline{b} \text{ is rejected} \quad \text{if both} \quad \underline{a} \bullet \underline{a} < \underline{a} \bullet \underline{b} < \underline{b} \bullet \underline{b} \quad \text{and} \tag{6.1}$$

$$\underline{a} \text{ is rejected} \quad \text{if both} \quad \underline{b} \bullet \underline{b} < \underline{a} \bullet \underline{b} < \underline{a} \bullet \underline{a} \tag{6.2}$$

Since the square of the length of any vector $\underline{v}$ is given by $\underline{v} \bullet \underline{v}$, if the vectors are pre-sorted into increasing length order $\left\{\underline{v_1}, \underline{v_2}, \cdots, \underline{v_n}\right\}$ and compared in turn so that $a < b$, the the following inequalities are always true:

$$\begin{array}{rcl} & \underline{a} \bullet \underline{a} & < \quad \underline{b} \bullet \underline{b} \\ \text{and} \quad & \underline{a} \bullet \underline{b} & < ab < bb = \underline{b} \bullet \underline{b} \end{array}$$

Thus the criterion in Equation 6.2 can be eliminated, and that in Equation 6.1 becomes:

$$\underline{b} \text{ is rejected if } \underline{a} \bullet \underline{a} < \underline{a} \bullet \underline{b}$$

```
Calculate and store the square of the vector length v_i • v_i for each point v_i
Sort the points into increasing order of (square of) vector length
For every vector v_i {
   If (v_i has not been previously rejected) {
      For every vector v_j {                          \* v_j = v_{i+1}, v_{i+2}, ..., v_n *\
         Reject v_j if v_i • v_i < v_i • v_j,     \* where v_i • v_i is already known *\
      }
   }
}
Inner core is remaining set of points
```

Figure 6–6: An algorithm describing the method for locating the inner core set of a set atoms.

An appropriate algorithm to select the inner core from the bonding centre-relative points $\{\underline{v_1}, \underline{v_2}, \cdots, \underline{v_n},\}$ is therefore as shown in Figure 6–6

Notice that whilst this algorithm has potential worst-case quadratic complexity by considering all possible point pairs, far points towards the end of the sorted list may well be rejected by an earlier point and rejected from all future computations. Average and best case complexity may therefore be expected to be better than worst case.

The sorting stage may be implemented with any appropriate algorithm such as the `qsort` Unix implementation of quick sort.

Figure 6–7 illustrates the inner core set selected by this approach in circumstances where the more naïve approaches discussed above would select an inner core whose convex hull would not contain the centre.

○ *Stage 2: Location of Possible Biggest Hole Sites*

The second stage in the location of the biggest appropriate hole considers only the inner core atoms. Again, all pairs of atoms are considered in turn. The closest neighbour of each inner core atom is calculated and a record is kept of their mid-point. The most separated neighbouring pair do not *necessarily* indicate the biggest appropriate hole, since whilst in an appropriate direction from

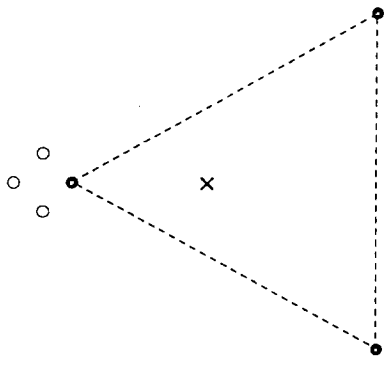


Figure 6–7: An example of the inner core set selected by vector comparison where more naïve approaches would fail.

the bonding centre, their mid-point has yet to be projected to the appropriate distance, both away from and through the centre of the metal atom, as shown in Figure 6–8.

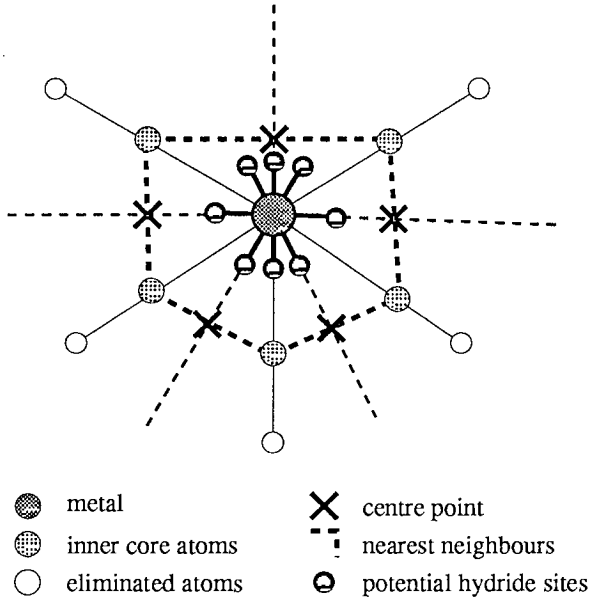


Figure 6–8: Potential sites for hydrides.

- *Stage 3: Location of Biggest Appropriate Hole*

All such mid-points are projected to the appropriate distance from the centre of the metal atom. One of these projected points is selected for the initial estimate of a hydride ligand position — that in the "biggest hole". The distance between each projected mid point and its nearest inner core atom neighbour is calculated. The projected mid point which is furthest away from its nearest such atom is selected as the biggest appropriate hole and is taken as the initial estimate of the hydride ligand position.

6.4 User Interface

In the interests of consistency, the Visualisor described in Chapter 4 was taken as a base for the interface to the Locator. Only those interface details unique to the Locator will be discussed in this section.

Figure 6–9 shows a screen dump of the Locator. The tool has the same display area for visualisation of the molecule and the same menu options along the top. The only difference between the two is extra "options" on the left hand side of the Locator. This panel is blank for the Visualisor unless the user is carrying out an operation which requires a display area. However, in the Locator the hydride ligand location options are permanently displayed on this panel.

6.4.1 Output and Experimental Control

As was mentioned in the previous chapter the original program allowed the output of several matrices and data which occasionally may be required by the user. The only default output for the current program is the revised model; the original skeleton plus the located hydride ligand(s). However, the output control option allows the user to select the output of any data available previously. The output itself has been tidied, but it is in the method of selecting which

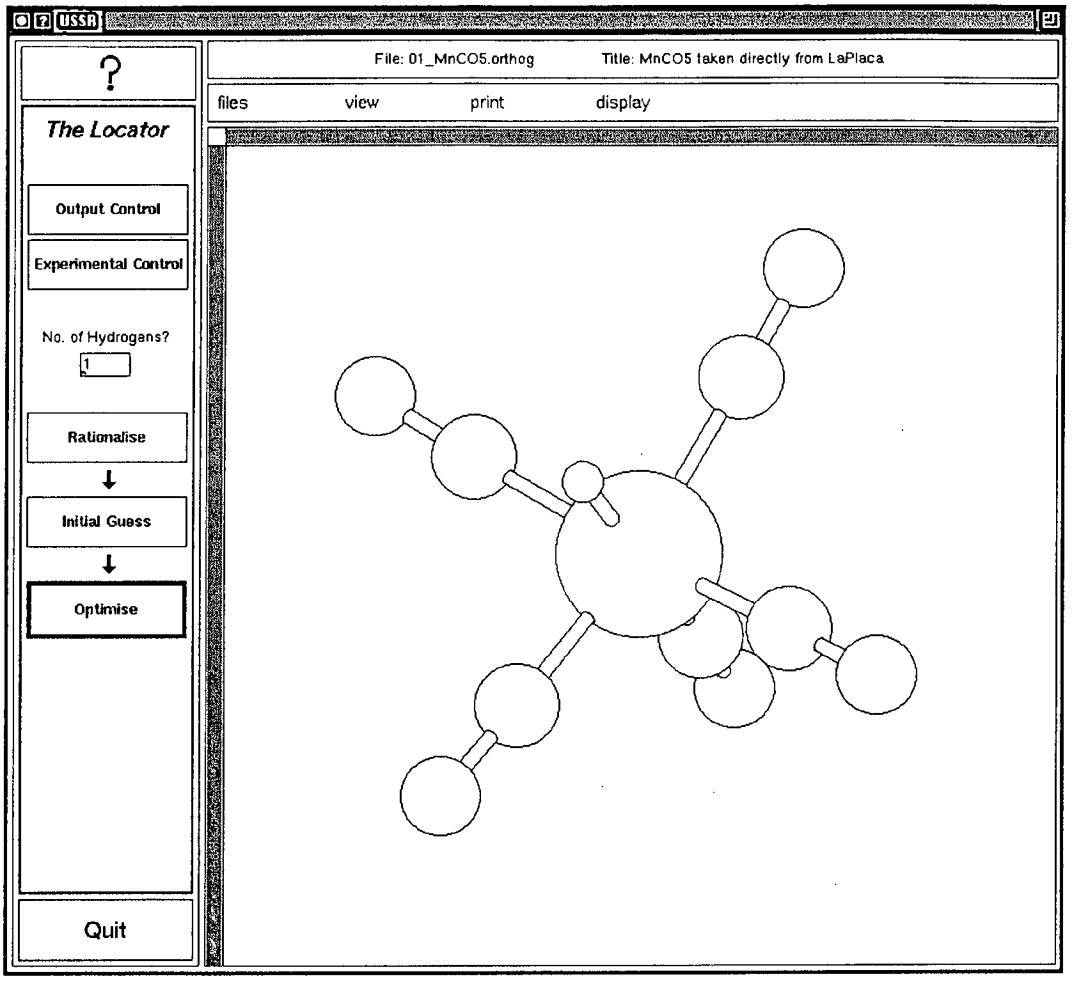


Figure 6–9: The user-interface.

data to output where there have been significant changes. In the original program the user could determine which data would be printed by setting twenty switches in an input file to either "T" (True) or "F" (False). It was necessary to consult the manual, determine the desired data number from the order it came in the manual list and then count along to the specified string in the input file to update the appropriate value. This method exposed the user to error, since

- there were many opportunities for an error to occur, such as the user selecting the wrong resource to be printed and
- there was no feedback as to which resource had been selected so that the user might not discover an error until the program had terminated.

In the current program, the user need only depress the *output control* button. A screen then appears containing a list of the different output data. Beside each item is a toggle for turning the output of the data on or off, as shown in Figure 6–10. This method has the advantages that

- no manual need be consulted as each printable resource is listed
- it is easy to see at a glance which resources are set for printing by the toggle to the right of each, black for on, white for off and
- the toggle to the side of the resource changes colour when selected or de-selected, giving the user feedback on any changes made.

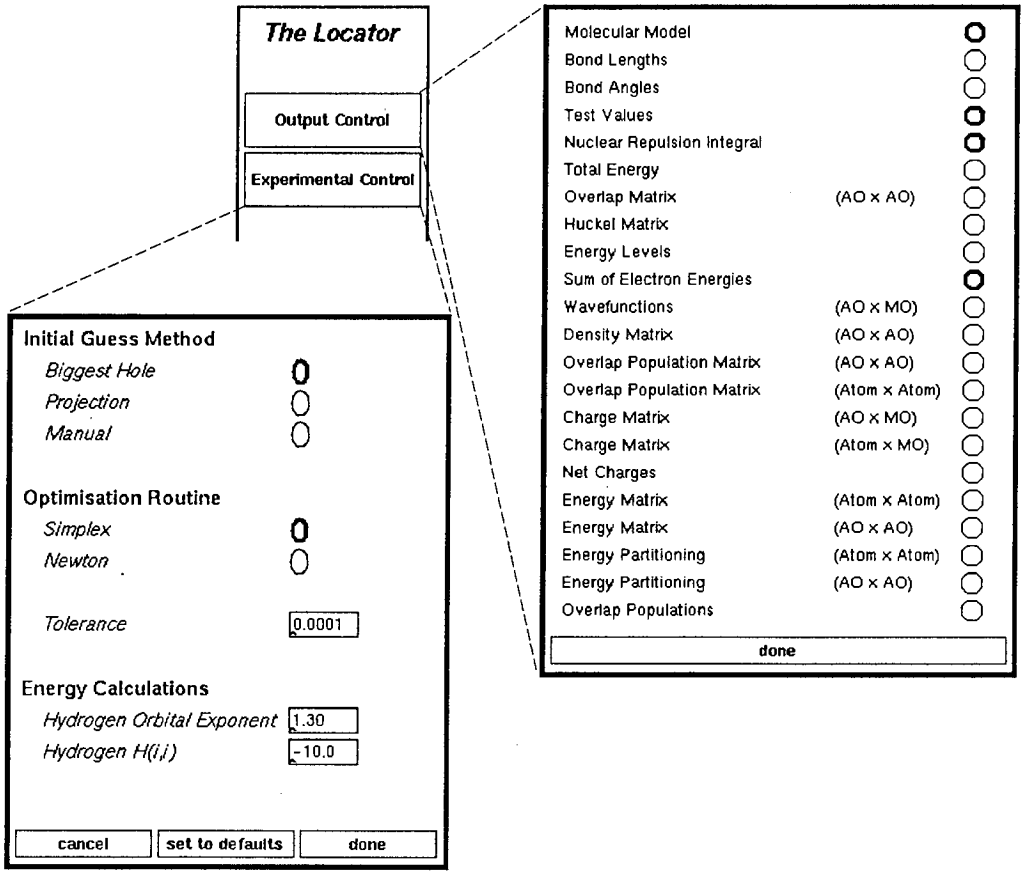


Figure 6–10: Output and experimental control.

The experimental control is similar; whereas in the previous program the experimental control parameters were entered in a combination of ways (see Figure 5–

1, Chapter 5), the output control parameters are collected under the heading of experimental control in the Locator. Figure 6–10 shows the result of depressing the *experimental control* button. Through the experimental control screen the user may;

- select between biggest hole and projection method for the initial point for the optimisation. However, if the model is recognised as having an interstitial hydride ligand the initial point will be in the centre of the metal atom cage irrespective of the method chosen by the user. Alternatively the user may choose to enter values manually in the form of coordinates. The default is the biggest hole method.

- select which optimisation routine will be used: Simplex (the default) or Newtonian.

- change the tolerance. The optimisation is an iterative process which terminates when subsequent improvements in position only result in a small decrease in energy, within a certain preset limit — the *tolerance.* The user may change the tolerance limit, for a coarser or finer optimisation, simply by editing the figure presented.

- the final result of the experiment is dependent on two parameters which describe the energy and orbital exponent of the metal-bound hydride ligand(s). These parameters may be set in a similar manner to the tolerance.

One part of the experimental control — the number of hydrogen atoms to be located — is accessed *via* the main panel rather than the experimental control screen. Arguably, the latter may at first seem more appropriate in terms of consistency. The reason for this location lies in a fundamental difference from other experimental control parameters. Whereas the others need only be accessed should the user choose to try different options, the number of hydrogen atoms is dependent on the particular model and should be set appropriately whenever a model has more than one hydrogen atom.

6.4.2 Hydride Ligand Location Procedures

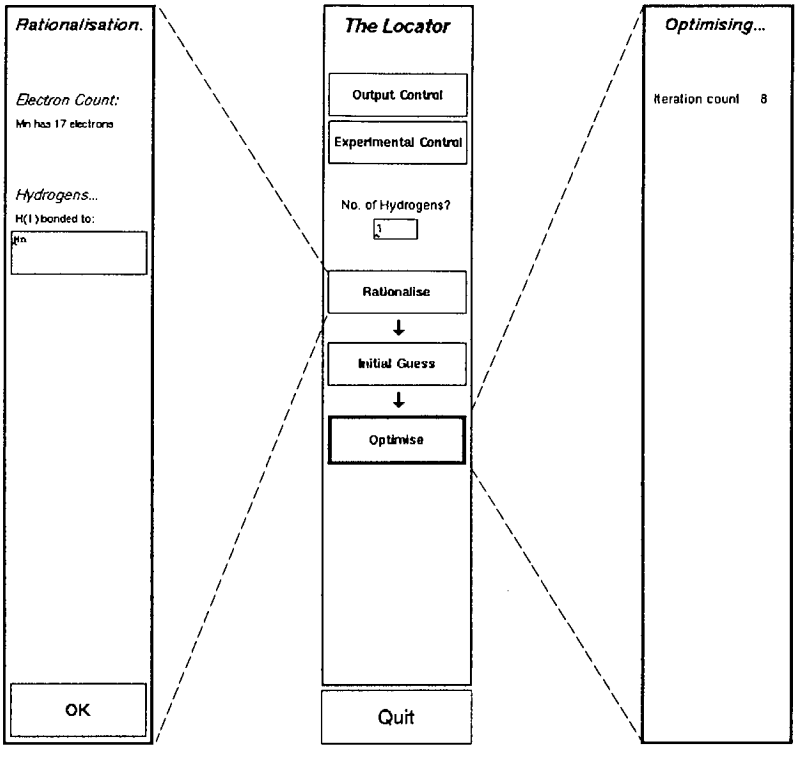


Figure 6–11: Rationalisation and optimisation panels.

The three buttons shown in Figure 6–11 correspond to the three stages discussed earlier

Rationalise:

Rationalise allows the user to control the rationalisation of the model, *i.e.* the determination of the M–H bond topology in the cluster. The rationalisation screen is shown in Figure 6–11. If the screen is being displayed because the program has failed to rationalise the model itself an explanation in the way of an apology is given as part of the display. Otherwise the screen consists of;

- a label "Rationalisation" to remind the user of the current state,
- a list of the transition metal atoms within the molecule and their calculated electron count and

- a space, for each hydrogen atom, to display the metal atoms to which it is bonded. If the model has been successfully rationalised this space will contain the metal atoms the program has calculated as being bonded to each hydrogen atom. Otherwise the space will be left blank, requiring this information to be input by the user. Even if the model has been successfully rationalised by the program, the user may still override this by simply typing in the display which may be edited.

Initial Guess:

Initial Guess allows the user to approve the initial estimate for the position of hydrogen ligands before any optimisation is undertaken. If the user does not agree with the initial point it may be changed in either of two ways. The user may return to the rationalisation procedure and redefine the bonding of the hydride ligands. Alternatively, the user may select the alternative initial estimate method from the experimental control screen.

Optimise:

Optimise initiates the optimisation procedure. On initiation the optimisation screen is displayed (see Figure 6–11). During the optimisation process the iteration count in the optimisation screen is updated and whenever a more suitable position for the hydride ligand is encountered the screen is updated showing the new position of the ligand.

The procedures behind the buttons must be carried out in a specific order (*rationalise model* → *make initial estimate* → *optimise*). However, the user does not have to press each of these three buttons in turn. The depression of any of the three buttons automatically performs the pre-requisite procedures before proceeding, as illustrated by the algorithm in Figure 6–12.

All of the operations which would normally be available to the user of the Visualisor are available to the user of Locator. Models may be rotated, labels turned on and off and bond-lengths queried *etc.* before, during and after the

```
Rationalise button depressed {
    If (model has not already been rationalised) {
        call rationalisation procedure
        If (rationalisation procedure was not successful) {
            write "apologies: I was unable to rationalise the model"
        }
    }
    turn atomic labels on
    display rationalisation screen
}

Initial Guess button depressed {
    If (model has not already been rationalised) {
        call rationalisation procedure
        If (rationalisation procedure was not successful) {
             write "apologies: I was unable to rationalise the model"
             turn atomic labels on
             display rationalisation screen
        }
    }

    If (initial guess has not been made or rationalisation has changed) {
        call initial guess routine
        display molecule with new hydride position
    }
}


Optimise button depressed {
    If (model has not already been rationalised) {
        call rationalisation procedure
        If (rationalisation procedure was not successful) {
            write "apologies: I was unable to rationalise the model"
            turn atomic labels on
            display rationalisation screen
        }

    }

    If (initial guess has not been made or rationalisation has changed) {
        call initial guess routine
        display molecule with new hydride position
    }
    display optimisation screen
    call optimisation procedure
}
```

Figure 6–12: *A series of algorithms illustrating the interaction of the rationalise, initial guess and optimise buttons.*

optimisation routine. The rotation of the molecule is particularly useful since the hydride ligand position(s) may move around substantially during the course of the optimisation and may disappear behind other atoms preventing the user from effectively visualising the current position of the hydride ligand.

6.5 Performance Evaluation of the Locator

The original objectives related with the performance of the Locator were that it should be consistent, successful, accurate, efficient and that it should not be restricted to an unreasonably limited number of atoms. The Locator is evaluated in terms of these criteria, both absolutely and by comparison with CCCP, to establish whether the addition of the initial estimate method to CCCP provided by the Locator constitutes an improvement. The twin stages of the initial estimate methods — rationalisation and initial position — are also evaluated in terms of these criteria.

In the evaluation of the Locator only the biggest appropriate hole method is considered. It became apparent during the course of the evaluation that the projection method was grossly inferior to the biggest appropriate hole method, and was not sufficiently successful to be considered further. Details of the comparison between the two initial estimate methods are given in Appendix F.

Due to the large number of models, summaries of the results experimental results are presented in this chapter. The full results for each experiment are tabulated in Appendix E. The results are categorised throughout this chapter by type of hydride ligand for simplicity. Further, a distinction is made between neutron and X-ray diffraction models, not because it is thought that the Locator may inherently be better at locating hydride ligands located by one technique than another, but because the neutron diffraction located hydride ligands are more accurate and the results for those models considered more reliable.

6.5.1 Models Used

The models examined in this evaluation are listed in Appendix C. These comprise:

- those models examined by Mitchell [13] in his evaluation of CCCP — and in the testing of CCCP in Chapter 5 of this thesis
- those models examined by Orpen [12] in his evaluation of Hydex.
- all transition metal hydride compounds for which neutron diffraction determined coordinates are available in the Cambridge Structural Database.
- a selection of transition metal hydride compounds for which coordinates determined by X-ray diffraction are available. In selecting these models, attempts were made to:
  - maintain reasonable proportions of different types (τ, μ_2, μ_3 or ι) of hydride ligands.
  - restrict attention to models where the X-ray diffraction experiment was performed at low temperature conditions, as the results from such experiments are considered more reliable.

A maximum of three of hydride ligands may be optimised simultaneously, as stated in Chapter 5. Neutron diffraction models are known to be the most reliable for the evaluation of the Locator, but models of these type are scarce. To maximise the number of neutron diffraction characterised models examined, those with more than three hydride ligands were included in the evaluation. These models were treated as a sequence of "sub-models", optimising a maximum of three hydride ligands at each stage whilst the other hydride ligands remained fixed at the literature positions. The breakdown of these models is detailed in Appendix C.

Table 6–1 shows the distribution of models, with regard to both their hydride ligand types and the analytical techniques by which their coordinates have been determined. Each model is illustrated in Appendix D complete with hydride ligand(s) located by energy optimisation. The original model is alongside each

for comparison with the literature hydride ligands positions as located by traditional analytical techniques.

	Neutron	Xray	Total
Terminal	17	47	64
Edge Bridging	30	29	59
Face Bridging	5	32	37
Interstitial	2	1	3
Mixed	6	9	15
Total	60	118	178

Table 6–1: The distribution of models according to hydride ligand type.

6.5.2 Evaluation of the Preliminary Stages of the Locator.

Rationalisation

Table 6–2 summarises the success, or otherwise,[3] of the rationalisation procedure. As can be seen from Table 6–2, the results of the automatic rationalisation procedure tend to depend on the ligand type. The rationalisation procedure predicts correct hydride ligand bonding interactions for 86% of the models containing only terminal hydride ligands, failing to rationalise only 9 out of the 74 models. However, the rationalisation procedure correctly rationalised just 26% of the models containing only edge-bridging ligands, whilst failing to rationalise 29 of the 66 models.

The overall success rate of rationalisation procedure, at 43% correctly and 24% incorrectly rationalised, would appear to be unsatisfactory on first consideration. However, the remainder of this chapter shows that, although correct

[3] "Incorrect" and "Failed" are distinct. Incorrect indicates that the Locator was able to suggest bonding interactions of the molecular model, but that these did not match those determined by traditional analytical techniques. Failure indicates that the Locator was unable to suggest any bonding interactions of the hydride ligands

		Neutron		Xray		Total	
Terminal	*Correct*	24	(89%)	40	(85%)	64	(86%)
	Incorrect	0	(0%)	1	(2%)	1	(1%)
	Failed	3	(11%)	6	(13%)	9	(12%)
Edge Bridging	*Correct*	10	(27%)	7	(24%)	17	(26%)
	Incorrect	10	(27%)	10	(34%)	20	(30%)
	Failed	17	(46%)	12	(41%)	29	(44%)
Face Bridging	*Correct*	0	(0%)	1	(3%)	1	(3%)
	Incorrect	2	(25%)	20	(63%)	22	(55%)
	Failed	6	(75%)	11	(34%)	17	(43%)
Interstitial	*Correct*	1	(50%)	0	(0%)	1	(33%)
	Incorrect	0	(0%)	0	(0%)	0	(0%)
	Failed	1	(50%)	1	(100%)	2	(67%)
Mixed	*Correct*	0	(0%)	0	(0%)	0	(0%)
	Incorrect	1	(100%)	3	(33%)	4	(40%)
	Failed	0	(0%)	6	(67%)	6	(60%)
Total	*Correct*	35	(47%)	48	(41%)	83	(43%)
	Incorrect	13	(17%)	34	(29%)	47	(24%)
	Failed	27	(36%)	36	(31%)	63	(33%)

Table 6–2: The success of the rationalisation procedure.

rationalisation may increase the accuracy of the optimisation it is by no means necessary for a correct final answer.

Some of the models where the bonding between metal atoms and the hydride ligands predicted by the theoretical rationalisation of the Locator does not match those observed experimentally are examined below in an attempt to identify the cause of these anomalies.

Model 77, $Os_3H_2(CO)_{10}(\mu_2\text{–}CH_2)$

The electron count of each metal atom shows two of the osmium atoms to have seventeen electrons whilst the other has eighteen. This suggests that the two hydride ligands bond terminally to the two electron deficient osmium atoms. Neutron diffraction reveals two edge-bridging hydride ligands.

Model 86, $Rh_3H_3(PH_3)_6$

Electron counting shows that in the absence of any hydride ligands the three rhodium atoms would be equivalent, each having fifteen electrons. Each rhodium atom requires another three hydrogen ligands to satisfy the eighteen electron rule. The only way for the three hydride ligands

to bond with the three metal atoms to satisfy the electron requirement of each is to form three face-bridges across the three metal atoms. Since this is impossible sterically, the program reverts to trying to satisfy the *sixteen* electron rule. In some oxidation states some metal atoms, particularly Rh^{I}, Ir^{I}, Pd^{II}, Pt^{II} and Au^{III}, require only sixteen electrons to be stable, and under these conditions show square planar rather than octahedral geometry. In this case, the Locator predicts three terminal hydride ligands. Neutron diffraction reveals that all three hydride ligands are in fact edge-bridging, apparently giving each metal atom seventeen electrons.

The actual explanation of the bonding in this and, in fact, the previous molecule is that the hydride ligand forms a three-centre two-electron *(3c–2e)* bond. Rather than the Rh–Rh *(2c–2e)* bond being *supplemented* by additional hydride bonding, two rhodium atoms and one hydrogen atom come together to form a *3c–2e* bond. There are three such bonding interactions in the complex, as shown schematically in Figure 6–13. Each (13e) $\{Rh(PH_3)_2\}$ fragment gains access to a total of three more electrons by this mechanism, affording sixteen electrons in all, consistent with their "square planar" geometries.

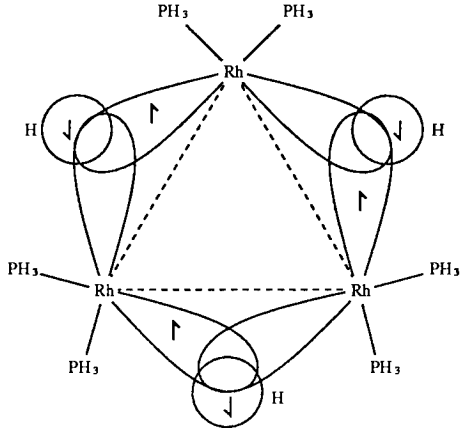


Figure 6–13: The forming of three-centre two-electron bonds in $Rh_3H_3(PH_3)_6$.

The Locator deals only with standard bonding interactions and is not programmed to deal with three-centre two-electron bonds.

Model 113, $[Os_4H_2(CO)_{12}]^{2-}$

The Locator correctly rationalised that each of the four osmium atoms in $[Os_4H_2(CO)_{12}]^{2-}$ require one electron for the completion of a quota of eighteen, and that the two hydride ligands could satisfy the osmium atoms by edge-bridging to two metal atoms each. The rationalisation falls short of identifying a unique bonding configuration because there were several ways in which the hydride ligands could bridge the four metal atoms.

Model 127, $[Os_6H(CO)_{18}]^-$

All the osmium atoms in $[Os_6H(CO)_{18}]^-$ are equivalent in that each has the same number of valence electrons, including those donated by carbonyl ligands. The hydride ligand would therefore seem to be interstitial, as indeed it is in the ruthenium analogue, $[Ru_6H(CO)_{18}]^-$. The face-bridging nature of the ligand is indicated only by the lengthening of the appropriate Os–Os bonds and the bending of the adjacent ligands. The Locator is not programmed to recognise such features.

Initial Estimate Method

The distance errors between the hydride ligand locations as predicted by energy optimisation and published in the literature are summarised in Table 6–3. The *success* of the final position is judged on whether the topology of the hydride ligand is correct. This criterion is not appropriate to evaluate the initial estimate, since the topology of the initial estimate depends not on the initial estimate itself but on the rationalisation. However the accuracy of the initial guesses may be measured, as summarised in Table 6–3.

As Table 6–3 shows, the initial estimate routine generally provided a reasonable estimate of the hydride ligand position — with an average error of 1.278 ±

	Automatic Rationalisation						User Rationalisation					
	Neutron		Xray		Total		Neutron		Xray		Total	
Terminal	50	0.944 ±0.106	59	1.083 ±0.097	109	1.019 ±0.071	55	0.944 ±0.096	69	1.049 ±0.083	124	1.002 ±0.063
Edge Bridging	38	1.139 ±0.152	20	0.861 ±0.213	58	1.044 ±0.124	67	0.466 ±0.075	35	0.491 ±0.040	102	0.475 ±0.052
Face Bridging	2	2.962 ±0.916	30	2.440 ±0.263	32	2.472 ±0.251	15	0.476 ±0.103	45	0.296 ±0.024	60	0.341 ±0.032
Interstitial	1	0.008 n/a	0	n/a n/a	1	0.008 n/a	2	0.016 ±0.008	1	0.177 n/a	3	0.070 ±0.054
Mixed	3	0.916 ±0.238	6	2.258 ±0.628	9	1.811 ±0.468	3	1.244 ±0.131	19	0.295 ±0.043	22	0.424 ±0.081
Total	94	1.055 ±0.090	115	1.460 ±0.114	209	1.278 ±0.076	142	0.662 ±0.057	169	0.643 ±0.045	311	0.644 ±0.035

Table 6–3: The accuracy, in Å, of the initial estimate routine.

0.076Å[4]. Indeed, the initial estimate is sometimes closer to the position revealed by neutron or X-ray diffraction than the result arrived at after optimisation — see models (72) $Mo_2H(CO)_4(\eta^5\text{–}C_5H_5)_2(PMe_2)$, (76) $Os_3H_2(CO)_{10}$ and (162) $[Ru_6H(CO)_{18}]^-$. Further, if erroneously rationalised models are corrected by user intervention before making the initial estimate, the average error falls to 0.644 ± 0.035Å.

6.5.3 Comparative Evaluation the Locator

A full comparison between these two programs is impossible since, with CCCP, the initial point estimation for subsequent optimisation is made by the user and there are no known recommendations for selecting or calculating this initial point. Users may have been assumed to employ their own experience and chemical knowledge for the initial estimate of the hydride location, and calculate the coordinates geometrically. However, given the time and effort required to perform this calculation and there being no expressed requirement for a reasonable initial point, users would be unlikely to go to such lengths. It seems more probable that some random point in the molecular space was chosen. Thus, the Locator is evaluated by comparison of its results with those produced by CCCP in experiments 1.1, 3.1, and 3.2 of Chapter 5 — for which CCCP was

[4]these, and subsequent measurements are presented as mean ± standard error

given initial estimates at arbitrary points, the centre of the molecule and the centre of the transition metals respectively.

The success rates of the two tools are compared in Table 6–4. Only Mitchell's models are examined in this experiment since only these are addressed by the testing of CCCP in Chapter 5.

		Neutron		*Xray*		*Total*	
Random	(1.1)	3 / 9	(33%)	0 / 4	(0%)	3 / 13	(23%)
Centre of Molecule	(3.1)	7 / 9	(78%)	2 / 4	(50%)	9 / 13	(69%)
Centre of Metals	(3.2)	6 / 9	(67%)	3 / 4	(75%)	9 / 13	(69%)

(a) Success rates of CCCP.

	Automatic Rationalisation						*User Rationalisation*					
	Neutron		*Xray*		*Total*		*Neutron*		*Xray*		*Total*	
The Locator	8 / 8	(100%)	3 / 4	(75%)	11 / 12	(92%)	9 / 9	(100%)	3 / 4	(75%)	12 / 13	(92%)

(b) Success rates of The Locator.

Table 6–4: *A comparison between CCCP and the Locator with respect to their success rates.*

The Locator compares favourably with CCCP in this table. When taking random coordinates as an initial estimate, the success rate of CCCP was only 23%, succeeding for just 3 of the 13 models. Taking the centre of the molecule and metal(s) each gave an improved success rate of 69%. However, when taking an the initial point provided by the biggest appropriate method, the Locator achieved a success rate of 92%, succeeding for 12 of the 13 models tested.

Table 6–5(a) compares the two tools with regards to their accuracy. This table shows that the positions predicted by the Locator are more accurate than those located by CCCP from the initial points described. The possibility of CCCP being more accurate over those models for which it succeeded is examined in Table 6–5(b) which applies only such models in the calculation. The figures given in the lower table are the equivalent accuracy figures for the Locator, averaged over the same models. The table shows the Locator is of comparative accuracy with CCCP even if those models where the latter failed are not considered.

	Neutron	Xray	Total
Random (1.1)	16 5.837 ±2.112	3 6.007 ±4.748	19 5.864 ±1.878
Centre of Molecule (3.1)	16 1.706 ±0.904	3 1.182 ±0.987	19 1.624 ±0.770
Centre of Metals (3.2)	16 1.702 ±0.659	3 0.263 ±0.076	19 1.475 ±0.566

	Automatic Rationalisation			User Rationalisation		
	Neutron	Xray	Total	Neutron	Xray	Total
The Locator	15 0.191 ±0.042	3 0.285 ±0.065	18 0.207 ±0.037	16 0.180 ±0.040	3 0.257 ±0.068	19 0.192 ±0.035

(a) Average accuracy, in Å, applied to Mitchell's Models.

	Neutron	Xray	Total
Random (1.1)	7 0.188 ±0.014	0 n/a n/a	7 0.188 ±0.014
Centre of Molecule (3.1)	11 0.192 ±0.059	2 0.198 ±0.114	13 0.193 ±0.051
Centre of Metals (3.2)	9 0.200 ±0.069	3 0.263 ±0.076	12 0.216 ±0.054

	Automatic Rationalisation			User Rationalisation		
	Neutron	Xray	Total	Neutron	Xray	Total
(1.1)	7 0.182 ±0.016	0 n/a n/a	7 0.182 ±0.016	7 0.184 ±0.015	0 n/a n/a	7 0.184 ±0.015
The Locator (3.1)	10 0.191 ±0.063	2 0.267 ±0.108	12 0.204 ±0.054	11 0.174 ±0.058	2 0.231 ±0.109	13 0.183 ±0.051
(3.2)	8 0.205 ±0.079	3 0.285 ±0.065	11 0.227 ±0.060	9 0.182 ±0.072	3 0.257 ±0.068	12 0.201 ±0.056

(b) Average accuracy, in Å, applied only to those models for which the various methods succeeded.

Table 6–5: A comparison of the Locator and CCCP with respect to accuracy.

The relative efficiencies of the Locator and CCCP are compared in Table 6–6. This comparison proves somewhat inconclusive. When the centre of the metal array is chosen as an initial estimate, CCCP appeared to converge in slightly fewer iterations. However, when the centre of the molecule is chosen as a starting point, CCCP converged in significantly more iterations. Moreover, the Locator is substantially more efficient when erroneous rationalisation is corrected, converging in a lower average number of iterations.

6.5.4 Absolute Evaluation of the Locator

Consistency

The Locator would appear to have largely achieved the consistency objective — that the program should not be dependent on an initial point given by the user. The addition of the initial estimate option reduces the required input from an absolute position in an infinite space to a choice between the two initial estimate

	Neutron	Xray	Total
Centre of Molecule (3.1)	9 222 ±58	4 372 ±219	13 268 ±75
Centre of Metals (3.2)	9 232 ±95	4 196 ±46	13 221 ±66

	Automatic Rationalisation			User Rationalisation		
	Neutron	Xray	Total	Neutron	Xray	Total
The Locator	8 260 ±104	4 179 ±81	12 233 ±73	9 175 ±48	4 104 ±18	13 153 ±34

(a) Average efficiency applied to Mitchell's Models.

	Neutron	Xray	Total
Centre of Molecule (3.1)	7 197 ±72	2 692 ±286	9 307 ±103
Centre of Metals (3.2)	6 194 ±115	3 205 ±63	9 198 ±76

	Automatic Rationalisation			User Rationalisation		
	Neutron	Xray	Total	Neutron	Xray	Total
The Locator (3.1)	7 241 ±140	2 273 ±148	9 249 ±106	7 147 ±58	2 129 ±4	9 143 ±44
(3.2)	6 246 ±156	3 203 ±110	9 230 ±109	6 135 ±67	3 103 ±25	9 124 ±38

(b) Average efficiency applied only to those models for which the various methods succeeded.

***Table 6–6:** A comparison of the Locator and CCCP with respect to efficiency.*

methods, and/or a user intervention request to correct an erroneous rationalisation. Admittedly, these choices have some bearing on the final outcome. However, the variability of the final result is reduced from the original method.

Atom Number Limitations

The atom count limit was increased from forty non-hydrogen atoms to fifty to accommodate $[Ru_6H(CO)_{18}]^-$ and $[Os_6H(CO)_{18}]^-$. Since the common practice described in Chapter 5 of reducing large functional groups to smaller groups mostly reduces the effective "atom" count to below this number, this increase was considered adequate.

Success

Table 6–7 summarises the results given in Appendix E and details the success of the program with respect to the different hydride ligand types. The success rate is calculated on a per-model rather than per-hydride basis. Unless the Locator

correctly locates all the hydride ligands of a particular model it is deemed to have failed for that model. For example the 71% rate for the Locator for neutron diffraction models with terminal hydride ligands means that the Locator succeeded for 17 out of the 24 models, despite the successful location of a total of 37 of the 50 hydride ligands.

	Automatic Rationalisation						*User Rationalisation*					
	Neutron		*Xray*		*Total*		*Neutron*		*Xray*		*Total*	
Terminal	17 / 24	(71%)	31 / 41	(76%)	48 / 65	(74%)	19 / 27	(70%)	35 / 47	(74%)	54 / 74	(73%)
Edge Bridging	18 / 20	(90%)	14 / 17	(82%)	32 / 37	(86%)	35 / 37	(95%)	27 / 29	(93%)	62 / 66	(94%)
Face Bridging	1 / 2	(50%)	9 / 21	(43%)	10 / 23	(43%)	8 / 8	(100%)	27 / 31	(87%)	35 / 39	(90%)
Interstitial	1 / 1	(100%)	0 / 0	n/a	1 / 1	(100%)	2 / 2	(100%)	0 / 1	(0%)	2 / 3	(67%)
Mixed	1 / 1	(100%)	2 / 3	(67%)	3 / 4	(75%)	1 / 1	(100%)	7 / 9	(78%)	8 / 10	(80%)
Total	**38 / 48**	**(79%)**	**56 / 82**	**(68%)**	**94 / 130**	**(72%)**	**65 / 75**	**(87%)**	**96 / 117**	**(82%)**	**161 / 192**	**(84%)**

Table 6–7: *A summary of success rates for the models with different hydride ligands.*

However, it must be borne in mind that the figures quoted above are for X-ray and neutron diffraction models combined. As was discussed earlier the hydride ligand coordinates determined by X-ray diffraction are not as reliable as those determined by neutron diffraction. Moreover some of the published coordinates or topological positions may be disputed, since they have been located by unreliable methods. For such models, a difference between the published coordinates or positions and those located by the Locator does not necessarily mean a failure on the part of the Locator. Some such examples are detailed below. More reliable statistics may, perhaps, be obtained when considering only the 60 neutron models. It must be noted, however, that 60 models, is in effect 75 because of the treatment of those models with more than three hydride ligands. When considering only models determined by neutron diffraction the success rate of the Locator is 79%, rising to 87% on user intervention at the rationalisation stage.

◦ *Models for which the Locator "Failed"*

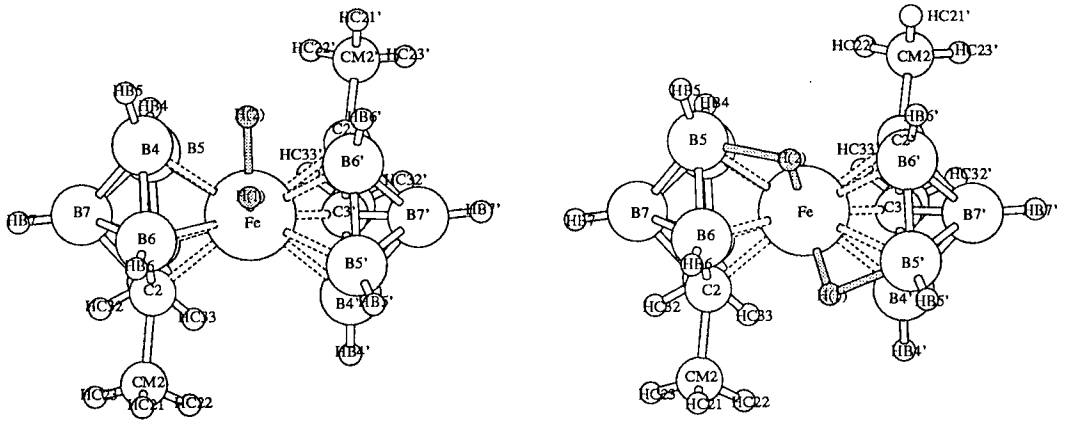
The biggest appropriate hole method failed for 36 out of 130 models which the Locator was able to rationalise. With appropriate user intervention during rationalisation the Locator failed for 31 out of 192 models. Some of the models for which the Locator failed to locate reasonable hydride ligand(s) were:

Model 20: $FeH_2\{(CCH_3)_2B_4H_4\}_2$

This model was originally studied by X-ray diffraction [123]. Therefore accurate positions of the two hydride ligands remain unknown. Using steric arguments, it was observed that the non-hydride ligands tilted in a direction opposite to that expected in the absence of the metal-bound hydride ligands. The hydride ligands are thereby indicated to be in the vicinity of the Fe–B5–B6 and Fe–B5′–B6′ faces. The positions estimated by the Locator are close to but not entirely consistent with the steric predictions.

The first biggest appropriate hole is located exactly where the bending back of the non-hydride ligands would indicate were there only one hydride, *i.e.* between the Fe–B5–B6 and Fe–B5′–B6′ faces as shown in Figure 6–14. The second biggest appropriate hole is located between the Fe–B4–B5 and Fe–C2′–B6′ faces. Optimisation predicts two edge-bridging hydride ligands, one bonded to Fe and B5, in the vicinity of the Fe–B5–B6 face (correct), the other bonded to Fe and B5′, but in the vicinity of the Fe–B4′–B5′ face (incorrect).

In the absence of a neutron diffraction study we cannot comment further on the appropriateness of these optimised positions. We note, however, that the molecule has crystallographically-imposed C_2 symmetry which suggests that the two hydride ligands should also be related in this way. At present neither CCCP nor the Locator are set up to bind together the positions of symmetry related groups of hydride ligands during optimisation. Furthermore, it may be that the biggest appropriate hole that could accommodate two hydride ligands may in fact not be spherical. However, due to the serial rather than parallel manner of locating the biggest appro-



(a) The initial estimate. *(b) After optimisation.*

Figure 6–14: *The hydride ligand location estimate and subsequent optimisation of* $FeH_2\{(CCH_3)_2B_4H_4\}_2$.

priate hole such a situation is not recognised. It would seem that a more detailed analysis of the shape of the biggest appropriate hole could feature in further development of the Locator.

Model 112: $Os_4H_3(CO)_{11}(C_6H_9)$

Again, the structure of this molecule was found by X-ray diffraction and thus the three hydride ligand locations are not accurately known. In the light of evidence from spectroscopic data, Bhaduri *et. al.* [124] proposed that the hydride ligands were edge-bridging, across Os1–Os3, Os2–Os3 and Os2–Os4 respectively. The Locator found two edge-bridging hydride across Os1–Os3 and Os2–Os4, but found the third to be capping the Os2, Os3, Os4 face, as shown in Figure 6–15.

On further investigation, however, it was discovered that the H(2) hydride ligand tended towards the Os2 and Os3 atoms (bond-lengths of 1.61Å and 1.57Å respectively) and away from the Os4 atom (bond-length of 1.81Å). This is not considered to be entirely inconsistent with the findings of Bhaduri *et. al.* since, positionally, the two predictions are similar. Furthermore, it must be remembered that indirect methods are not particularly reliable. The hydride ligand position predicted by the Locator

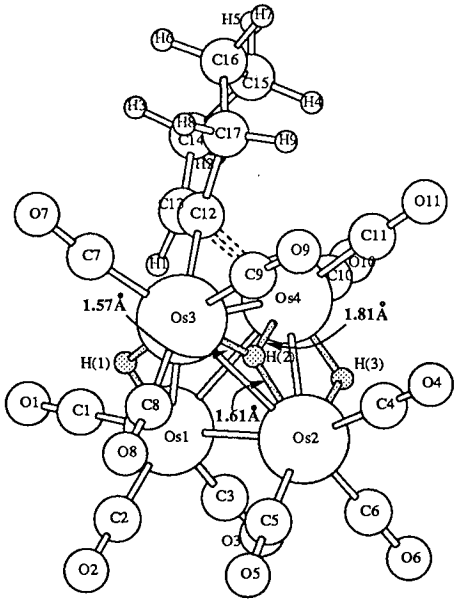


Figure 6–15: $Os_4H_3(CO)_{11}(C_6H_9)$, with the hydride ligand positions predicted by the Locator.

could actually be correct, and the third hydride ligand face- rather than edge-bridging.

Model 114: $Os_5H_2(CO)_{16}$

The Locator failed to correctly predict the site of one of the two edge-bridging ligands with automatic rationalisation, settling on a face-bridging position (Os1,Os2,Os3) instead of the edge-bridge across Os2–Os4. However, on user intervention at the rationalisation stage, the hydride ligand sites were correctly located.

Model 115: $Os_3PtH_2(CO)_{10}(PH_3)$

The $Os_3PtH_2(CO)_{10}(PH_3)$ model could not be rationalised automatically by the Locator and thus required rationalisation from the user. There would appear to be some dispute over the two edge-bridging hydride ligand locations, proposed in the light of steric evidence by Farrugia *et. al.* [125]. In the cited article, it was proposed that the hydride ligands bridged the two longest M–M edges of their type, *i.e.* across Os2–Os3 (2.789Å) and Pt–

Os1 (2.863Å). However the molecular model provided by the Cambridge Structural Database placed one of the hydride ligands across Pt–Os3 (2.832 Å) rather than Pt–Os1. M–M bond lengthening and ligand bend back provided the evidence in the estimation of the hydride ligand location, and would seem to indicate that the hydride ligand locations given in the paper (Os2–Os3 and Pt–Os1) are those intended.

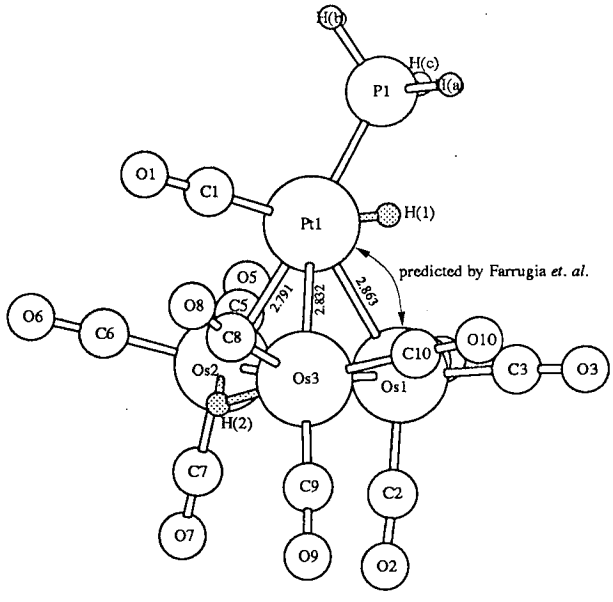


Figure 6–16: *$Os_3PtH_2(CO)_{10}(PH_3)$, with the hydride ligand positions predicted by the Locator.*

The Locator successfully located the hydride bridging between Os2 and Os3, but determined that the second was a terminal hydride on Pt1. Figure 6–16 shows the model with the located hydride ligands, and also illustrates the position that Farrugia *et. al.* predicted for H(1). It can be seen that, although the located hydride ligand is terminal, it is in the vicinity of that originally predicted, *i.e.* between the Pt and Os1. The reasoning employed to deduce that H(1) was a bridging hydride across Pt–Os1 is not entirely inconsistent with it being terminal. The respective bond lengths of Pt–Os1, Pt–Os2 and Pt–Os3 are 2.863Å, 2.791Å and 2.832Å. Note that there is a larger difference of 0.041Å between the two shortest bonds than 0.031Å between the two longest. There is no apparent reason for this dif-

ference between the two shorter bonds. Thus it is feasible that the bond lengthening may be produced by factors other than a bridging hydride.

Model 125: $FeCo_3H(CO)_9(PH_3)_3$

The Locator failed for this model when using automatic rationalisation because the rationalisation was grossly incorrect, predicting that the hydride ligand is bonded to the iron atom instead of face-bridging the three cobalt atoms. A local minimum would seem to be encountered and accepted without the actual global minimum ever being encountered. The Locator has no problem in pinpointing the correct hydride ligand site if the rationalisation is corrected by the user. It must be remembered that providing the topology of the hydride ligand is a requirement of Hydex. The fact that it is necessary to *occasionally* provide the Locator with this information is not considered unreasonable.

Accuracy

Table 6–8 shows measurements of the accuracy — the displacement of the located hydride(s) from their neutron diffraction coordinates in Å — categorised by model type. Table 6–8 (a) includes the results of all experiments, whereas Table 6–8 (b) includes only those experiments for which the Locator is deemed to have succeeded. The latter may be considered a more relevant measurement of accuracy.

It can be seen that for the 192 experiments performed, the hydride ligand locations predicted by the Locator were only displaced from those located by neutron diffraction by an average of 0.261 $\pm$ 0.031Å. The accuracy figure is slightly improved to 0.256 $\pm$ 0.009Å when erroneous rationalisation is corrected by the user. Considering only the neutron models, the figures for all models, and those for which the Locator succeeded were 0.210 $\pm$ 0.028Å and 0.236 $\pm$ 0.013Å respectively. These figures are certainly satisfactory, given that a typical metal–hydrogen ligand bond length is between 1.6Å and 1.8Å.

	Automatic Rationalisation			*User Rationalisation*		
	Neutron	*Xray*	*Total*	*Neutron*	*Xray*	*Total*
Terminal	50 0.616 ±0.126	59 0.773 ±0.205	109 0.701 ±0.125	55 1.178 ±0.431	69 1.959 ±1.130	124 1.612 ±0.656
Edge Bridging	38 0.370 ±0.124	20 0.422 ±0.200	58 0.388 ±0.105	67 0.378 ±0.085	35 0.292 ±0.033	102 0.349 ±0.058
Face Bridging	2 1.789 ±1.064	30 1.651 ±0.359	32 1.660 ±0.339	15 0.282 ±0.052	45 0.800 ±0.254	60 0.670 ±0.193
Interstitial	1 0.031 n/a	0 n/a n/a	1 0.031 n/a	2 0.032 ±0.001	1 3.059 n/a	3 1.041 ±1.009
Mixed	3 0.122 ±0.007	6 3.760 ±3.259	9 2.547 ±2.189	3 0.117 ±0.007	19 1.004 ±0.296	22 0.883 ±0.003
Total	94 0.519 ±0.088	115 1.097 ±0.224	209 0.837 ±0.131	142 0.667 ±0.174	169 1.204 ±0.468	311 0.958 ±0.267

(a) Average accuracy, in Å, applied to all models.

	Automatic Rationalisation			*User Rationalisation*		
	Neutron	*Xray*	*Total*	*Neutron*	*Xray*	*Total*
Terminal	34 0.210 ±0.017	46 0.354 ±0.097	80 0.293 ±0.057	39 0.200 ±0.016	52 0.252 ±0.019	91 0.238 ±0.013
Edge Bridging	33 0.207 ±0.019	19 0.223 ±0.026	52 0.213 ±0.015	63 0.259 ±0.020	33 0.245 ±0.022	96 0.254 ±0.015
Face Bridging	1 0.725 n/a	11 0.248 ±0.050	12 0.288 ±0.060	15 0.282 ±0.052	39 0.286 ±0.024	54 0.286 ±0.021
Interstitial	1 0.031 n/a	0 n/a n/a	1 0.031 n/a	2 0.032 ±0.001	0 n/a n/a	2 0.032 ±0.001
Mixed	3 0.122 ±0.007	4 0.313 ±0.054	7 0.231 ±0.048	3 0.117 ±0.007	15 0.372 ±0.113	18 0.349 ±0.047
Total	72 0.210 ±0.028	80 0.306 ±0.057	152 0.261 ±0.031	122 0.236 ±0.013	139 0.273 ±0.013	261 0.256 ±0.009

(b) Average accuracy, in Å, applied only to those models for which the Locator succeeded.

Table 6–8: A summary of the accuracy of the Locator.

Efficiency

Table 6–9 shows the average number of iterations for the Locator to converge. Figures are tabulated for automatic rationalisation and, where appropriate, user-corrected rationalisation. As earlier, figures are averaged over all experiments and over those for which the Locator succeeds. On average, the Locator is seen to converge with fewer iterations if erroneous rationalisation is corrected by the user where appropriate — in this case study, by nearly 25%.

6.5.5 Dihydrogen Models

Intuitively, the projection method may be more appropriate than the biggest hole for a certain group of models. These are the models which have dihydrogen as a ligand, rather than distinct hydrogen atoms. The hydrogen atoms are separated by approximately 0.85Å in such models. By its very nature, the biggest hole method will not place the hydride ligands so close together. However the

	Automatic Rationalisation						User Rationalisation					
	Neutron		Xray		Total		Neutron		Xray		Total	
Terminal	24	269 ±38	41	150 ±22	65	194 ±21	27	251 ±35	47	145 ±19	74	184 ±19
Edge Bridging	20	374 ±95	17	210 ±59	37	299 ±59	37	203 ±30	29	155 ±29	66	182 ±21
Face Bridging	2	61 ±2	21	145 ±29	23	138 ±26	8	161 ±26	31	90 ±12	39	104 ±12
Interstitial	1	34 n/a	0	n/a n/a	1	34 n/a	2	41 ±7	1	47 n/a	3	43 ±5
Mixed	1	482 n/a	3	339 ±70	4	375 ±61	1	825 ±1	9	146 ±36	10	214 ±75
Total	48	303 ±45	82	168 ±18	130	218 ±21	75	220 ±22	117	132 ±11	192	166 ±11

(a) Average efficiency applied to all models.

	Automatic Rationalisation						User Rationalisation					
	Neutron		Xray		Total		Neutron		Xray		Total	
Terminal	17	253 ±47	31	148 ±26	48	184 ±24	19	256 ±44	35	142 ±23	54	182 ±23
Edge Bridging	18	355 ±101	14	204 ±71	32	289 ±65	35	199 ±23	27	130 ±24	62	169 ±21
Face Bridging	1	59 n/a	9	114 ±26	10	108 ±24	8	161 ±26	27	100 ±13	35	114 ±12
Interstitial	1	34 n/a	0	n/a n/a	1	34 n/a	2	41 ±7	0	n/a n/a	2	24 ±17
Mixed	1	482 n/a	2	410 ±11	3	434 ±25	1	825 n/a	7	175 ±41	8	256 ±89
Total	38	298 ±54	56	166 ±24	94	218 ±27	65	216 ±24	96	129 ±12	161	164 ±12

(b) Average efficiency applied only to those models for which the Locator succeeded.

***Table 6–9:** A summary of the efficiency (number of iterations) of the Locator.*

projection method will place the two atoms in the same effective position, since they will be equivalent.

This hypothesis was tested by examining eight models with dihydrogen ligands, taking both the projection and the biggest appropriate hole methods to provide an initial estimate. The results of these experiments are shown in Table E–6. In each case only the hydrogen atoms constituting the dihydrogen ligand were optimised. In practice the Locator failed to correctly locate the two hydrogen atoms in the dihydrogen ligand for each model tested, regardless of the initial estimate method. To investigate whether the initial guess methods were inadequate or whether the CCCP code was unable to deal with dihydrogen ligands, the actual locutions as determined by X-ray or neutron diffraction were taken as a starting point for optimisation. However, even with such an accurate starting point, the Locator still failed to successfully locate the hydrogen atoms in the dihydrogen ligand for each of the models.

Consider the accuracy produced by the Locator for each of the methods. Table 6–10 shows a summary of the results for each of the initial estimate methods. It

can be seen that, for these models, using the actual coordinates does not produce more accurate results than either of the initial estimate methods, and that the projection method does not produce more accurate results than the biggest appropriate hole method.

	Automatic Rationalisation			User Rationalisation		
	Neutron	*Xray*	*Total*	*Neutron*	*Xray*	*Total*
Projection	6 4.198 ±2.807	8 2.742 ±2.498	14 3.366 ±2.636	6 4.198 ±2.807	10 3.213 ±2.420	16 3.583 ±2.526
Biggest Hole	6 2.457 ±2.120	8 2.838 ±1.926	14 2.674 ±1.940	6 2.457 ±2.120	10 2.529 ±1.869	16 2.502 ±2.502
Actual				6 8.092 ±7.196	10 5.516 ±6.445	16 6.482 ±6.621

Table 6–10: *A summary of the accuracy, in Å, of the Locator when applied to models with dihydrogen ligands.*

The Locator seems to be inappropriate for hydride ligand location within this group of models. The optimisation code itself rather than the the initial position estimates seems inadequate in some way. By its very nature the MEHMO method involves two body repulsion. The two body repulsion contribution in the CCCP implementation of the MEHMO method seems to be inappropriate for dihydrogen ligands.

6.5.6 HCI Evaluation

The user interface to this tool is itself discussed in Chapter 4. In the evaluation here it is considered more appropriate to discuss the manner in which the interface and data presentation problems of CCCP have been addressed by the Locator. The CCCP interface has been discussed in Chapter 5 from which the key points are summarised here in italics for the purposes of comparison.

1. *Obscurity: There was no explicit relation between the data items in the input file and the associated control action. The user was continually required to cross-check entries with the manual. This resulted in the user not being aware of exactly what could be controlled.*

Several of the improvements made to CCCP reduced the problem of obscurity as described in Chapter 5.

The separation of different types of data into logical categories reduced obscurity. For example, when dealing with a molecular model file, the user will know that the data items therein pertain to the properties of a molecule.

The complete removal of output and experimental control from input file options, being replaced with interactive buttons and menus (see Figure 6–10) made choosing the items to be printed and the experiment to be run more intuitive. For example, the user of CCCP was obliged to edit the appropriate "T" to a "F". However, to determine which character is the appropriate one the user must first refer to the manual. In the Locator the user need only click button beside the "overlap matrix" label in the output control box. This button will then turn from white to black, providing visual feedback to indicate selection. It would seem that the latter method is both more intuitive and less error prone. Furthermore, the user can see immediately what data items can be output and what changes can be made for running the experiment, thus promoting a greater use of the available resources.

2. *Inconsistency: several authors were involved in the development of CCCP. As a result, different styles are evident in the input file. Moreover there are two methods for controlling output without any indication which has priority;* T *(for True) enables some functions yet represses others. The alignment of a data entry within its field is dependent on its type.*

The replacement of all experimental control with interactive switches and buttons and the removal of unnecessary parameters to be stored or calculated internally leaves very little data in the required input file. The specification of each remaining datum is consistent in that it follows one basic rule; it must have an identifying keyword. An on/black button al-

ways means that the facility is turned on in the interactive operations for experimental and output control, and *vice versa.*

3. *Superfluity: The user is required to provide data which would best be provided by the program as defaults.*

   The only data which the user is required to provide are the molecular model and the number of hydride ligands to be located. The former is specified *via* the molecular model input file whilst the latter may be set interactively from the main hydrogen location panel as shown in Figure 6–9, and defaults to 1. All other parameters are stored internally, or may be set interactively through the user interface.

4. *Inflexibility: Parts of the input file enforce an excessively rigid structure on the user. CCCP could crash or produce incorrect results if this was not conformed to.*

   The input file to the Locator has much greater flexibility than that of CCCP. The identification of data items with keywords allows data to appear in any order in the input file. There is no restriction on the spacing of the data within a line. Provided the data is on the same line as the keyword it will be recognised. Some of the data items in the file may have a varying number of fields. For example, the `CELL` item may contain an extra field called `lambda` and the `ATOM` item may have an atom label and/or an element identifier. In either case the Locator will recognise the form of the data and act accordingly. The flexibility of the Locator input file allows data provided by the Cambridge Structural Database to be accepted as a valid input file by the simple addition of keywords.

5. *Lack of logical structure: Different types of data items are intermingled with nothing to separate or distinguish them.*

   The data comprising the CCCP input file were categorised and addressed accordingly:

Molecular Model: This category contains only those items particular to the molecule. These are independent of both the experiment being performed and the program being run. The molecular model is discussed in Chapter 3.

Output and Experimental Control: Items in this category are removed completely from the input file. Each form of control is now controlled interactively within the program, *via* buttons and toggles, as shown in Figure 6–10.

Atomic Parameters: Much as for output and experimental control, items in this category are no longer required as input from the user. Some, such as the atomic parameters, are stored internally. Others, such as the number of atoms, are calculated.

Arranging the data in this manner encourages the user to partition the data being addressed logically. This leads to greater intuitiveness and fewer errors.

6. *Output: Those items selected in the input file are printed at every iteration. This often results in a large output file. It can be difficult to pick out useful information from such a mass of data. The actual information sought — the located hydride positions — is presented as simple coordinates. There is no indication of the relation of the hydride ligands to the rest of the molecule — in particular, their topology within the molecule. Another tool must be employed to visualise the results.*

   The presentation of data by the Locator differs fundamentally to that by CCCP. The output file from CCCP is often large, containing all those matrices which the user explicitly or inadvertently requested. The user must extract the final hydride ligand atom coordinates — the output actually sought — from the CCCP output file. These must then be incorporated into an Ortep or Pluto input file together with the original coordinates for the remainder of the molecule in order to view the results of the hydride ligand location procedure and infer the hydrogen ligand topology.

Furthermore, the user must wait until CCCP has terminated before any iterated ligand position can be viewed. The user will be unaware of any instance were the hydride ligand location procedure is failing, perhaps by being caught in local minimum or leaving the bounds of the molecule, until the program has terminated. This can lead to valuable time being wasted. The Locator, however, gives a three-dimensional view of the molecule which can be manipulated interactively throughout the whole of the hydride ligand location procedure. This is complete with the most recent iterated hydrogen ligand position, and enables the user to recognise both an erroneous convergence early on and act accordingly, and the true ligand topology immediately on the convergence of the iteration — and sometimes before. Furthermore, by default the only output file is a revised molecular model file. Users may explicitly request information to be output to another file by a simple procedure, should this be required. Generally, however, no such extra output should be needed.

7. *General: The user has so much to remember in setting up an input file, so many rules and regulations to follow, that even the more experienced user may find it difficult to create a file for the simplest model, and may not be confident that, in fact, the file is correctly created.*

   The interface to the Locator has been designed so as to be easily and fully utilised by even the most inexperienced novice user.

6.6 Conclusions

The original hydride ligand location program, CCCP, had two shortcomings: an unnecessarily awkward use and a high dependency on initial point(s) provided by the user.

The addition of initial point estimate routines has provided a reasonable initial point for optimisation, for all but the dihydrogen ligand model group. However,

such models constitute a very small group, only eight example could be found the the Cambridge Structural Database. With this feature, the Locator succeeded for most of the 178 models tested and converged with greater accuracy and in fewer iterations.

Further, the incorporation of an appropriately modified version of CCCP into the tool-kit with the accompanying inheritance of all the design issues and objectives of the tool-kit has greatly increased the usability of CCCP.

6.7 Further Work

6.7.1 Efficiency

Several of the matrices calculated in the determination of the overall energy of the molecule are dependent on the interaction of the atoms within the molecule. Since the positions of only a few of these atoms — the hydride ligands — are modified during the energy optimisation, only the interactions with these atoms required recalculation. Essentially, this is a manifestation of "loop invariance" or temporal data coherency. Due to memory restrictions, ICON8, the original program from which CCCP was developed, overwrote many of the matrices dependent on atom/atom positions during the course of the calculation. Thus each had to be recalculated for each iteration. The computational load imposed during convergence by the Locator would be significantly decreased if the data calculated during the first iteration were retained for reuse in subsequent iterations where appropriate. However, the most tidy and time efficient manner to achieve this would be to redesign and implement the program from scratch in a more appropriate language.

6.7.2 Charge Iteration

The parameters describing the molecular atoms are currently read from a file (see Chapter 5, Table 5–3). These parameters describe the energy levels and diffuseness of the *s,p* and *d* orbitals in the atom and determine inter-atom bonding ability. There is some evidence that these parameters vary with the context of an atom's environment. These atomic parameters differ from molecule to molecule due to inter-atomic electron flow in the molecule.

A procedure called *charge iteration* can allow for such molecular context. This produces more realistic molecular energy levels by varying orbital energies of chosen atoms, starting from tabulated values. The charge iteration need usually only be applied to the *metal* atoms in the molecule. This charge iteration procedure could be included in the hydride ligand location program to enhance the accuracy of energy calculations as follows:

Objectives

1. The charge iteration procedure should form an integral part of the hydrogen ligand location program.

2. The charge iteration should be enabled to be optionally selected or deselected. The default would be for charge iteration to be selected.

3. Being an iterative process, the procedure is terminated when the decrease in total molecular energy falls within a certain tolerance limit. This tolerance limit should be user-selectable.

4. Experimental parameters should be allowed to be modified for the current experiment whilst retaining default parameters for further experiments.

5. Defaults should be user-selectable.

6. Charge iteration should be allowed to be applied to elements either serially on in parallel.

The Interface

Some implementation details have already been considered.

A charge iteration button could be appended to the main hydrogen ligand location panel with the three control buttons. This would increase the number of procedures preceding the optimisation of hydrogen ligand energy to four: charge iteration → rationalisation → starting point → optimisation. The button would be governed by the same rules as the others in that the charge iteration procedure will be called whenever;

- the button itself is depressed. In this case *the charge iteration control panel* would open to allow user-customisation or

- any of the subsequent buttons are depressed *and* the charge iteration itself had not been depressed before-hand. In this case the charge iteration panel would not open.

Figure 6–17 shows a possible design for the interface. The modified hydrogen ligand location panel is central. The results of pressing the *charge iteration* button are shown under two different circumstances to the left and right — without and with charge iteration respectively. From top to bottom, the display shows two sections for the selection of the charge iteration facility and the tolerance limit. The first selects the parameters particular to the current experiment. The second is for the defaults where any changes made would be persistent. The remainder of the screen would be blanked if the charge iteration button for the current experiment were turned off, as shown on the left hand side of the diagram. A third section lists those transition metal elements present in the molecule. It is uncommon for one transition metal cluster compound to comprise more than three or four different transition metal atoms. There need

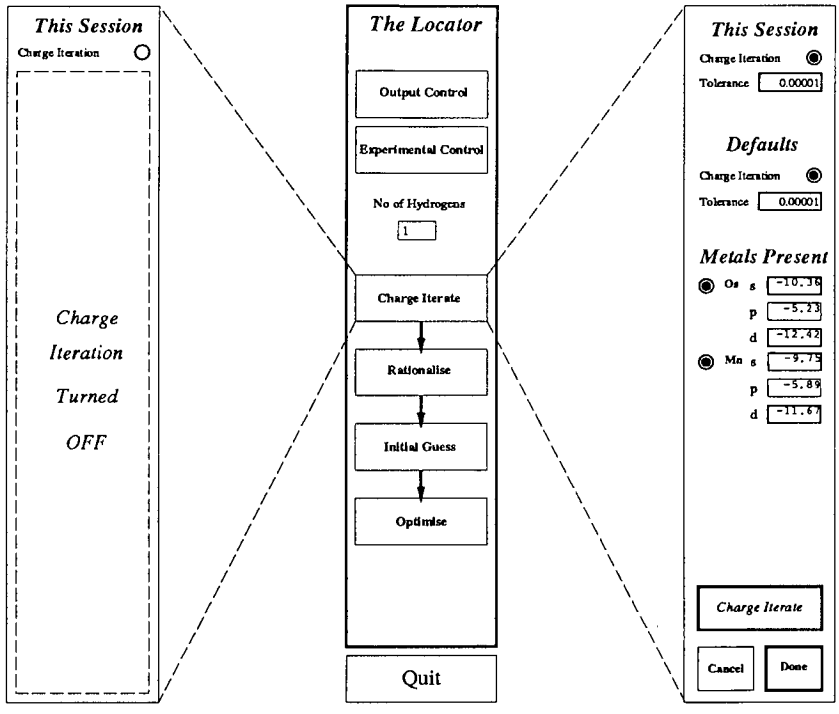


Figure 6–17: A possible control panel for the charge iteration process.

only be space left to accommodate about this number of atoms. Each metal atom is listed with a toggle button to the left, allowing selection of the elements for charge iteration. To the right the *s*, *p* and *d* orbital energies are listed in *asciiText* widgets [60]. The values displayed in these widgets would be updated on charge iteration as activated by the depression of the *charge iteration* button. The most recently iterated values would be displayed in a similar manner to the molecular display in the main tool. If required, particular energy values could be specified manually by direct entry into the asciiText widgets and depressing the *done* button.

6.7.3 Tuning of the orbital energy and exponent of the metal-bound hydride ligand(s)

Two parameters in the energy optimisation — the hydride orbital energy and exponent — were originally tuned by Mitchell [13] using $MnH(CO)_5$. This

resulted in the acceptance of an orbital energy of -12.6eV and exponent of $0.85au^{-1}$. An attempt was made to reproduce these results. The hydride ligand was located for this molecule with energy values ranging over -6.0eV → -14.0ev and exponent values ranging over $0.6au^{-1}$ → $1.4au^{-1}$. The distance of the located value from the actual (neutron diffraction) value was plotted for each input energy/exponent pair in a three dimensional graph with GnuPlot[5] as shown in Figure 6–18.

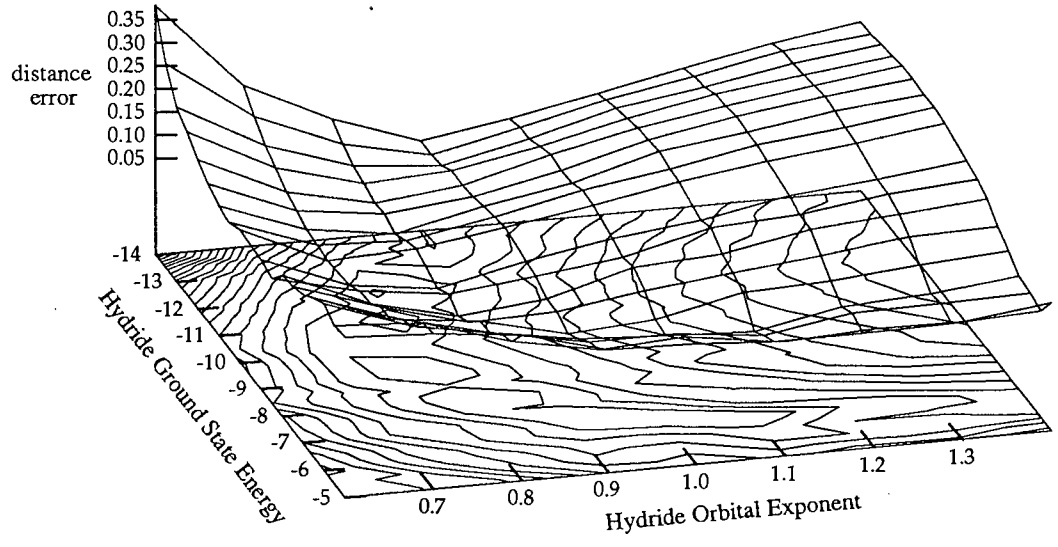


Figure 6–18: Tuning the Hydride orbital energy and exponent using $MnH(CO)_5$.

There is no clearly-defined explicit minimum "well" in the figure. The optimal values for the parameters appear to be mutually dependent, with the graph surface describing a valley. It is difficult to extract globally optimal energy and exponent values from this plot. The global minimum distance error of 0.0191Å occurs at the energy/exponent pair $(-9.0ev, 0.8au^{-1})$. However, this is only 0.0001Å less than the distance error of 0.0192Å at $(-8.0ev, 0.8au^{-1})$. The data seem inappropriate for the determination of optimal exponent and energy

[5]GnuPlot is a graph-plotting program produced by the Free Software Foundation

	exponent								
energy	0.6	0.7	**0.8**	0.9	1.0	1.1	1.2	1.3	1.4
-5.0	0.1819	0.1726	0.1530	0.1034	0.0972	0.0835	0.0943	0.0927	0.1141
-5.5	0.1796	0.1667	0.1087	0.0737	0.0842	0.0538	0.0552	0.0582	0.0680
-6.0	0.1970	0.1286	0.0957	0.0610	0.0616	0.0395	0.0701	0.0730	0.0779
-6.5	0.1807	0.1170	0.0683	0.0380	0.0284	0.0185	0.0542	0.0636	0.0846
-7.0	0.1584	0.0865	0.0659	0.0199	0.0332	0.0749	0.0809	0.0992	0.1107
-7.5	0.1461	0.0862	0.0330	0.0385	0.0680	0.0879	0.1066	0.1316	0.1358
-8.0	0.1414	0.0625	**0.0192**	0.0373	0.0872	0.0996	0.1359	0.1544	0.1692
-8.5	0.1057	0.0605	0.0381	0.0588	0.0827	0.1277	0.1414	0.1680	0.1829
-9.0	0.1172	0.0442	**0.0191**	0.0606	0.0928	0.1154	0.1534	0.1706	0.1919
-10.0	0.1254	0.0397	0.0478	0.0776	0.1097	0.1423	0.1687	0.1901	0.2156
-10.5	0.1124	0.0456	0.0364	0.0735	0.1129	0.1412	0.1760	0.1958	0.2211
-11.0	0.1225	0.0514	0.0245	0.0673	0.1000	0.1499	0.1742	0.2019	0.2199
-11.5	0.1415	0.0644	0.0402	0.0589	0.0986	0.1458	0.1796	0.2039	0.2249
-12.0	0.1646	0.0770	0.0331	0.0547	0.1024	0.1421	0.1720	0.1985	0.2247
-12.5	0.1938	0.0985	0.0477	0.0585	0.0927	0.1348	0.1694	0.1934	0.2156
-13.0	0.2209	0.1180	0.0572	0.0395	0.0785	0.1276	0.1559	0.1875	0.2114
-13.5	0.2845	0.1519	0.0739	0.0335	0.0757	0.1162	0.1491	0.1854	0.2053
-14.0	0.3801	0.1896	0.0995	0.0351	0.0637	0.1015	0.1389	0.1725	0.2029

Table 6–11: *Tuning the Hydride orbital energy and exponent using $MnH(CO)_5$, the value given is the distance, in Å, of the located hydride from that found by neutron diffraction.*

values to application to general models. It was concluded that an alternative molecular model should be sought for this tuning in order to gain more definite results. To appreciate why these results do not concur with those of Mitchell, it should be borne in mind that:

1. it is unknown which atomic parameters were used by Mitchell, although it is suspected that they are the same as in this experiment.

2. the starting point of Mitchell's experiment is unknown. It is suspected that the neutron position was taken. In practice however, the numerical location of hydrogen ligands is undertaken specifically when this data is *not* available. Since the initial position estimate provided by the *biggest appropriate hole* method is that recommended when using the Locator, this position was thought to be the more appropriate from which to undertake this tuning.

An ideal model for tuning the parameters should have a single transition metal atom from the second or third row of the Periodic Table and a single terminal hydride ligand. Metals from the second or third row transition metals are preferable since their orbital energies and exponents are generally regarded as better defined. The metal hydride $IrH(PH_3)_3Cl_2$ which models $IrH((PH_3)P)_3Cl_2$ was selected for the tuning.

Figure 6–19 shows that this model proves yet more unsatisfactory. The surface describes a gentle curve with a high border around the edge which is ineffectual in providing globally optimal values for the energy and exponent.

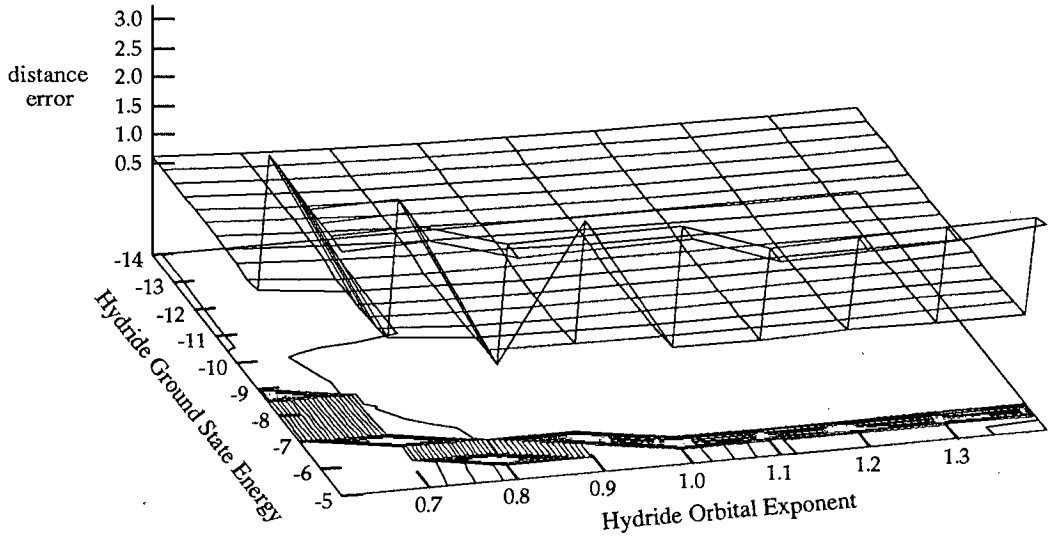


Figure 6–19: *Tuning the Hydride orbital energy and exponent using $IrH(PH_3)_3Cl_2$.*

It is recommended that these parameters should be tuned more effectively.

These parameters may depend significantly on the type of metal atom(s) to which the hydride ligand is bonded. It may be possible to take this into account when tuning these parameters. Several options lie open for tuning the parameters:

	exponent								
energy	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4
-5.0	3.1410	3.1710	2.7391	2.7696	2.7885	2.2517	2.2912	2.3431	2.4713
-5.5	3.1752	3.1950	2.7441	2.7669	2.7809	2.2526	2.2976	2.3483	2.3773
-6.0	3.1707	3.2194	0.4306	2.7748	0.4367	0.4240	0.4559	0.4292	0.4415
-6.5	3.1870	3.2238	0.4516	0.4098	0.4045	0.4056	0.4015	0.4178	0.4139
-7.0	3.1924	0.5629	0.4389	0.3917	0.3717	0.3777	0.3729	0.3993	0.3938
-7.5	3.1948	0.4997	0.3915	0.3754	0.3734	0.3717	0.3566	0.3657	0.3646
-8.0	3.2151	0.4745	0.4026	0.3681	0.3639	0.3493	0.3680	0.3701	0.3529
-8.5	3.2101	0.4425	0.3929	0.3533	0.3605	0.3513	0.3466	0.3287	0.3601
-9.0	0.5908	0.4421	0.3847	0.3492	0.3501	0.3399	0.3412	0.3362	0.3425
-9.5	0.5444	0.4125	0.3806	0.3620	0.3453	0.3463	0.3460	0.3344	0.3441
-10.0	0.6133	0.5266	0.4624	0.4185	0.3774	0.3603	0.3419	0.3368	0.3326
-10.5	0.6133	0.5266	0.4624	0.4185	0.3774	0.3603	0.3419	0.3368	0.3326
-11.0	0.6133	0.5266	0.4624	0.4185	0.3774	0.3603	0.3419	0.3368	0.3326
-11.5	0.6133	0.5266	0.4624	0.4185	0.3774	0.3603	0.3419	0.3368	0.3326
-12.0	0.6133	0.5266	0.4624	0.4185	0.3774	0.3603	0.3419	0.3368	0.3326
-12.5	0.6133	0.5266	0.4624	0.4185	0.3774	0.3603	0.3419	0.3368	0.3326
-13.0	0.6133	0.5266	0.4624	0.4185	0.3774	0.3603	0.3419	0.3368	0.3326
-13.5	0.6133	0.5266	0.4624	0.4185	0.3774	0.3603	0.3419	0.3368	0.3326
-14.0	0.6133	0.5266	0.4624	0.4185	0.3774	0.3603	0.3419	0.3368	0.3326

Table 6–12: *Tuning the Hydride orbital energy and exponent using $IrH(PH_3)_3Cl_2$, the value given is the distance, in Å, of the located hydride from that found by neutron diffraction.*

- finding an ideal molecular model which produces a well defined minimum.

- examining several non-ideal models from different groups, preferably with metals from the second or third row of the Periodic Table, to establish a relationship between groups with respect to the influence exerted on the hydride ligands by the metals to which they are bonded.

- using accepted, well-researched knowledge of atomic trends to calibrate the effect of different metal atoms on the hydride ligand.

- a combination of all or some of the above, gaining as much knowledge on as many transition metal hydride ligands as possible and applying these data to their full advantage.

6.7.4 Dihydrogen Models

It is thought that the reason that the Locator fails to correctly locate dihydrogen ligands may be due to the nuclear repulsion between the two hydrogen atoms within the dihydrogen ligand being too great. It would be useful for the user to have control over the extent of nuclear repulsion between these two atoms when it is suspected that a ligand might be dihydrogen.

6.7.5 The Failure of the Locator for Some Models

Looking more closely at the models for which the Locator fails it appears that it is particularly unsuccessful for models with *chloride* ligands. Models with such ligands comprise 6% of the models tested. However, of the models for which the Locator fails, those with chloride ligands comprise 19%, or 26% after user rationalisation. In fact, the success rate for the Locator when considering only models with such ligands was only 36%, or 43% with user intervention at the rationalisation stage. This is compared with the general figures of 72% and 84%.

It is proposed that the general success rate of the Locator may be improved further by looking more closely at the models for which the Locator does fail in an attempt to identify common features which may be causing the failure of the Locator.

Chapter 7

Conclusions and Further Work

7.1 Conclusions

It would seem that the overall success of this project would be appropriately evaluated by considering the extent to which the initial objective of this project has been achieved. The specific points of this objective as presented in Chapter 1 are addressed in turn. The first part of the objective was:

> *"To design a base for an integrated tool set and illustrate the tool set by implementing some example tools to be constituents of the set."*

This thesis has described the design of an integrated tool kit in Chapter 3, and the implementation of three tools therein. The principle of these tools was the Locator, a tool which exploits atomic spatial constraints, and performs molecular orbital calculations and Simplex optimisation to locate hydride ligands in transition metal cluster compounds. The development of the Locator and the testing thereof are discussed in Chapter 6.

Although the original hydride ligand location procedure was shown to be highly dependent on the initial point(s) given by the user, no method was recommended by the author of CCCP for calculating such points. Thus, two methods —

the projection and the biggest appropriate hole methods — were implemented which attempted to apply chemical knowledge to the formulation of an initial estimate of location(s) of the sought hydride ligand(s). This knowledge comprised the known electronic requirements and the spatial distribution of atoms, and the bonding interactions within the molecular model. In order to keep computation time to a minimum, both methods consider only those atoms likely to have a bearing on the position of the hydride ligand(s), the "inner core" set. The estimated hydride ligand location(s) were then taken as starting point(s) for subsequent optimisation by the original CCCP code.

The two implemented initial estimate methods were initially tested against each other in order to determine which, if either, of the two methods was the superior overall, and which was more appropriate for the different types of hydride ligands. Since the biggest appropriate hole method proved to be more successful than the projection method, the former was tested against CCCP with *ad hoc* estimates of starting point(s). The results of this test showed that the addition of a routine to estimate initial hydride ligand location(s) led to an improved success rate of 92%, succeeding for 12 of the 13 models tested as opposed to 23% and 69% for the two *ad hoc* initial estimate methods used for CCCP. Further, the Locator proved to give greater accuracy and sometimes efficiency than CCCP with *ad hoc* initial estimates.

The Locator was then extensively on tested 178 models. It was shown to be reasonably successful, with a success rate of at least 72% and an average accuracy of 0.261 $\pm$ 0.031Å, rising to 84% and 0.256 $\pm$ 0.009Å respectively with user intervention at the rationalisation stage. The term "at least" is adopted here since the figures quoted are for X-ray and neutron diffraction models combined. As was discussed in Chapters 5 and 6 the hydride ligand coordinates determined by X-ray diffraction are not as reliable as those determined by neutron diffraction. Further, as discussed in Chapter 6, some of the *topological* descriptions of the located hydrides may be disputed. When considering only the 60 neutron models — assessed in 75 experiments — the success rate of the Locator is 79%, and the

accuracy rate 0.210 ± 0.028Å, rising to 87% and 0.236 ± 0.013Å respectively on user intervention at the rationalisation stage.

There is, however, one small group of models (only eight examples could be found on the Cambridge Structural Database) for which the Locator was particularly unsuccessful — models with dihydrogen ligands. It would appear that both the initial estimate methods and the actual CCCP code itself are inappropriate for the location of such ligands. The biggest appropriate hole initial estimate method, by its very nature, will not place the atoms in the dihydrogen ligand sufficiently close together. It was thought that the projection method might be more suitable for such ligands, but testing showed otherwise. The CCCP code itself was found to be unsuitable since the Locator was unsuccessful for all eight appropriate models even when the actual positions of the atoms in the dihydrogen ligand were taken as starting points. Suggestions are made in the "Future Work" section of Chapter 6 as to how this situation could be improved.

The addition of the initial estimate routine has demonstrated that, although the knowledge and experience of a chemist may be too complex to be completely encapsulated, elements thereof may be successfully applied. The use of encapsulated chemical knowledge rather that full automation, which would be both impracticable and arguably undesirable, may provide valuable *support* for the research chemist, providing information to the chemist in the decision making process.

The other two tools discussed during the course of this thesis are The Orthogonalisor and the Visualisor. The Orthogonalisor is a simple utility tool to convert a model from fractional coordinates to orthogonalised Ångström coordinates. The Visualisor is a molecular visualisation tool and is discussed in Chapter 4.

Although the Visualisor is in itself a useful molecular visualisation tool, its most valuable feature is proposed to be the provision of a standard visualisation interface for adoption by other tools dealing with molecular models. This has been illustrated in Chapter 6 which documents how the Visualisor is adopted

as an interface by the Locator. This allowed the user to monitor the progress of the hydride ligand location and intervene when the location appeared to be failing rather than wait until the termination of the optimisation routine before visualising the results, as was previously the case with CCCP. Further applications for which such an interface might be suitable are discussed later in this chapter.

> *"The tool set would be designed for general use by any scientist, but for the purposes of this research the science of chemistry, specifically in the Chemistry Department at the University of Edinburgh, would be addressed as a case study."*

An analysis of user requirements was undertaken to investigate the current state of software support for chemists at Edinburgh University, through the distribution of a questionnaire as discussed in Chapter 2. Although the success of the questionnaire was limited, light was shed on some areas relevant to this project. From the questionnaire responses an approximate picture of the working habits of a chemist was constructed, resulting in a user model. Furthermore, a list of requirements could be constructed for future software support. The responses to this questionnaire together with personal experience led to both the decision to implement the tools described above and the identification of appropriate objectives.

> *"The tool-set would be designed to bring together latest techniques in computer science* and *chemistry."*

The implementation of the Locator demonstrated the use of chemical knowledge in the initial estimate routine. The electron requirements of each of the transition metal atoms in the molecule, the calculated bonding interactions therein and the spatial distribution of atoms each contribute to the calculation of the initial estimate of the hydride ligand location(s). Modified Extended Hückel Molecular Orbital *(MEHMO)* calculations and Simplex optimisation are then used locate the most energetically stable position for the hydride ligand(s)

A study was made of current human-computer interaction and software engineering techniques in Chapter 3. The application of those techniques considered relevant to this project were discussed.

> *"This would promote the design of tools with maximum utility to chemists, which allow them to concentrate on the work in hand, with minimum attention to the tools they use to complete their task."*

It is considered that the addition of the initial estimate to the CCCP code in the Locator implementation significantly increased the usability of the tool in two major areas: the user is no longer required to provide an initial estimate and can be more confident that the final results produced are correct. Further, at each stage of both the initial estimate and the subsequent optimisation, all relevant data are presented to the user in a form which may be readily assimilated. This allows the user to monitor the progress of the Locator and facilitates decision-making such as correcting inaccurate rationalisation or the position of the initial estimate where appropriate. The Locator was demonstrated at a chemistry conference[1] where the interface and ease of use of this tool received favourable feedback from several delegates.

The application of the HCI principles discussed in Chapter 3, such as direct manipulation, intuitiveness and consistency, would appear to have greatly enhanced the usability of the tools within the tool-kit. The tools are designed to ensure that users neither have any need to consult complex and extensive manuals before using them, nor are required to use a tool many times before becoming confident, or even competent, in their use. It is hoped that users need not even be confident in the use of a computer before exploiting the tools within this tool-kit; that they are exactly that — tools, which may be "picked up" and used whenever needed much as one would use a hammer or a screwdriver.

---

[1]The International Conference on the Chemistry of the Copper and Zinc Triads at Edinburgh in July 1992

Furthermore, the adoption of the software engineering principles has aided the creation of a software base on which other tools may be built. The modular design of the tools and demonstration of resource re-use will aid any subsequent adaptations or extensions of the tools or the tool-kit as a whole.

7.2 Further Work

7.2.1 Improvements to Current Tools

Several ideas for future work have been described in Chapters 4 and 6. These are summarised below.

The Locator

Suggestions were made for further development of the Locator. The first suggestion would improve the efficiency, *i.e.* the speed of execution. The others are potential improvements to its success and accuracy.

- The efficiency of the Locator could be improved by the re-use or "caching" of data calculated in the first iteration to be used in subsequent iterations.
- The inclusion of a *charge iteration* procedure is discussed in Chapter 6. This would produce more appropriate molecular energy levels in the metal atoms by allowing for the environment in which the metal atoms lie.
- The molecular model $MnH(CO)_5$ was originally used for the tuning of the orbital energy and exponent parameters. However, this model appeared to be less than ideal since a plot of the accuracy of the final position of the hydride ligand against over a range of values for these two parameters did not produce a well-defined minimum. It is recommended that these parameters be tuned more effectively with a more appropriate molecular model.

- There are still some models with terminal, edge-bridging, face-bridging and interstitial hydride ligands for which the Locator fails. It is considered that the success rate of the Locator could be improved generally by attempting to identify the reason(s) for the failure of the Locator to locate the hydride ligands in such models. One approach would be to look for common features in the molecular models for which the Locator does fail. The high failure rate for models containing chloride ligands was highlighted in Chapter 6, and may provide one area for investigation.
- The Locator is singularly unsuccessful at locating dihydrogen ligands. An investigation for the failure of the Locator in locating such ligands is recommended.

The Visualisor

Suggestions for further development of the Visualisor are summarised below:

- the removal of atom number restrictions,
- improvement of the display by the inclusion of
  - perspective,
  - variable atom and bond sizes,
  - movable atom labels and
  - alternative representations of molecular models.
- improvement of the PostScript output from the Visualisor and
- the inclusion of keyboard shortcuts for the operation of the Visualisor.

In addition to the ideas for future work detailed in previous chapters there are several possibilities for further developments of the tool-kit itself:

7.2.2 A Construction Tool

Some applications which could be readily incorporated into the tool-kit are identified in this section. The Visualisor is a clear candidate as an interface to any tool in which molecular models play a large part. However, tools in which molecular models play only a minor rôle could adopt the Visualisor to display the molecular model data with suitable additions for the display of other data such as energy plots. One such tool could be a theoretical molecular construction tool. This would allow the user to build a molecular model from sub-units of atoms and functional groups, or simply edit an existing model by the substitution, deletion or addition of sub-units. The Visualisor could be adopted in a similar manner as for the Locator. The user could control the construction of the model through the simple addition of a tool-specific control panel on the side panel of the Visualisor. The existing tool CALC available to chemists in the Department of Chemistry supports molecular construction amongst other functions, and could provide the basis for a tool. CALC is discussed in more detail in Chapter 2. Currently this tool accepts molecular models and textual commands. The addition of a graphical interface to the relevant part of this tool should be straight-forward.

7.2.3 Future Developments to the Tool-Kit

The tool kit is written in a modular form with a well defined application procedural interface (*API*)[2]. This modularity could be taken to its logical extreme by converting the implementation of the tool kit to a language which actively encourages modularity and code re-use rather than merely allowing it, thus further enhancing portability, wide re-use and extensibility. Furthermore, the adoption of a language which encourages modularity and code re-use will ensure that further developments remain true to the original objectives applied to

[2]The API is the set of procedures providing the interface between the application and the tool-kit.

this project which arose from computer science issues. *Object-oriented* languages would allow just this. Since the implementation is currently in C it would seem logical to choose the C++ language if this conversion were to be undertaken. ANSI C is almost a subset of C++ and the two have many structures in common. It is considered that this translation would not involve a great amount of work. The benefits of objected-oriented programming are well documented.

The tool-kit currently runs under the X Window system, a standard windowing system available on many machines. Any machine running the X Window system may also run the tool-kit, thereby ensuring a reasonable level of portability for the tool-kit. However, for even greater portability each call to the host windowing system is abstracted to a dedicated "driver" routine. For example, an associated driver routine is called whenever a "toggle widget" is required, which then calls the appropriate X Window system routine. This indirection would simplify a port of the tool-kit to a window system other than the X Window system, such as *MS-Windows* system for PCs. Such a port would only require the driver routines which currently call the X Window system functions to be reimplemented to call equivalent Windows functions instead. The code need only be modified at a single point, rather than repeatedly for each call. It must be remembered however that a bitmapped screen and pointing device are still requirements; this must be taken into consideration when porting to another system.

7.2.4 Broadening the Scope of Applications

Although the original project was directed at chemical research at Edinburgh it was only meant to be a *case study*. Thus, it would be worthwhile to apply the methods used in this project to other scientific subjects, *e.g.* physics or biology. The first step would be to repeat a user requirements analysis, since the work practices, and existing levels of software support, would be unlikely to be the same in all subjects. Moreover, similar studies in academic chemistry departments other than Edinburgh may also be worthwhile since the current

state of affairs is unlikely to be the same in all. The lessons learned in this project, particularly in relation to the design of the questionnaire in the user requirements analysis, could be applied in any subsequent studies.

Appendix A

Questionnaire

This appendix contains a blank copy of the questionnaire described in Chapter 2.

Questionnaire

Section 1

...for everyone

Name .....

Research Topic .....

.....

.....

1. How do you spend your working day? e.g.

bench work	..... hrs
calculations	..... hrs
running spectra	..... hrs
.....	..... hrs
.....	..... hrs
.....	..... hrs
.....	..... hrs
.....	..... hrs
.....	..... hrs
.....	..... hrs

2. Are there any of the activities mentioned above which could possibly be benefited by using a computer? (*such as a) the comparison of spectra with theoretical models or from a library of spectra, b) complex mathematical calculations, c) storage of information etc.*)

.....

.....

.....

.....

3. Are there any other facilities (such as those listed above) you would like to see provided on a central computer?

.....

.....

.....

.....

4. If you use a personal computer...

   (a) What kind is it?.....

   (b) What do you use it for?.....

   (c) What programs do you run (and what is the purpose of each) .....

   .....

   .....

   .....

   .....

   .....

Please use the rest of this page (and any more if necessary) to list any relevant comments etc.

Section 2

... for people who use the on-line Library Catalogues

1. How often do you use each of the on-line library catalogues?
   (a) never
   (b) you have used them once or twice
   (c) when you have difficulty locating material
   (d) whenever you are looking for material in the library

   EULCAT a b c d GEAC a b c d

2. How would you describe the use of each facility?
   (a) not applicable
   (b) difficult
   (c) tedious
   (d) reasonable
   (e) simple

   EULCAT a b c d e GEAC a b c d e

3. What do you think of the standard of the two systems?
   (a) not applicable
   (b) bad
   (c) could be improved
   (d) reasonable
   (e) good

   EULCAT a b c d e GEAC a b c d e

   If you have any specific complaints, general comments, suggestions for improvement or ideas for any new features which you would like to see included in any of the above facilities please list them of one of the sheets provided

4. If you have experience of both catalogue systems which do you prefer?

   .....

   Please give reasons for your preference (advantages and disadvantages of each)

   .....

   .....

   .....

   .....

Section 3

... for people who use *EMAS*.

Username(s) .....

1. How often do you use *EMAS*? *(On Average)*

   (a) At least once a week
   (b) A few days a month
   (c) About once a month
   (d) Less than once a month

2. For what purposes do you find *EMAS* most useful?

   .....

   .....

   .....

   .....

   The remaining questions in this section are concerned with actual facilities and programs provided on *EMAS*. The most common facilities are listed, but please use the space provided to include any you may use, but which are not listed, in the space provided.

3. How often do you use each of the following facilities?

   (a) Never
   (b) Rarely
   (c) Quite often
   (d) Regularly

CALC	a	b	c	d	Alert	a	b	c	d
CCCP	a	b	c	d	Batch	a	b	c	d
ICON	a	b	c	d	Calculator	a	b	c	d
SHELX	a	b	c	d	Chat	a	b	c	d
EASYGRAPH	a	b	c	d	Electronic Mail	a	b	c	d
NOTICE	a	b	c	d	Epoch	a	b	c	d
ORTEP	a	b	c	d	Help	a	b	c	d
PLUTO	a	b	c	d	Kermit	a	b	c	d
ECCE	a	b	c	d	Tell	a	b	c	d
EDIT	a	b	c	d	X-Talk	a	b	c	d
EMACS	a	b	c	d	.....	a	b	c	d
VECCE	a	b	c	d	.....	a	b	c	d
SCRIBE	a	b	c	d	.....	a	b	c	d
.....	a	b	c	d	.....	a	b	c	d

4. How long did it take you to become competent in each of the facilities?

   (a) not applicable
   (b) you are not yet competent
   (c) a month or less
   (d) a week or less
   (e) a day or less

CALC	a	b	c	d	e	Alert	a	b	c	d	e
CCCP	a	b	c	d	e	Batch	a	b	c	d	e
ICON	a	b	c	d	e	Calculator	a	b	c	d	e
SHELX	a	b	c	d	e	Chat	a	b	c	d	e
EASYGRAPH	a	b	c	d	e	Electronic Mail	a	b	c	d	e
NOTICE	a	b	c	d	e	Epoch	a	b	c	d	e
ORTEP	a	b	c	d	e	Help	a	b	c	d	e
PLUTO	a	b	c	d	e	Kermit	a	b	c	d	e
ECCE	a	b	c	d	e	Tell	a	b	c	d	e
EDIT	a	b	c	d	e	X-Talk	a	b	c	d	e
EMACS	a	b	c	d	e	.....	a	b	c	d	e
VECCE	a	b	c	d	e	.....	a	b	c	d	e
SCRIBE	a	b	c	d	e	.....	a	b	c	d	e
.....	a	b	c	d	e	.....	a	b	c	d	e

5. How would you describe the use of each facility?

   (a) not applicable
   (b) difficult
   (c) tedious
   (d) reasonable
   (e) simple

CALC	a	b	c	d	e	Alert	a	b	c	d	e
CCCP	a	b	c	d	e	Batch	a	b	c	d	e
ICON	a	b	c	d	e	Calculator	a	b	c	d	e
SHELX	a	b	c	d	e	Chat	a	b	c	d	e
EASYGRAPH	a	b	c	d	e	Electronic Mail	a	b	c	d	e
NOTICE	a	b	c	d	e	Epoch	a	b	c	d	e
ORTEP	a	b	c	d	e	Help	a	b	c	d	e
PLUTO	a	b	c	d	e	Kermit	a	b	c	d	e
ECCE	a	b	c	d	e	Tell	a	b	c	d	e
EDIT	a	b	c	d	e	X-Talk	a	b	c	d	e
EMACS	a	b	c	d	e	.....	a	b	c	d	e
VECCE	a	b	c	d	e	.....	a	b	c	d	e
SCRIBE	a	b	c	d	e	.....	a	b	c	d	e
.....	a	b	c	d	e	.....	a	b	c	d	e

6. How did you find out about this facility?

   (a) not applicable

   (b) from fellow students or member of staff?

   (c) from *User Support*?

   (d) from the *EMAS* user guide?

   (e) other *please state* .....

CALC	a	b	c	d	e	Alert	a	b	c	d	e
CCCP	a	b	c	d	e	Batch	a	b	c	d	e
ICON	a	b	c	d	e	Calculator	a	b	c	d	e
SHELX	a	b	c	d	e	Chat	a	b	c	d	e
EASYGRAPH	a	b	c	d	e	Electronic Mail	a	b	c	d	e
NOTICE	a	b	c	d	e	Epoch	a	b	c	d	e
ORTEP	a	b	c	d	e	Help	a	b	c	d	e
PLUTO	a	b	c	d	e	Kermit	a	b	c	d	e
ECCE	a	b	c	d	e	Tell	a	b	c	d	e
EDIT	a	b	c	d	e	X-Talk	a	b	c	d	e
EMACS	a	b	c	d	e	.....	a	b	c	d	e
VECCE	a	b	c	d	e	.....	a	b	c	d	e
SCRIBE	a	b	c	d	e	.....	a	b	c	d	e
.....	a	b	c	d	e	.....	a	b	c	d	e

7. If you have never used the facility or only use it rarely why do you not use it more regularly?

   (a) you have never heard of the facility

   (b) you have never had need to use it

   (c) you find it difficult and tedious to use

   (d) you don't know how to use the facility

   (e) other *please state* .....

CALC	a	b	c	d	e	Alert	a	b	c	d	e
CCCP	a	b	c	d	e	Batch	a	b	c	d	e
ICON	a	b	c	d	e	Calculator	a	b	c	d	e
SHELX	a	b	c	d	e	Chat	a	b	c	d	e
EASYGRAPH	a	b	c	d	e	Electronic Mail	a	b	c	d	e
NOTICE	a	b	c	d	e	Epoch	a	b	c	d	e
ORTEP	a	b	c	d	e	Help	a	b	c	d	e
PLUTO	a	b	c	d	e	Kermit	a	b	c	d	e
ECCE	a	b	c	d	e	Tell	a	b	c	d	e
EDIT	a	b	c	d	e	X-Talk	a	b	c	d	e
EMACS	a	b	c	d	e	.....	a	b	c	d	e
VECCE	a	b	c	d	e	.....	a	b	c	d	e
SCRIBE	a	b	c	d	e	.....	a	b	c	d	e
.....	a	b	c	d	e	.....	a	b	c	d	e

8. What do think of the *standard* of *information* available regarding the facility?

   (a) not applicable
   (b) bad
   (c) could be improved
   (d) reasonable
   (e) good

CALC	a	b	c	d	e	Alert	a	b	c	d	e
CCCP	a	b	c	d	e	Batch	a	b	c	d	e
ICON	a	b	c	d	e	Calculator	a	b	c	d	e
SHELX	a	b	c	d	e	Chat	a	b	c	d	e
EASYGRAPH	a	b	c	d	e	Electronic Mail	a	b	c	d	e
NOTICE	a	b	c	d	e	Epoch	a	b	c	d	e
ORTEP	a	b	c	d	e	Help	a	b	c	d	e
PLUTO	a	b	c	d	e	Kermit	a	b	c	d	e
ECCE	a	b	c	d	e	Tell	a	b	c	d	e
EDIT	a	b	c	d	e	X-Talk	a	b	c	d	e
EMACS	a	b	c	d	e	.....	a	b	c	d	e
VECCE	a	b	c	d	e	.....	a	b	c	d	e
SCRIBE	a	b	c	d	e	.....	a	b	c	d	e
.....	a	b	c	d	e	.....	a	b	c	d	e

Please give any specific complaints, comments or suggestions for improvement on one of the sheets provided

9. What do think of the *availability* of *information* regarding the facility?

   (a) not applicable
   (b) bad
   (c) could be improved
   (d) reasonable
   (e) good

CALC	a	b	c	d	e	Alert	a	b	c	d	e
CCCP	a	b	c	d	e	Batch	a	b	c	d	e
ICON	a	b	c	d	e	Calculator	a	b	c	d	e
SHELX	a	b	c	d	e	Chat	a	b	c	d	e
EASYGRAPH	a	b	c	d	e	Electronic Mail	a	b	c	d	e
NOTICE	a	b	c	d	e	Epoch	a	b	c	d	e
ORTEP	a	b	c	d	e	Help	a	b	c	d	e
PLUTO	a	b	c	d	e	Kermit	a	b	c	d	e
ECCE	a	b	c	d	e	Tell	a	b	c	d	e
EDIT	a	b	c	d	e	X-Talk	a	b	c	d	e
EMACS	a	b	c	d	e	.....	a	b	c	d	e
VECCE	a	b	c	d	e	.....	a	b	c	d	e
SCRIBE	a	b	c	d	e	.....	a	b	c	d	e
.....	a	b	c	d	e	.....	a	b	c	d	e

Please give any specific complaints, comments or suggestions for improvement on one of the sheets provided

10. What do think of the *standard* the facility *itself*?

    (a) not applicable
    (b) bad
    (c) could be improved
    (d) reasonable
    (e) good

CALC	a	b	c	d	e	Alert	a	b	c	d	e
CCCP	a	b	c	d	e	Batch	a	b	c	d	e
ICON	a	b	c	d	e	Calculator	a	b	c	d	e
SHELX	a	b	c	d	e	Chat	a	b	c	d	e
EASYGRAPH	a	b	c	d	e	Electronic Mail	a	b	c	d	e
NOTICE	a	b	c	d	e	Epoch	a	b	c	d	e
ORTEP	a	b	c	d	e	Help	a	b	c	d	e
PLUTO	a	b	c	d	e	Kermit	a	b	c	d	e
ECCE	a	b	c	d	e	Tell	a	b	c	d	e
EDIT	a	b	c	d	e	X-Talk	a	b	c	d	e
EMACS	a	b	c	d	e	.....	a	b	c	d	e
VECCE	a	b	c	d	e	.....	a	b	c	d	e
SCRIBE	a	b	c	d	e	.....	a	b	c	d	e
.....	a	b	c	d	e	.....	a	b	c	d	e

If you have any specific complaints, general comments, suggestions for improvement or ideas for any new features which you would like to see included in any of the above facilities please list them of one of the sheets provided

Supplementary sheet

for Section 3, Questions 8 and 9

(facility) .....

Comments about documentation or information about facility.

for Section 3, Question 10

(facility) .....

Comments about or suggestions for improvement of facility.

Appendix B

Questionnaire Results

This appendix contains the raw results of the questionnaire discussed in Chapter 2.

B.1 Section 1 — Working practices

All 35 respondents filled in this part of the questionnaire.

1. How do you spend your working day?

The responses to this question are summarised in Table B–1

2. Are there any of the activities mentioned above which could possibly be benefited by using a computer? ***(such as a) the comparison of spectra with theoretical models or from a library of spectra, b) complex mathematical calculations, c) storage of information etc.)***

"Literature Searches (already used), anything to do with availability of a much greater level of professional secretarial services — at present all report typing is done privately outside." *Respondent 1*

"Calculations and graphics." *Respondent 2*

"Yes — mainly greater integration of structure solving / structure refinement / presentation of results / preparation of reports & papers. Also better on-line documentation of local programs." *Respondent 3*

"yes (a)–(c) as stated." *Respondent 4*

"Mass calibration of data from known samples by comparison with library data." *Respondent 5*

Respondent	Benchwork	Washing Glassware	Data Collection	Experimental	Running Spectra	Chromatography	Electrochemistry	Crystallography	Analysing Spectra	Calculations	Data Analysis	Computer Graphics	literature	Document Preparation	Teaching	Supervision	Administration	Miscellaneous	Attending Lectures
1													1	3	3	2	1		
2										4			3						
3			2	0.5				1							1	2.5	1	1	
4	6									2									
5										8									
6	4									4									
7					3					3				3					
8	3				1					3			0.5	0.5					
9	6				0.5													0.5	
10	2									2				4	2				
11	3				2					4									
12				10							8								
13													8						
14	2				0.5								1	1	3				
15	1				0.5								1	1	2.5	1	1		
16	0.5				1					5		1.5							
17	5				1					1			0.5						0.5
18				8									2						
19	2				2					2									
20	1					6				1									
21					2				6										
22	6.5				1					0.5									
23	4				2		2			1							1		
24															8				
25					1					1				2	2				
26	3				3					2									
27	6				0.5			0.5					1						
28	7				1								1						
29	3				2					3									
30														8					
31	6									1									
32	8				0.5					0.5									
33	4	2			9					1.5			2						
34	5				2														
35	1									2									

Table B–1: *The responses to question 1, section 1.*

"I use a PC and EMAS for simulating experiments, collecting data, and running trajectory programs." *Respondent 6*

"Yes — essential in all I do." *Respondent 7*

"Storage of Spectra as recorded with ability to manipulate and plot them." *Respondent 8*

"Use of a database containing a library of spectra from which comparisons between previously spectroscopically characterised compounds (IR, Raman, NMR *etc.*) and new products would be advantageous." *Respondent 9*

"Yes." *Respondent 10*

"Yes — (a) computer comparison / simulation for NMR / ESR spectra, (b) more data / results / papers in prep. held on computer (personal computers?)" *Respondent 11*

"Molec-geom expt. is run under computer control (IBM PC) spectra are analysed using software run on both the PC and mainframe (EMAS)" *Respondent 12*

"Comparison of Spectra with theoretical models, storage of spectral data, display of spectral data, control of lasers." *Respondent 13*

"Yes — (a) comparison of spectra, (b) storage of information, (c) basic word processing, (d) chem. abs. on line" *Respondent 14*

"a) and c) above, chemical abstracts searching, preparation of research papers." *Respondent 15*

"Apart from bench work, all the above *[working activities]* is done with the aid of a computer." *Respondent 16*

"No" *Respondent 17*

"No" *Respondent 18*

"Mathematical calc. (Jouoptics), Data acquisition and storage of spectra, graphics, word processing, CAD." *Respondent 19*

"Analysis of chromatograms, calculations, storage of chromatograms" *Respondent 20*

"I use a computer for all of them but could do with more sophisticated software for running spectra and storing them." *Respondent 21*

"All of above, also computerised chemical abstracts would be great." *Respondent 22*

"Yes — library spectra would be a good idea, as would information retrieval. I don't use complicated mathematical calculations every day." *Respondent 23*

"Could be worthwhile to set up ones own literature database, & be able to retrieve from it quickly." *Respondent 24*

"W.P. for writing, computer used for calculations" *Respondent 25*

"Could save a lot of time calculating g-values *etc.* for ESR spectra, also storage of spectra — v. useful." *Respondent 26*

"X-ray structure determination." *Respondent 27*

"Library of Spectra, molecular modelling — graphics system." *Respondent 28*

"Calculation and storage of data." *Respondent 29*

"Everything that I did would benefit from computational back-up in its various forms. *e.g.* computer controlled data acquisition, storage and manipulation (all controlled by PC). Followed by transfer to EMAS for graphical handling and theoretical modelling." *Respondent 30*

"The storage of experimental data, and then the use of this data in mathematical calculations and the plotting of this data." *Respondent 31*

"No." *Respondent 32*

"The analysis of my kinetic data from $\hat{3}1$P NMR spectra comparison w theoretical models; refining experimental parameters; storage of info." *Respondent 33*

"Comparison of spectra *etc.*" *Respondent 34*

"All of the calculational work (but already done by a computer)" *Respondent 35*

3. Are there any other facilities (such as those listed above) you would like to see provided on a central computer?

"No." *Respondent 1*

"Yes — graph plotting and presenting." *Respondent 4*

"Facility for taking standard (*e.g.* HPGL) graphics files & inserting into documents for laser printing." *Respondent 5*

"(b) and (c) essential" *Respondent 7*

"Simulation and comparison of NMR spectra" *Respondent 8*

"Yes." *Respondent 10*

"More on-line data. Better / more available graphics." *Respondent 11*

"Too early to know yet." *Respondent 13*

"(d) Chem. Abs. on line" *Respondent 14*

"A 3-D fully interactive version of Pluto would be a great advantage, especially if it were one of the CCP4 versions which would allow the simultaneous examinations of both the structure and electron density maps. Failing this, a version of Pluto allowing at least 500 input atoms (not including symmetry-related ones) would be better than the present version." *Respondent 16*

"No." *Respondent 17*

"No." *Respondent 18*

"Backup service. Software packages like spreadsheets." *Respondent 19*

"Same as above." *Respondent 20*

"On-line abstracts — Chemical Physics, Engineering and Citation Index. E-mail addresses for major (and minor!) suppliers. On-line stores information." *Respondent 21*

"Chemical abstracts." *Respondent 22*

"A better word processing system made readily available. A better calculator. How about on-line examples classes for some of the programs to be run *e.g.* Notice, Easygraph *etc.* rather than rely on unclear, useless, out of data user notes!!" *Respondent 23*

"On line chemical abstracts would be very beneficial." *Respondent 28*

"Data transfer via Kermit is very slow, however it is the most reliable. EMAS does provide everything one requires. The biggest limitations being data transfer and graphical display for analysing digitised/manipulated data." *Respondent 30*

"No." *Respondent 31*

"A basic package for plotting & printing out and "interpretation" of various forms of kinetic data — straight lines, exponential curves, *etc.* (incl. variations in concentrations of substrates (all on same plot))." *Respondent 33*

"No." *Respondent 35*

B.2 Section 2 — Library System

20 of the respondents filled in this part of the questionnaire.

1. How often do you use each of the on-line Library Catalogues

		EULCAT	GEAC
(a)	never	4	2
(b)	you have used them once or twice	5	5
(c)	when you have difficulty locating material	2	4
(d)	whenever you are looking for material in the library	6	7
(f)	*not answered*	3	2

2. How would you describe the use of each facility?

		EULCAT	GEAC
(a)	not applicable	3	1
(b)	difficult	1	0
(c)	tedious	3	9
(d)	reasonable	6	4
(e)	simple	3	4
(f)	*not answered*	4	2

3. What do you think of the standard of the two systems?

		EULCAT	GEAC
(a)	not applicable	3	1
(b)	difficult	1	3
(c)	tedious	7	7
(d)	reasonable	3	5
(e)	simple	1	1
(f)	*not answered*	4	2

4. If you have experience of both catalogue systems which do you prefer?

		EULCAT	GEAC
(a)	Used solely	3	7
(b)	Both used - one preferred	7	2

One user expressed no preference.

B.3 Section 3 — Use of EMAS

22 of the respondents filled in this part of the questionnaire.

1. How often do you use EMAS? (On Average)

(a)	At least once a week	19
(b)	A few days per month	2
(c)	About once a month	0
(d)	Less than once a month	1

2. For what purposes do you find *EMAS* most useful?

Performing Calculations	9	Resources	4	Data Transfer	3
Crystal Structure Solution	6	Graphics	3	Data Analysis	2
Word Processing	5	Programming	3	Email	2

The rest of these questions are concerned with the tools available on EMAS. The numbers in the table refer to the tools below.

A. CALC	E. Easygraph	I. Edit	M. Help
B. CCCP	F. Ortep	J. μEmacs	
C. ICON8	G. Pluto	K. VECCE	
D. SHELX	H. ECCE	L. Scribe	

3. How often do you use each of these facilities?

	A	B	C	D	E	F	G	H	I	J	K	L	M
(a) Never	12	18	19	11	7	11	12	14	7	15	16	4	0
(b) Rarely	1	1	0	0	7	1	2	1	1	2	0	6	7
(c) Quite Often	2	0	1	3	3	6	2	1	2	0	1	5	6
(d) Regularly	5	0	0	5	3	1	4	4	11	0	2	4	8
not answered	*2*	*3*	*2*	*3*	*2*	*3*	*2*	*2*	*1*	*5*	*3*	*3*	*1*

4. How long did it take you to be competent in each of the facilities?

	A	B	C	D	E	F	G	H	I	J	K	L	M
(a) not applicable	13	19	19	11	7	12	13	13	7	18	17	7	0
(b) not yet competent	3	1	1	2	3	1	0	1	0	1	0	3	1
(c) a month or less	2	0	0	4	0	4	3	1	4	1	0	4	1
(d) a week or less	1	0	0	1	10	2	3	1	4	0	3	6	2
(e) a day or less	2	0	0	1	0	1	1	4	6	0	0	0	16
not answered	*1*	*2*	*2*	*3*	*2*	*2*	*2*	*2*	*1*	*2*	*2*	*2*	*2*

5. How would you describe the use of this facility?

	A	B	C	D	E	F	G	H	I	J	K	L	M
(a) not applicable	11	17	17	10	7	10	11	13	4	16	16	6	0
(b) difficult	0	0	0	1	0	0	0	0	0	1	0	2	0
(c) tedious	3	1	1	1	1	3	2	0	2	0	0	6	3
(d) reasonable	4	0	0	5	10	5	6	2	5	1	1	2	3
(e) simple	1	0	0	1	2	0	0	5	9	0	2	3	15
not answered	*3*	*4*	*4*	*4*	*2*	*4*	*3*	*2*	*2*	*4*	*3*	*3*	*1*

6. How did you find out about this facility?

	A	B	C	D	E	F	G	H	I	J	K	L	M
(a) not applicable	12	16	16	11	7	11	11	12	3	16	16	4	0
(b) from fellow students or members of staff?	7	2	1	6	10	6	6	4	13	1	3	13	13
(c) from *User Support*?	0	0	0	0	1	0	0	1	1	0	0	0	1
(d) from the *EMAS* user guide?	0	0	0	0	1	0	0	1	0	0	0	0	4
(e) other	0	0	1	1	1	1	1	2	1	2	0	1	2
not answered	*3*	*4*	*4*	*4*	*2*	*4*	*4*	*2*	*4*	*3*	*3*	*4*	*2*

7. If you have never used the facility before why do you not use it more often?

	A	B	C	D	E	F	G	H	I	J	K	L	M
(a) you have never heard of the facility	9	14	12	8	1	10	8	1	1	11	6	1	1
(b) you have never had need to use it	3	5	7	5	3	2	4	5	3	4	4	5	1
(c) you find it difficult and tedious to use	0	0	0	0	1	0	1	2	3	2	1	4	1
(d) you don't know how to use the facility	0	0	0	0	3	0	1	6	1	3	6	2	0
(e) other	2	0	0	1	1	1	0	1	1	0	0	1	1
not answered	*8*	*3*	*3*	*8*	*13*	*9*	*8*	*7*	*12*	*2*	*5*	*9*	*18*

8. What do you think of the *standard* of *information* available regarding the facility?

	A	B	C	D	E	F	G	H	I	J	K	L	M
(a) not applicable	10	15	15	9	5	9	8	11	6	15	14	4	1
(b) bad	1	1	1	0	0	1	0	0	2	1	0	3	1
(c) could be improved	3	1	1	3	3	2	4	1	1	2	2	2	1
(d) reasonable	1	0	0	2	8	4	2	3	6	0	1	4	8
(e) good	3	0	0	3	3	1	3	3	3	0	0	5	7
not answered	*4*	*5*	*5*	*5*	*3*	*5*	*5*	*4*	*4*	*4*	*5*	*4*	*4*

9. What do you think of the *availability* of information regarding the facility?

	A	B	C	D	E	F	G	H	I	J	K	L	M
(a) not applicable	10	15	15	9	5	9	9	11	4	14	13	2	1
(b) bad	2	1	1	2	0	4	1	0	3	1	0	3	0
(c) could be improved	2	1	1	3	4	2	3	2	2	1	3	2	2
(d) reasonable	0	0	0	2	5	2	2	2	4	0	0	4	8
(e) good	3	0	0	1	4	0	2	3	4	1	1	7	6
not answered	*5*	*5*	*5*	*5*	*4*	*5*	*5*	*4*	*5*	*5*	*5*	*4*	*5*

10. What do you think of the *standard* of the facility *itself?*

	A	B	C	D	E	F	G	H	I	J	K	L	M
(a) not applicable	10	16	16	9	6	9	10	13	4	17	15	2	1
(b) bad	0	0	0	0	1	1	1	0	1	0	0	2	1
(c) could be improved	1	0	0	3	1	3	4	1	5	1	1	3	2
(d) reasonable	1	1	0	2	7	2	3	3	5	0	0	2	5
(e) good	5	0	1	3	5	2	0	2	4	0	2	7	11
not answered	*5*	*5*	*5*	*5*	*3*	*5*	*4*	*3*	*3*	*4*	*4*	*4*	*1*

B.4 Section 4 — User Suggestions

B.4.1 Library System

"Ability to recognise journal abbreviations would be useful." *Respondent 8*

"Checking for titles is not easy on EULCAT if you have a common word in the title, as it doesn't appear to take full titles but only keywords" *Respondent 23*

"How about an on-line "Dictionary of chemicals index". For example, including IUPAC names, trivial names *etc.* together with relevant information *e.g.* physical data, I.P., E.A., refractive index, m.pt, b.pt. properties such as reactions and toxicity *etc.* (similar in nature to the Cambridge dictionary of crystal structures.)" *Respondent 30*

"Without a doubt, the best improvement that could be made would be for all users to be able to "log-on" for on-line Chemical Abstracts literature search, using *e.g.* reference code number and keywords, phrases *etc.*... This would save many hours of trivial time wasting." *Respondent 30*

B.4.2 Applications

Graphics

"Ortep in the raw is a total dinosaur, but Easy Ortep lacks explicit orienting commands as Pluto has" *Respondent 3*

"Pluto: A more effective system allowing the image to be directly manipulated on a graphics screen." *Respondent 8*

"*[Easygraph]* can be a bit cumbersome to use. It would help in scaling diagrams if there was a graduated scale along two sides of the VDU display" *Respondent 20*

"Ortep: This program should be made interactive with a plotting program which should allow you to view molecules and at the same time change your viewpoint. Similarly the Pluto program" *Respondent 23*

"Easygraph — I hate it because it often takes ages to get it to do what you want it to do. However, it does do the job so it is satisfactory" *Respondent 30*

"Can we get the department to 'buy' a dedicated high resolution graphics machine — with mouse *etc. etc.* that would allow you to draw diagrams and position script *etc.* and then get hard back copies" *Respondent 30*

Document Preparation Facilities

"I find Scribe confusing and unsatisfactory to use, and its documentation isn't clear. Other chemists have commented about problems with referencing, producing tables of references *etc.*, which the set up options (*e.g.* REPORT) don't cater for. I can't even centre a numbered subheading" *Respondent 6*

"Scribe: (1) Need full definitions of the precise way each environment lays out the text. (2) Need more manuals giving a full description of the program available."
Respondent 8

"Scribe: (a) some sort of previewer for the output text to allow special fonts *etc.* to be checked. (b) Ability to use paper in landscape mode. (c) The draft facility could be improved by making it mirror more closely the output which will be obtained on laser printer. This is particularly the case for tables." *Respondent 8*

"The integrated word processor / graphics package is a good idea. It would be useful if this could incorporate files produced by other packages, such as RCO plotter files."
Respondent 8

Crystal Structure Solution

"SHELX is inflexible and cumbersome when dealing with non-routine cases ... modelling disorder in molecules containing symmetry elements is cumbersome and potentially confusing." *Respondent 3*

"On the crystallography side this *[more readily available and accessible information]* would allow person 1 and person 2 to do what they are best at *i.e.* crystallography & the research student could learn more about computing by being able to accomplish more for him/herself" *Respondent 11*

"I think that, in general, larger versions of all the currently available crystallographic programs would be a good thing. By larger versions I mean those allowing more input atoms." *Respondent 16*

"SHELX: ... a better manual is needed and more automation of the system should be available preferably in the form of questions like "Do you want to update your channel 5 (input) file Y/N?". In the manual I think examples should be given."
Respondent 23

Editors

"There are too many editors and the instruction sets aren't that freely available"
Respondent 30

"VECCE: (a) Availability of a private dictionary (for the spelling checker) like that on ECCE. (b) Ability to set up own key definitions (This is supposed to be available — does not work on chem. terminals)" *Respondent 8*

Miscellaneous

"CALC: A better help manual for CALC and perhaps the insertion of "comment lines" would be useful." *Respondent 23*

"Calculator would be good if it weren't reverse polish" *Respondent 30*

B.4.3 User Interface Issues

"In general, programs which require data (other that from a pre-existing file) should accept this data as free format entries" *Respondent 3*

"Since we essentially deal; with real time data handling with high levels of data acquisition one really does require a fully dedicated P.C. because the demands on EMAS are too variable. Once the data is taken you could help — but there is always the problem of having data in the correct format and quite often we have to change this before downlining it." *Respondent 30*

B.4.4 Support

Documentation/Help

"Manuals for packages such as VECCE, Easygraph *etc.* could be available" *Respondent 8*

"Generally more information more readily accessible and available is required." *Respondent 11*

"While there is generally enough information on specific facilities it is occasionally presented in a manner more complex than is necessary which may have the effect that rather than use a better facility (viz. the various editors) the casual user will tend to stick with a "tried and tested" one" *Respondent 16*

"Not enough publicity in the chemistry department about this system *[VECCE]*" *Respondent 23*

"The unfortunate problem with any information system is that unless you know it exists you can't exploit it...I found the way I learned was directly from people about me, that is, I learned the folk-lore of all the smart tricks from the grape-vine of accrued common knowledge." *Respondent 30*

"You could improve Help by adding catchments *e.g.* "help graphics" leading to a menu of facilities" *Respondent 30*

"Improved documentation should be available — why don't you try having documentation levels *e.g.* (1) first-timer (easy to read essentials), (2) common practitioner, (3) smart ass (nitty-gritty, everything you want to know, as technical as possible." *Respondent 30*

User Support

"Also since we sometimes downline loads of data it would be nice if we could get things like archiving to go at regular periods." *Respondent 30*

Appendix C

List of Molecular Models

Terminal Hydride Ligands

Neutron Cases

1. $IrHCl_2(PMe_2Ph)_3$ [126] (FIRCEO01)
   modelled by $IrHCl_2(PH_3)_3$
2. $[IrH(OH)(PMe_3)_4]^+$ [127] (DUPSEM01)
   modelled by $[IrH(OH)(PH_3)_4]^+$
3. $IrH_2(SiEt_3)_2(\eta^5\text{–}C_5Me_5)$ [128] (CIWJAT10)
   modelled by $IrH_2(SiEt_3)_2(\eta^5\text{–}C_5H_5)$
4. $[IrH_3(PMe_3)(\eta^5\text{–}C_5H_5)]^+$ [129] (GATDOU01)
5. $MnH(CO)_5$ [130] (FOKCEN02)
6. $MoH_2(\eta^5\text{–}C_5H_5)_2$ [98] (HCYPMO02)
7. $OsH_4(PMe_2Ph)_3$ [131] (THMPOS01)
   modelled by $OsH_4(PHMe_2)_3$
   (a) H1 and H4
   (b) H2 and H3
8. $OsH_6\{P(CHMe_2)_2Ph\}_2$ [132] (CIDWER11)
   modelled by $OsH_6\{PH(CHMe_2)_2\}_2$
   (a) H1, H2 and H3
   (b) H4, H5 and H6
9. $ReH_5(PMePh_2)_3$ [133] (CUHHES01)
   modelled by $ReH_5(PH_3)_3$
   (a) H11′ and H12′
   (b) H13′, H14′ and H15
10. $ReH_5\{P(CHMe_2)_2Ph\}_2(SiPh_3)_2$ [134] (KUJLIK)
    modelled by $ReH_5\{PH(CHMe_2)_2\}_2(SiH_3)_2$
    (a) H1, H2 and H5
    (b) H3 and H4
11. $ReH_7\{(PPh_2CH_2)_2\}$ [135] (JACNUW)
    modelled by $ReH_7\{(PH_2CH_2)_2\}$

(a) H2, H3, H4
(b) H1

12. $RhH(PPh_3)_4$ [136] (VEKFOG)
modelled by $RhH(PH_3)_4$
13. $RhH_2(SiEt_3)_2(\eta^5\text{–}C_5Me_5)$ [137] (CONFEQ01)
modelled by $RhH_2(SiEt_3)_2(\eta^5\text{–}C_5H_5)$
14. $TaH_3(\eta^5\text{–}C_5H_5)_2$ [138] (TACPTH)
15. $TaH_2Cl(PMe_3)_4$ [139] (BUXFIJ01)
modelled by $TaH_2Cl(PH_3)_4$
16. $W_2H_2(PMe_3)_4(\eta^2\text{–}CH_2CH_2)_2(C_3H_3O_2)$ [140] (KAYDET)
modelled by $W_2H_2(PH_3)_4(\eta^2\text{–}CH_2CH_2)_2(C_3H_3O_2)$
17. $\{ZnH(NMe_2EtNMe)\}_2$ [141] (MAEMAZ11)

X-ray Cases

18. $CoH(CO)(PPh_3)_3$ [142] (CAHCOP01)
modelled by $CoH(CO)(PH_3)_3$
19. $FeH(CO)(NO)(PPh_3)_2$ [143] (BUNYIS)
modelled by $FeH(CO)(NO)(PH_3)_2$
20. $FeH_2\{(CCH_3)_2B_4H_4\}_2$ [123] (MECBFE)
21. $FeH_2(N_2)(PEtPh_2)_3$ [144] (SIGFOD)
modelled by $FeH_2(N_2)(PH_3)_3$
22. $IrH(PMe_3)(\eta^5\text{–}C_5Me_5)(CHCH_2)$ [145] (DAMLOS)
23. $IrHCl(PMe_3)_3(CH_2COH)$ [146] (BIYJIC)
24. $IrHCl(PEt_3)_2(NHPhC_7H_{10})$ [147] (GINWEF)
modelled by $IrHCl(PMe_3)_2(NHPhC_7H_{10})$
25. $[IrH(PMe_3)_4(OMe)]^+$ [148] (DUPSIQ)
26. $IrH_2(PMe_3)_3(BC_8H_{14})$ [149] (JIHFOV)
modelled by $IrH_2(PH_3)_3(BC_8H_{14})$
27. $[IrH_2(PPh_3)_2(N_2C_{10}H_8)]^+$ [150] (FONDUH)
modelled by $[IrH_2(PH_3)_2(N_2C_{10}H_8)]^+$
28. $[Ir_2H_2(OH)Cl(PEt_3)_4(NPh)]^+$ [147] (GINWIJ)
modelled by $[Ir_2H_2(OH)Cl(PH_3)_4(NPh)]^+$
29. $IrReH_2\{PH(C_6H_{11})_2\}_2\{P(C_6H_{11})_2\}_2(C_8H_{12})$ [151] (JOVLEL)
modelled by $IrReH_2(PH_3)_2(PH_2)_2(C_8H_{12})$
30. $[MoH(CO)(\eta^5\text{–}C_5H_5)_2]^+$ [152] (CPCBMO)
31. $[MoH\{P(OMe)_3\}_2(CCH_2CMe_3)(\eta^5\text{–}C_5H_5)]^+$ [153] (BENFEF10)
modelled by $[MoH\{P(OMe)_3\}_2(CCH_2CH_3)(\eta^5\text{–}C_5H_5)]^+$
32. $[NbH_2(\eta^5\text{–}C_5H_5)_2]^-$ [154] (GABFEU)
33. $[OsH(\mu_2\text{–}CH_2CH_2)_2(PMe_2Ph)_3]^+$ [155] (KARMEV)
modelled by $[OsH(\mu_2\text{–}CH_2CH_2)_2(PHMe_2)_3]^+$

34. $[OsH_3(PMe_2Ph)_3]^-$ [156] (DAMMEJ)
    modelled by $[OsH_3(PH_3)_3]^-$
35. $[Pd_2HClMe\{(PPh_2)_2CH_2\}]^+$ [157] (GEWHUL)
    modelled by $[Pd_2HClMe\{(PH_2)_2CH_2\}]^+$
36. $PtH(CH_2CMe_3)[\{P(C_6H_{11})_2CH_2\}_2]$ [158] (FARLOZ10)
    modelled by $PtHMe\{(PH_2CH_2)_2\}$
37. $PtH(CH_2CMe_3)[\{P(CMe_3)_2\}_2CH_2]$ [159] (VETZAV)
    modelled by $PtH(CH_2CMe_3)\{(PH_2)_2CH_2\}$
38. $ReH(\eta^5\text{–}C_5H_5)(\eta^5\text{–}C_5Me_4Et)$ [160] (JASRAW)
39. $ReH_2(PPh_3)_2(\eta^5\text{–}C_5H_5)$ [161] (FIZLUV)
    modelled by $ReH_2(PH_3)_2(\eta^5\text{–}C_5H_5)$
40. $ReH_2(PMe_2Ph)_2(C_8H_{12})$ [162] (CUKMEA)
    modelled by $ReH_2(PH_3)_2(C_8H_{12})$
41. $ReH_2(PMe_2Ph)_2(\eta^5\text{–}C_5Me_5)$ [163] (SAWNUZ)
    modelled by $ReH_2(PHMe_2)_2(\eta^5\text{–}C_5Me_5)$
42. $[ReH_2(PPh_3)_2(NOC_6H_6)_2]^+$ [164] (KINTEG10)
    modelled by $[ReH_2(PH_3)_2(NOC_6H_6)_2]^+$
43. $ReH_3(\eta^2\text{–}CH_2CH_2)_2\{P(CHMe_2)_2Ph\}_2$ [165] (COXXES)
    modelled by $ReH_3(\eta^2\text{–}CH_2CH_2)_2(PH_3)_2$
44. $[ReRhH\{P(C_6H_{11})_2\}_4(C_8H_{12})]^+$ [166] (GIKPOF)
    modelled by $[ReRhH(PH_2)_4(C_8H_{12})]^+$
45. $[ReRhH_2\{PH(C_6H_{11})_2\}\{P(C_6H_{11})_2\}_3\{(PMe_2CH_2)_2\}]^+$ [151] (JOVLIP)
    modelled by $[ReRhH_2(PH_3)(PH_2)_3\{(PMe_2CH_2)_2\}]^+$
46. $RhH\{P(O)(OMe_2)_2\}_2\{P(OH)(OMe_2)_2\}_2(CO)$ [167] (DEMCUT)
47. $RhH\{P(CHMe_2)_3\}_3$ [168] (IPRHRH)
    modelled by $RhH(PMe_3)_3$
48. $RhH(PPh_3)_2(B_{10}NH_{11})$ [169] (KOJKUP)
    modelled by $RhH(PH_3)_2(B_{10}NH_{11})$
49. $RhHCl\{P(CHMe_2)_3\}_2(BO_2C_6H_4)$ [170] (SITKUB)
    modelled by $RhHCl(PH_3)_2(BO_2C_6H_4)$
50. $RhHCl(PMe_3)_3(CH_2COMe)$ [171] (VEZFAH)
51. $[RhH(PMe_3)_4(CCCH_2CH_2OH)]^+$ [172] (FOMLUO)
    modelled by $[RhH(PMe_3)_4(CCH)]^+$
52. $[RhH(CNCH_2CMe3)_2\{(N_2C_5H_7)_3BH\}]^+$ [173] (VIPVUL)
    modelled by $[RhH(CNMe)_2\{(N_2C_5H_7)_3BH\}]^+$
53. $RhH_2(O_2COH)\{P(CHMe_2)_3\}_2$ [174] (IPHCIR)
    modelled by $RhH_2(O_2COH)(PMe_3)_2$
54. $RhH_2Cl\{P(CHMe_2)_3\}_2$ [175] (JOWLEM)
    modelled by $RhH_2Cl(PMe_3)_2$
55. $[RhHN\{P(CHMe_2)_3\}_2]_2$ [176] (HNIPRH)
    modelled by $\{RhHN(PH_3)_2\}_2$
56. $Rh_2H_2\{(CC_6H_5)_2\}\{P(CHMe_2)_3\}_3(CO_3)$ [177] (CESSOI)

modelled by $Rh_2H_2(C_2H_2)(PH_3)_3(CO_3)$

57. $[Rh_3H_2Cl_2\{(PPh_2CH_2)_2PPh\}_2(CO)_2]^+$ [178] (DOBLAH10)
modelled by $[Rh_3H_2Cl_2\{(PH_2CH_2)_2PH\}_2(CO)_2]^+$
58. $RhZnH_2\{(PPh_2CH_2)_3CCH_3\}\{N(SiMe_3)_2\}$ [179] (SACKAI)
modelled by $RhZnH_2\{(PH_2CH_2)_3CCH_3\}\{N(SiMe_3)_2\}$
59. $RuH\{(PMe_2C_3H_6)_3P\}Ph$ [180] (SACGEI)
modelled by $RuH\{(PH_2C_3H_6)_3P\}Ph$
60. $RuH(\eta^5\text{–}C_5H_5)\{(PPh_2CH_2)_2CH_2\}$ [181] (VUBXUL)
modelled by $RuH(\eta^5\text{–}C_5H_5)\{(PH_2CH_2)_2CH_2\}$
61. $TaH(\eta^5\text{–}C_5Me_5)_2(\eta^2\text{–}CH_2CH_2)(AlEt_3)$ [182] (GACMAY)
modelled by $TaH(\eta^5\text{–}C_5H_5)_2(\eta^2\text{–}CH_2CH_2)(AlEt_3)$
62. $TaH_2Cl_2\{(PMe_2CH_2)_2\}_2$ [183] (CACJEV)
63. $WH(CO)_2(NO)(PMe_3)_2$ [184] (SOLMUB)
64. $WH_2Cl_2(PHMe_2)_4$ [185] (CALNAE)
modelled by $WH_2Cl_2(PMe_2Ph)_4$

Edge Bridging Hydride Ligands

Neutron Cases

65. $[Cr_2H(CO)_{10}]^-$ [186] (KCPTCR01)
66. $[Cr_2H(CO)_{10}]^-$ [187] (PIMCRC01)
67. $Fe_4H(CH)(CO)_{12}$ [188] (HMYCFE01)
68. $IrH_2Cl(PMe_2Ph)_3$ [189] (GEVNIE05)
modelled by $IrH_2Cl(PH_3)_3$
69. $[Ir_4H(CO)_{11}]^-$ [190] (CUSGAY01)
70. $MnH(CO)_2(\eta^5\text{–}C_5H_4Me)(SiPh_2F)$ [191] (BOTGAS)
71. $Mn_2ReH(CO)_{14}$ [192] (VITMIU01)
72. $Mo_2H(CO)_4(\eta^5\text{–}C_5H_5)_2(PMe_2)$ [193] (PHCCMO01)
73. $[Mo_2HCl_8]^{3-}$ [194] (CEHDAU01)
74. $Os_3H(CO)_{10}(CHCH_2)$ [195] (CHVINO)
75. $Os_3H(CO)_{10}(NCHCF_3)$ [196] (BAJXIT)
76. $Os_3H_2(CO)_{10}$ [197, 198] (FOKNAU02)
77. $Os_3H_2(CO)_{10}(\mu_2\text{–}CH_2)$ [199] (DCHMOS01)
78. $Os_3H_2(CO)_9S$ [101] (FUZKAM01)
79. $Os_3H_3(CO)_9(\mu_3\text{–}CH)$ [200] (COTPOQ01)
80. $Os_4H_3(CO)_{11}(CHCHPh)$ [201] (HPETOS)
81. $Os_4H_3I(CO)_{12}$ [202] (FUZKEQ01)
82. $Os_4H_4(CO)_{11}\{P(OMe)_3\}$ [203] (BAJSUA01)

(a) H12 and H23
(b) H41 and H34

83. $Re_3H_3(CO)_8[\{P(OEt)_2\}_2O]_2$ [204] (BEGPOS01)
modelled by $Re_3H_3(CO)_8\{(PH_2)_2O\}_2$

84. $Re_3H_3(CO)_{11}(PPh_3)$ [203] (BAJTAH01)
modelled by $Re_3H_3(CO)_{11}(PH_3)$

85. $Rh_2H_2\{P(OCHMe_2)_3\}_4$ [205] (IPXHRH01)
modelled by $Rh_2H_2(PH_3)_4$

86. $Rh_3H_3\{P(OMe)_3\}_6$ [106] (HRHMOP11)
modelled by $Rh_3H_3(PH_3)_6$

87. $Ru_3H(CO)_9(CCBu^t)$ [206] (HCMBRU11)

88. $Ru_3H(CO)_7\{(AsPh_2)_2CH_2\}(AsPh_2CH_2AsPh)$ [207] (CEPPES01)
modelled by $Ru_3H(CO)_7\{(AsH_2)_2CH_2\}(AsH_2CH_2AsH)$

89. $[Ru_3FeH(CO)_{13}]^-$ [208] (PIHFRB10)

90. $Ru_4H_4\{P(OMe)_3\}_4(CO)_8$ [209] (BOPZAH)
modelled by $Ru_4H_4\{P(OH)_3\}_4(CO)_8$
(a) H1310 and H2410
(b) H1410 and H2310

91. $W_2H(CO)_9(NO)$ [210] (FUZRIB01)

92. $W_2H(CO)_8(NO)\{P(OMe)_3\}$ [211] (HWCNMP11)
modelled by $W_2H(CO)_8(NO)(PH_3)$

93. $[W_2H_2(CO)_8]^{2-}$ [212] (BIKBEC01)

94. $[W_2H(CO)_{10}]^-$ [213] (DEKMUB01)

X-ray Cases

95. $AuCrH(PPh_3)(CO)_5$ [214] (BIHXEV)

96. $[AuRuH_2(PPh_3)\{(PPh_2)_2CH_2\}_2]^+$ [150] (FONDOB)
modelled by $[AuRuH_2(PH_3)\{(PH_2)_2CH_2\}_2]^+$

97. $[Au_2RuH_2(PPh_3)_2\{(PPh_2)_2CH_2\}_2]^{+2}$ [215] (DURSUE)
modelled by $[Au_2RuH_2(PH_3)_2\{(PH_2)_2CH_2\}_2]^{+2}$

98. $[Co_2H(\eta^5\text{–}C_5H_5)_2\{(PPh_2CH_2)_2CH_2\}]^+$ [216] (CUKXAH)

99. $[Fe_2H_3\{(PPh_2CH_2)_3CCH_3\}_2]^+$ [217] (PPMEFE)
modelled by $[Fe_2H_3\{(PH_2CH_2)_3CH\}_2]^+$

100. $Fe_3H(CO)_9\{S(C_6H_{11})\}$ [218,219] (BEVTIF)

101. $Fe_3H(CO)_9(SC_3H_7)$ [219] (IPSHFE)

102. $Fe_3H(CO)_9(NHCMe)$ [220] (ACICFE)

103. $Fe_5HN(CO)_{14}$ [221] (FUZLAN)

104. $Ir_2H(\eta^5\text{–}C_5Me_5)_2Ph(C_3H_4)$ [222] (FALNAH)
modelled by $Ir_2H(\eta^5\text{–}C_5H_5)_2Ph(C_3H_4)$

105. $Mn_3H_3(CO)_{12}$ [223] (FUZPOF)

106. $Mo_2H(CSiMe_3)(CH_2SiMe_3)_2(NC_5H_5)_2(OC_6H_4Me_2)_2$ [224] (CALVUG)
modelled by $Mo_2H(CSiH_3)(CH_2SiH_3)_2(NC_5H_5)_2(OMe)_2$
107. $[Ni_2H(\mu_2\text{–}CH_2CH_2)_4]^-$ [225] (BIPJIT10)
108. $LiNi_2H(\eta^2\text{–}CH_2CH_2)_4\{(NMe_2C_2H_4)_2NMe\}$ [226] (DUZCEG)
modelled by $LiNi_2H(\eta^2\text{–}CH_2CH_2)_4\{(NH_2C_2H_4)_2NMe\}$
109. $[Os_2H_3(PMe_2Ph)_6]^+$ [227] (CACNAV)
modelled by $[Os_2H_3(PH_3)_6]^+$
110. $[Os_3H(CO)_9S]^-$ [228] (PIMOSS)
111. $Os_3H_2(CO)_9(PPh_3)_3$ [229] (NCHPOS)
modelled by $Os_3H_2(CO)_9(PH_3)$
112. $Os_4H_3(CO)_{11}(C_6H_9)$ [124] (UCHXOS)
113. $[Os_4H_2(CO)_{12}]^{2-}$ [230] (IMPOSC)
114. $Os_5H_2(CO)_{16}$ [231] (FONGAQ)
115. $Os_3PtH_2(CO)_{10}\{P(C_6H_{11})_3\}$ [125] (HXPTOS)
modelled by $Os_3PtH_2(CO)_{10}(PH_3)$
116. $PtWH(CO)_2(\eta^5\text{–}C_5H_5)(PPh_3)_2(C_8H_8)$ [232] (BATJUB10)
117. $Re_2H_2(CO)_6\{(PPh_2)_2CH_2\}$ [233] (PPECRE)
modelled by $Re_2H_2(CO)_6\{(PH_2)_2CH_2\}$
118. $Rh_2H\{P(OCHMe_2)_3\}_4(CPhCHPh)\}$ [234] (BUWRAM)
modelled by $Rh_2H\{P(OH)_3\}_4(CPhCHPh)\}$
119. $Ru_2H_2(C_8H_{12})_2(N_2C_3H_3)(N_3C_5H_6)$ [235] (FOYCIF)
modelled by $Ru_2H_2(\mu_2\text{–}CH_2CH_2)_4(N_2C_3H_3)(N_3C_5H_6)$
120. $Ru_3H(CO)_9(C_6H_9)$ [236] (MALCRU)
121. $Ru_3H(CO)_{10}(CNMe_2)$ [237] (DCTRIM10)
122. $Ru_3H(CO)_{10}(OCNMe_2)$ [238] (DMFHRU)
123. $Ru_3H_3(CO)_9(\mu_3\text{–}CMe)$ [239] (NCEYRU)

Face Bridging Hydride Ligands

Neutron Cases

124. $Cu_6H_6\{P(C_6H_4Me)_3\}_6$ [240] (CICRIP01)
modelled by $Cu_6H_6(PH_3)_6$
(a) H1 and H2
(b) H3 and H4
(c) H5 and H6
125. $FeCo_3H(CO)_9\{P(OMe)_3\}_3$ [241] (HMPCIC01)
modelled by $FeCo_3H(CO)_9(PH_3)_3$
126. $Ni_4H_3(\eta^5\text{–}C_5H_5)_4$ [242,99] (TCPNIH11)
127. $[Os_6H(CO)_{18}]^-$ [243,244] (FUCPUO01)
128. $[Rh_4H_4(\eta^5\text{–}C_5Me_5)_4]^{2+}$ [245] (HPMCRH11)

modelled by $[Rh_4H_4(\eta^5\text{–}C_5H_5)_4]^{2+}$
(a) H1 and H2
(b) H3 and H4

X-ray Cases

129. $AuRe_3H(CO)_6(PPh_3)(PPh_2)_3$ [246] (KOFHIW)
modelled by $AuRe_3H(CO)_6(PH_3)(PH_2)_3$

130. $AuRhRu_3H(CO)_{11}(PPh_3)_2(COMe)$ [247] (KIPZOY)
modelled by $AuRhRu_3H(CO)_{11}(PH_3)_2(COMe)$

131. $Co_4H_2(\eta^5\text{–}C_5H_5)_4(B_2H_2)$ [248] (FELMIS)

132. $Co_3FeH(CO)_{12}(PPh_3)_2$ [249] (COXSAJ)
modelled by $Co_3FeH(CO)_{12}(PH_3)_2$

133. $Co_3FeH(CO)_{11}(PHPh_2)$ [250] (KOKVEL)
modelled by $Co_3FeH(CO)_{11}(PH_3)$

134. $CoFeGeMoH(CO)_8(\eta^5\text{–}C_5H_5)(Bu^t)$ [251] (DENJAH)
modelled by $CoFeGeMoH(CO)_8(\eta^5\text{–}C_5H_5)Me$

135. $CoRh_2RuH(CO)_{10}(PPh_3)_2$ [252] (SEWKOU)
modelled by $CoRh_2RuH(CO)_{10}(PH_3)_2$

136. $Co_2RhRuH(CO)_{11}(PMe_2Ph)$ [252] (SEWKEK)
modelled by $Co_2RhRuH(CO)_{11}(PH_3)$

137. $Co_3RuH(CO)_{11}(SMe_2)$ [253] (KEHSIZ)

138. $Cr_2CuH(CO)_{10}(N_2C_{10}H_8)$ [254] (SOTRUO)

139. $Cr_2CuH(CO)_{10}(N_2H_4C_6H_4)$ [254] (SOTSAV)

140. $Cr_2Cu_2H_2(CO)_{10}(NC_5H_5)_2$ [254] (SOTSEZ)

141. $CuRu_4H_3(CO)_{12}(PMePh_2)$ [255] (DUXVOH)
modelled by $CuRu_4H_3(CO)_{12}(PH_3)$

142. $Cu_2Ru_4H_2(CO)_{12}(PPh_3)_2$ [256] (CALMIL10)
modelled by $Cu_2Ru_4H_2(CO)_{12}(PH_3)_2$

143. $Cu_2Ru_4H_2(CO)_{12}(PPh_2CH_2CH_2)_2CH_2\}$ [257] (FURVOD)
modelled by $Cu_2Ru_4H_2(CO)_{12}(PH_2CH_2CH_2)_2CH_2\}$

144. $Cu_2Ru_4H_2(CO)_{12}(PPh_2CH_2\eta^2\text{–}Ph)$ [258] (KAHBIE)
modelled by $Cu_2Ru_4H_2(CO)_{12}(PH_2CH_2\eta^2\text{–}Ph)$

145. $FeLi_4H_2Ph_4(OEt_2)_4$ [259] (VUCKOT)
modelled by $FeLi_4H_2(\mu_2\text{–}CH_2)_4(OH_2)_4$

146. $[Fe_3NiH(CO)_{12}]^-$ [260] (ICORHC10)

147. $Fe_3PtH(CO)_{10}(COMe)(PPh_3)$ [261] (BEBKIC)
modelled by $Fe_3PtH(CO)_{10}(COMe)(PH_3)$

148. $Hg_2Os_5H(CO)_{15}(OCOCF_3)_3$ [262] (KOBLAO)
modelled by $Hg_2Os_5H(CO)_{15}(OCOH)_3$

149. $Ni_3H_2\{\eta^5\text{–}C_5(Bu^t)_2H_3\}_3$ [263] (KIMDOZ)
modelled by $Ni_3H_2(\eta^5\text{–}C_5H_5)_3$

150. $Os_3WH(CO)_{10}(\eta^5\text{–}C_5H_5)\{(CCOOCH_2CH_3)_2\}$ [264] (GISBIT)
modelled by $Os_3WH(CO)_{10}(\eta^5\text{–}C_5H_5)\{(CCOOH)_2\}$

151. $Pt_2Ru_8H_2(CO)_{23}$ [265] (KUFLOM)

152. $Re_3H(CO)_6I_2(PPh_2)_2$ [266] (JIJZUX)
modelled by $Re_3H(CO)_6I_2(PH_2)_2$

153. $Re_3H(CO)_6Br(PPh_2)_3$ [267] (VIJJIH)
modelled by $Re_3H(CO)_6Br(PH_2)_3$

154. $Rh_3H(C_8H_{12})_3\{(CH_2)_3C\}$ [268] (KIRMUT)
modelled by $Rh_3H(\eta^2\text{–}CH_2CH_2)_6\{(CH_2)_3C\}$

155. $Rh_2Ru_3H_2(CO)_{12}(PPh_3)_2$ [247] (KIPZIS)
modelled by $Rh_2Ru_3H_2(CO)_{12}(PH_3)_2$

156. $Rh_3RuH(CO)_{10}(PPh_3)_2$ [269] (DEGLAC)
modelled by $Rh_3RuH(CO)_{10}(PH_3)_2$

157. $Ru_5H(CO)_{12}S(C_7H_{11})$ [270] (VADSAU)

158. $Ru_6H(CO)_{15}S(C_7H_{11})$ [270] (VADSEY)

159. $Ru_6H_2(CO)_{17}$ [271] (SIMBAR)

160. $Ru_6H_2(CO)_{17}S$ [272] (GOBWID)

Interstitial Hydride Ligands

Neutron Cases

161. $[Co_6H(CO)_{15}]^-$ [273, 274] (PIMHCO10)

162. $[Ru_6H(CO)_{18}]^-$ [104, 108, 275] (PAHCRU)

X-ray Cases

163. $Ru_7H(CO)_{19}(CNMe_2)$ [276] (GINHUG)

Mixed Hydride Ligands

Neutron Cases

164. $[IrPtH_4(PEt_3)_4]^+$ [277] (BORFOD10)
modelled by $[IrPtH_4(PMe_3)_4]^+$
(a) H1 and H2
(b) H3 and H4

165. $[Pt_2H_3\{(PPh_2CH_2)_2\}_2]^+$ [278] (CAKNEH01)
modelled by $[Pt_2H_3\{(PH_2CH_2)_2\}_2]^+$

166. $Re_2H_8(PEt_2Ph)_4$ [279] (OHEPRH01)
modelled by $Re_2H_8(PH_3)_4$

(a) H1 and H3
(b) H5 and H7

167. $[Re_2H_9\{(PPh_2CH_2)_3CCH_3\}]^-$ [280] (BIXSEG01)
modelled by $[Re_2H_9\{(PH_2CH_2)_3CCH_3\}]^-$
(a) H3, H4 and H7
(b) H8

168. $WH_5(PMe_3)_3[Na\{O(CH_2)_2\}_5]$ [281] (DECVAI01)
modelled by $WH_5(PH_3)_3[Na\{O(CH_2)_2\}_5]$
(a) H2, H3 and H5
(b) H1 and H4

169. $WH_5(PMe_3)_3[K\{O(CH_2)_2\}_6]$ [281] (DECVEM01)
modelled by $WH_5(PH_3)_3\{K(OH_2)_6\}$
(a) H53, H54 and H56
(b) H52 and H55

X-ray Cases

170. $AuRu_4H_3(CO)_{12}(PPh_3)$ [282] (FERSOK)
modelled by $AuRu_4H_3(CO)_{12}(PH_3)$

171. $Au_2Ru_4H_2(CO)_{12}(PPh_3)_2$ [256] (FECMAB)
modelled by $Au_2Ru_4H_2(CO)_{12}(PH_3)_2$

172. $Au_2Ru_4H_2(CO)_{12}\{(PPh_2)_2CH_2\}$ [283] (DOCWEX)
modelled by $Au_2Ru_4H_2(CO)_{12}\{(PH_2)_2CH_2\}$

173. $Au_2Ru_4H_2(CO)_{12}\{(PPh_2CH_2)_2\}$ [284] (KAPLES)
modelled by $Au_2Ru_4H_2(CO)_{12}\{(PH_2CH_2)_2\}$

174. $Os_3H_2(CO)_{10}(PPh_3)_3$ [103] (DCTPHO)
modelled by $Os_3H_2(CO)_{10}(PH_3)$

175. $Co_2Ru_2H_2(CO)_{12}$ [285] (FOKWUX)

176. $Re_3H_2(CO)_{10}(\eta^5\text{–}C_7H_9)$ [286] (COHGEL)

177. $Ru_6H_2(CO)_{15}(PMe_3)(OC_6H_4)$ [287] (KIYNOV)
modelled by $Ru_6H_2(CO)_{15}(PH_3)(OC_6H_4)$

178. $Ru_6H_2(CO)_{16}(OC_6H_4)$ [288] (FOSTAI)

DiHydrogen Ligands

Neutron Cases

179. $[Fe(H_2)H\{(PPh_2CH_2)_2\}_2]^+$ [289] (KECVUJ)
modelled by $[Fe(H_2)H\{(PH_2CH_2)_2\}_2]^+$

180. $Fe(H_2)H_2(PEtPh_2)_3$ [144] (SIGFUJ)
modelled by $Fe(H_2)H_2(PH_3)_3$

181. $ReH_7\{P(C_6H_4Me)_3\}_2$ [290] (KILPEA)
modelled by $ReH_7(PH_3)_2$

X-ray Cases

182. $FeRu_2(H_2)H_2Cl_2(PPh_3)_2(\eta^5\text{–}C_5H_5)(NMe_2CHMeC_5H_3P(CHMe_2)_2)$ [291] (VABMAM)
modelled by $FeRu_2(H_2)H_2Cl_2(PH_3)_2(\eta^5\text{–}C_5H_5)(NMe_2CH_2C_5H_3PH_2)$

183. $Re(H_2)Cl(PMePh_2)_4$ [292] (GIKGOW01)
modelled by $Re(H_2)Cl(PH_3)_4$

184. $Re(H_2)Cl(PMe_3)_4$ [292] (VIPPOZ)
modelled by $Re(H_2)Cl(PH_3)_4$

185. $Ru(H_2)HI\{P(C_6H_{11})_3\}_2$ [293] (JISZOA)
modelled by $Ru(H_2)HI(PH_3)_2$

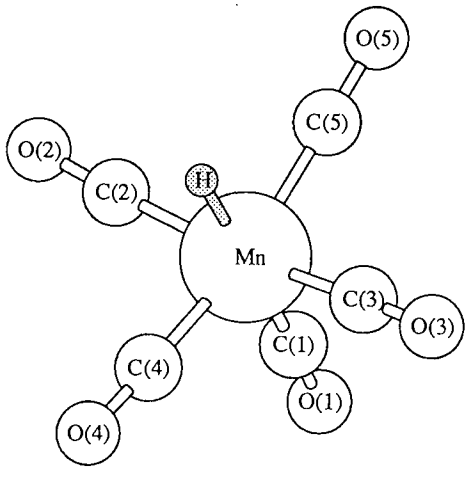
186. $W(H_2)(CO)_3\{P(CHMe_2)_3\}_2$ [294] (CEJDEA)
modelled by $W(H_2)(CO)_3(PH_3)_2$

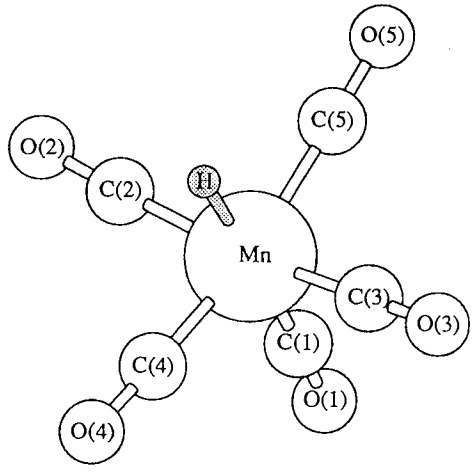
Appendix D

Molecular Models

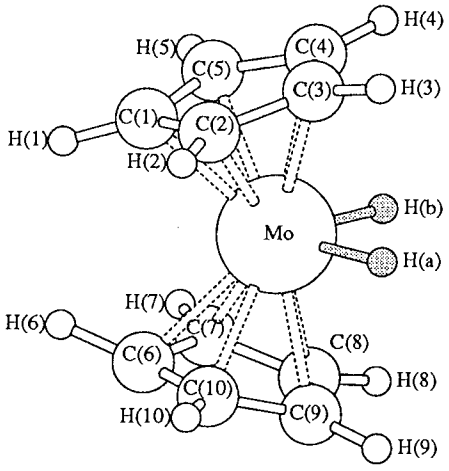
In this appendix several of the models used in the testing of the Locator are illustrated. In the model on the left is the hydride ligand has been located by conventional techniques, in that on the right the hydride ligand has been located by the Locator.

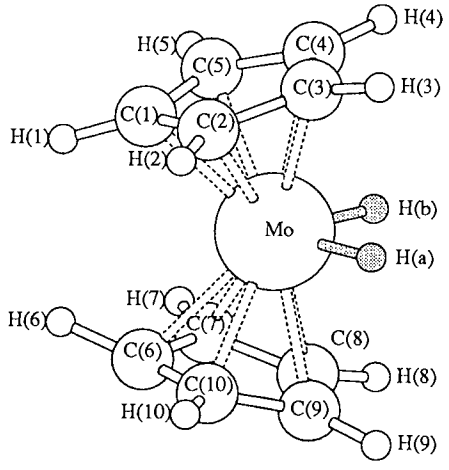
5. $MnH(CO)_5$. (FOKCEN02).



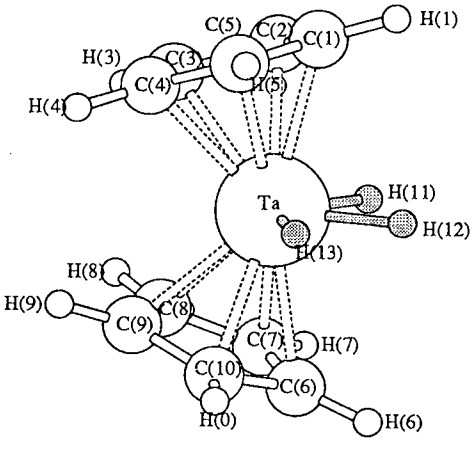


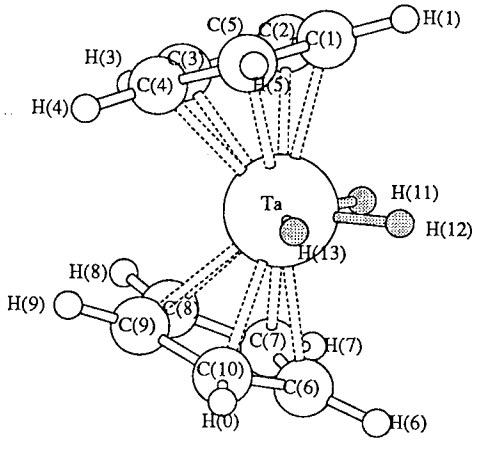
6. $MoH_2(\eta^5\text{–}C_5H_5)_2$. (HCYPMO02).



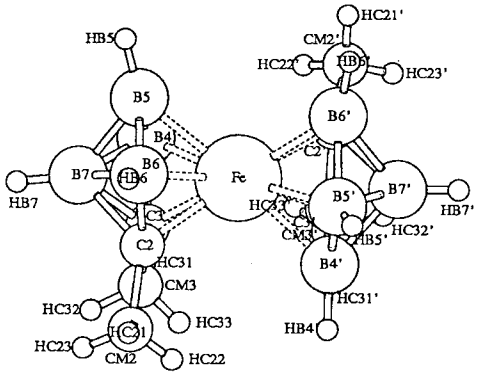


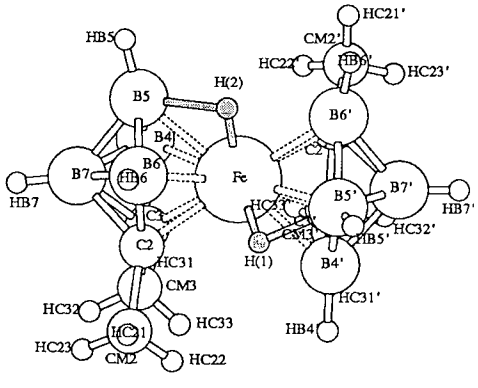
14. $TaH_3(\eta^5\text{–}C_5H_5)_2$. (TACPTH).



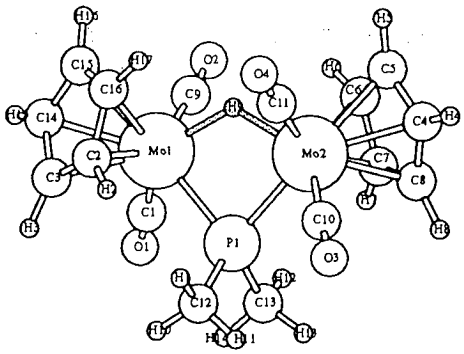


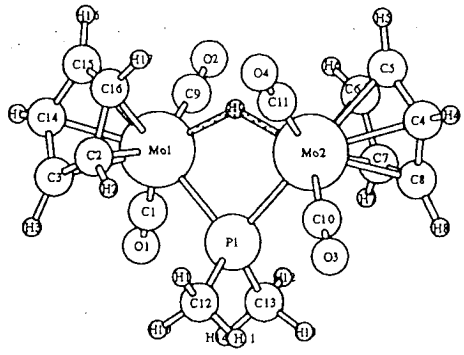
20. $FeH_2\{(CCH_3)_2B_4H_4\}_2$.



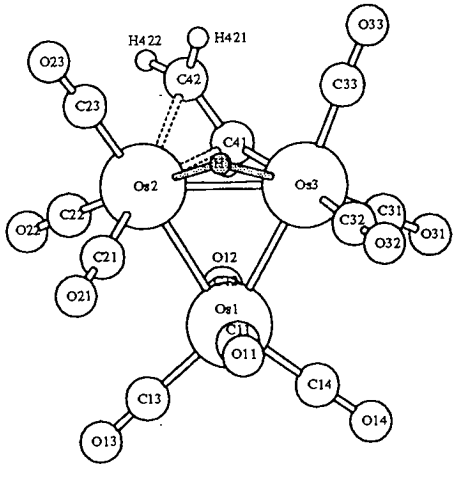


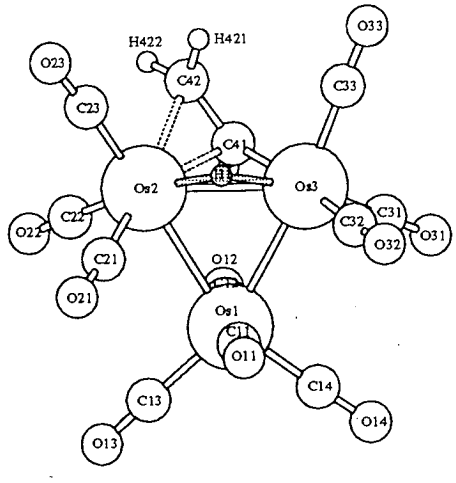
72. $Mo_2H(CO)_4(\eta^5\text{-}C_5H_5)_2(PMe_2)$. (PHCCMO01).



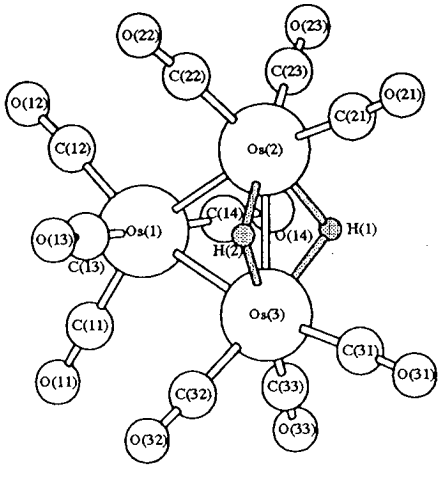


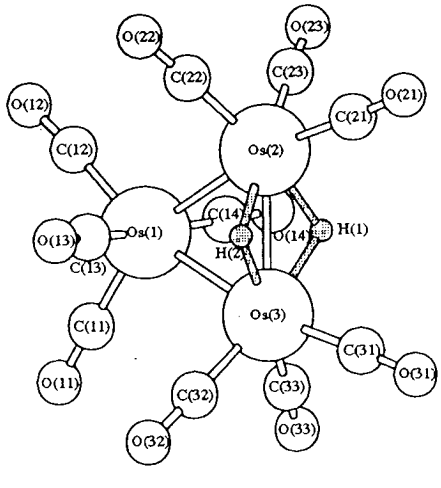
74. $Os_3H(CO)_{10}(CHCH_2)$. (CHVINO).



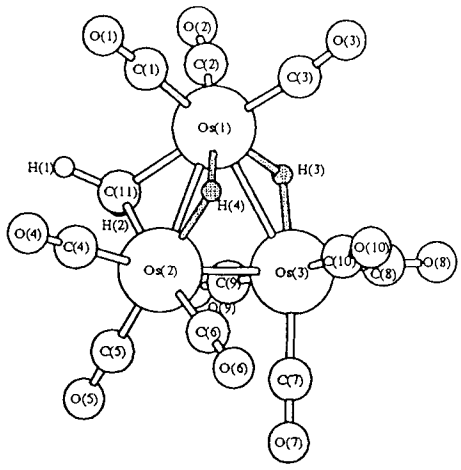


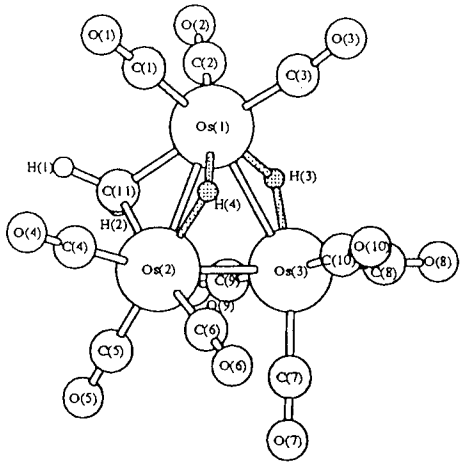
76. $Os_3H_2(CO)_{10}$. (FOKNAU02).



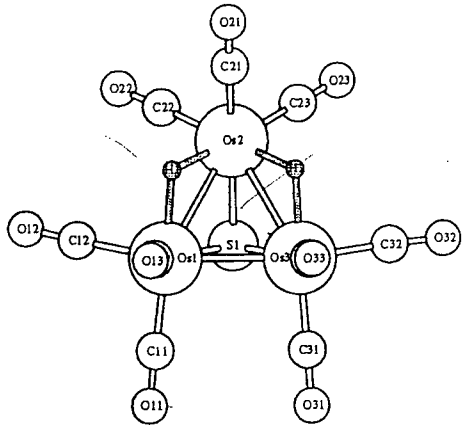


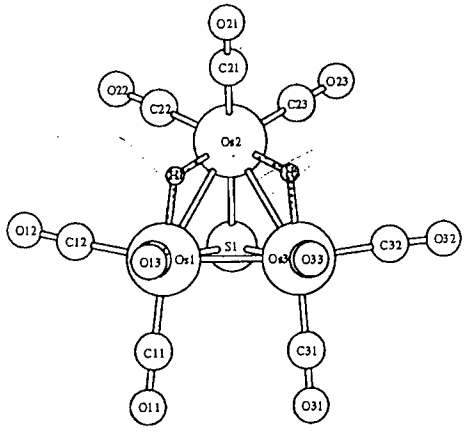
77. $Os_3H_2(CO)_{10}(\mu_2\text{–}CH_2)$. (DCHMOS01).



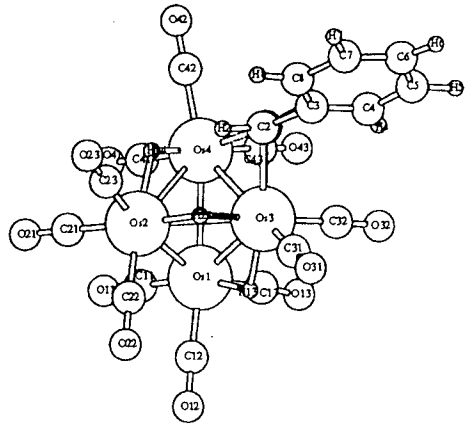


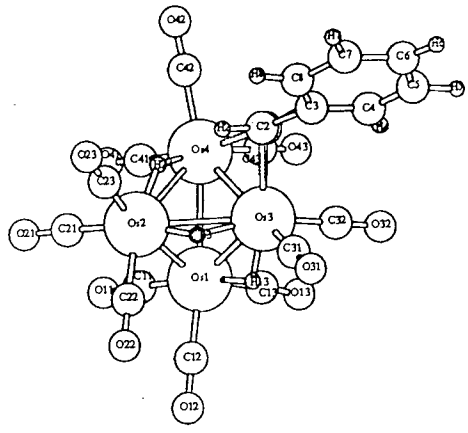
78. $Os_3H_2(CO)_9S$. (FUZKAM01).



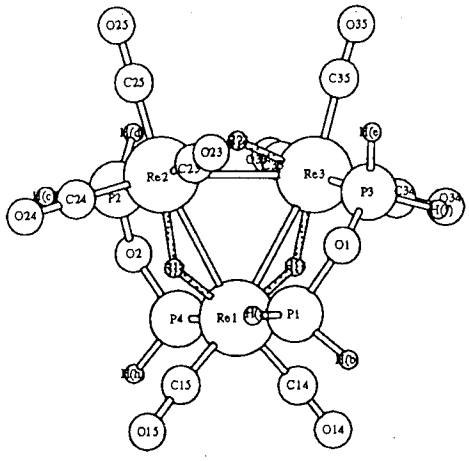


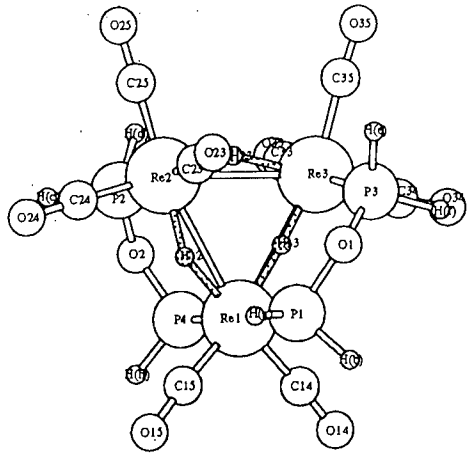
80. $Os_4H_3(CO)_{11}(CHCHPh)$. (HPETOS).



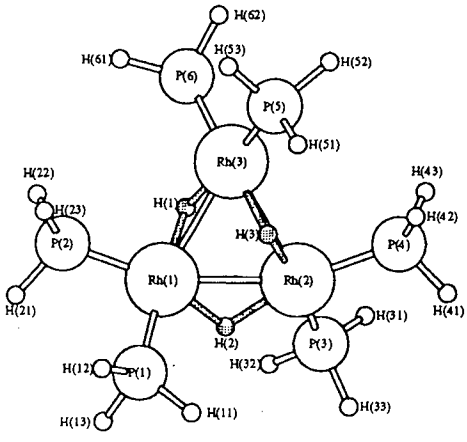


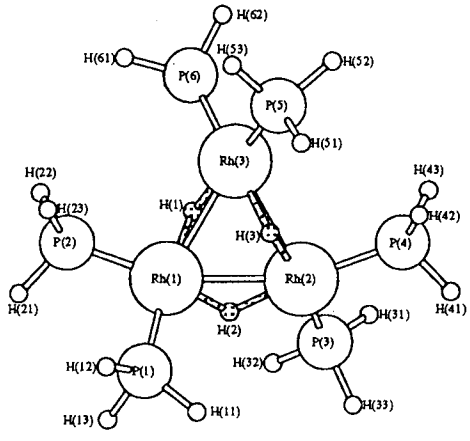
83. $Re_3H_3(CO)_8\{(PH_2)_2O\}_2$, a model for $Re_3H_3(CO)_8[\{P(OEt)_2\}_2O]_2$. (BEGPOS01).



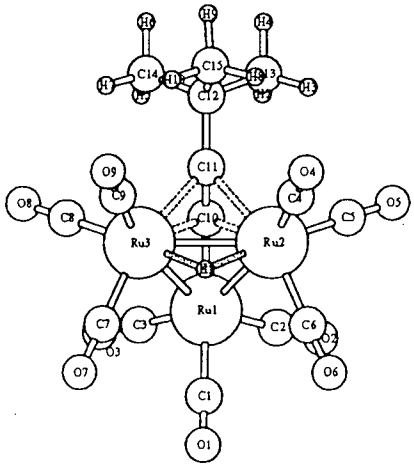


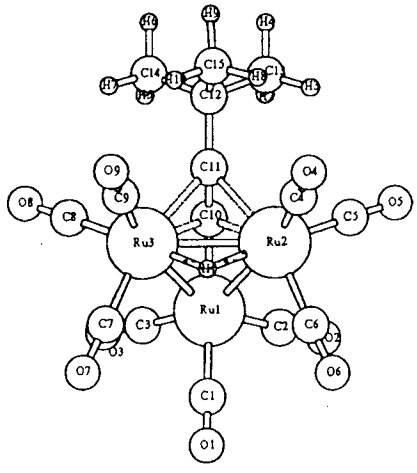
86. $Rh_3H_3(PH_3)_6$, a model for $Rh_3H_3\{P(OMe)_3\}_6$. (HRHMOP11).



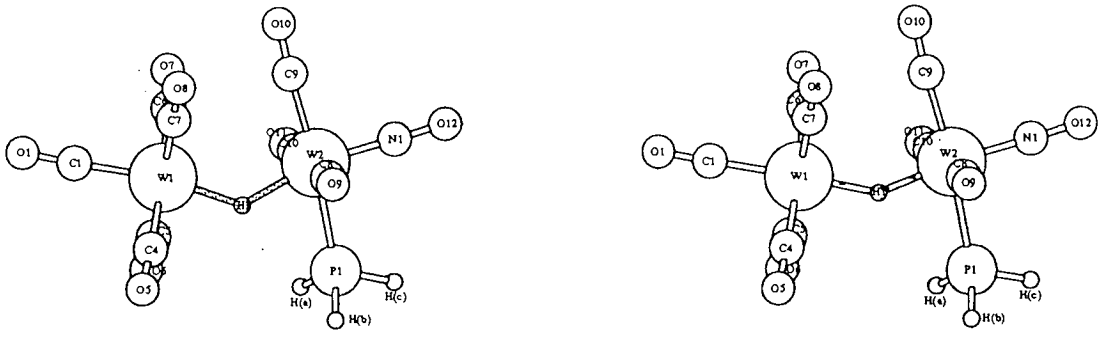


87. $Ru_3H(CO)_9(CCBu^t)$. (HCMBRU11).

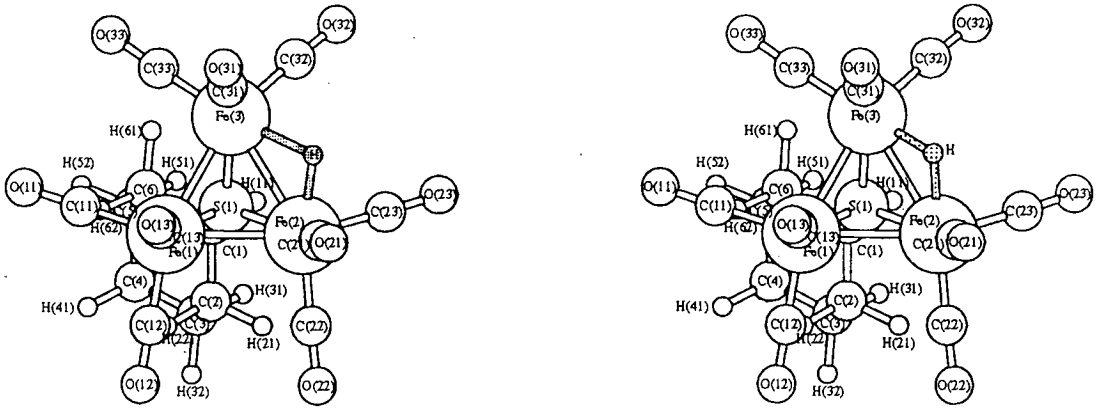




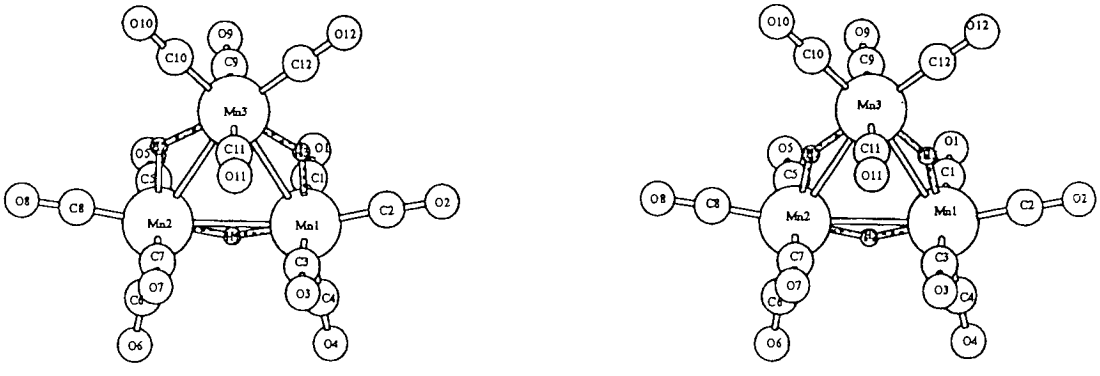
92. $W_2H(CO)_8(NO)(PH_3)$, a model for $W_2H(CO)_8(NO)\{P(OMe)_3\}$. (HWCNMP11).



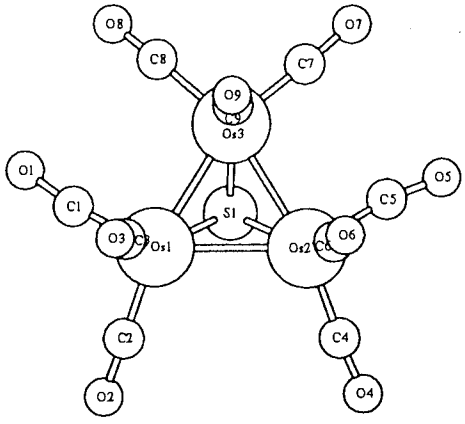
100. $Fe_3H(CO)_9\{S(C_6H_{11})\}$.

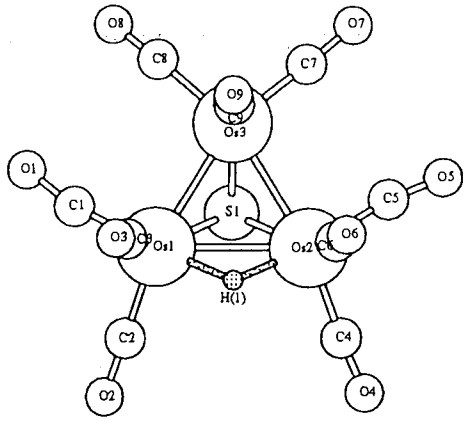


105. $Mn_3H_3(CO)_{12}$. (FUZPOF).

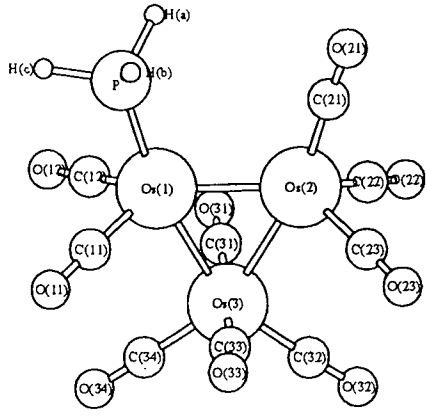


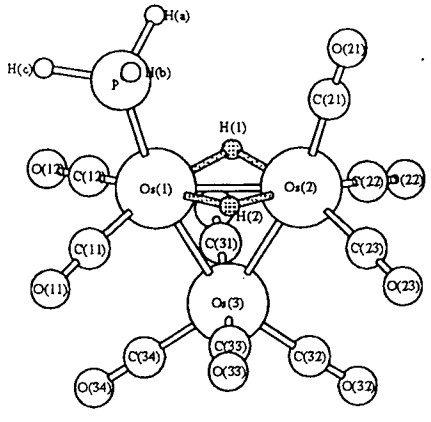
110. $[Os_3H(CO)_9S]^-$. (PIMOSS).



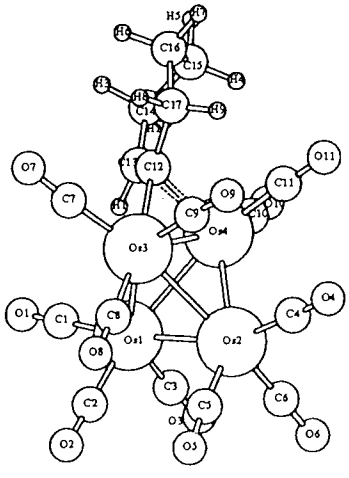


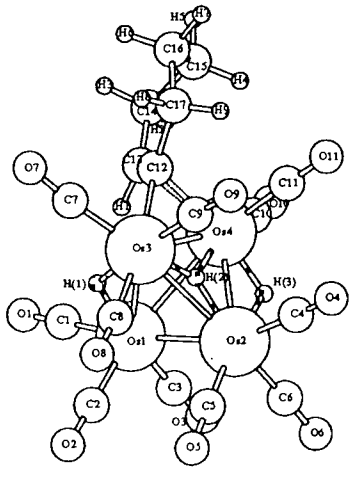
111. $Os_3H_2(CO)_9(PH_3)$, a model for $Os_3H_2(CO)_9(PPh_3)_3$.



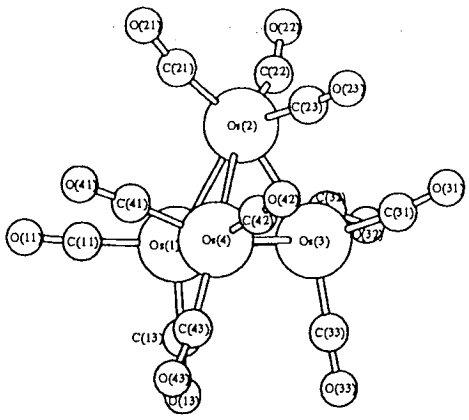
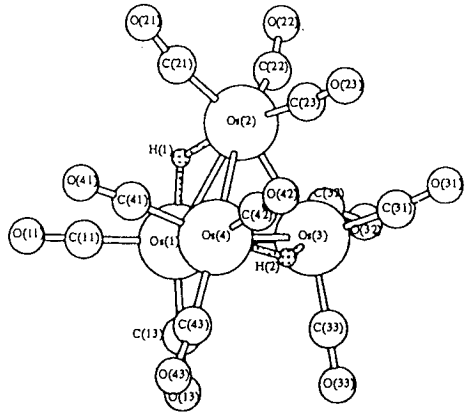


112. $Os_4H_3(CO)_{11}(C_6H_9)$. (UCHXOS).

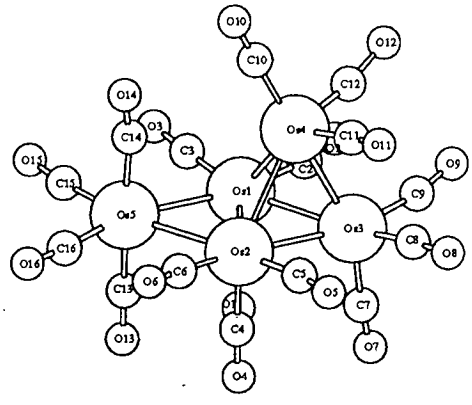
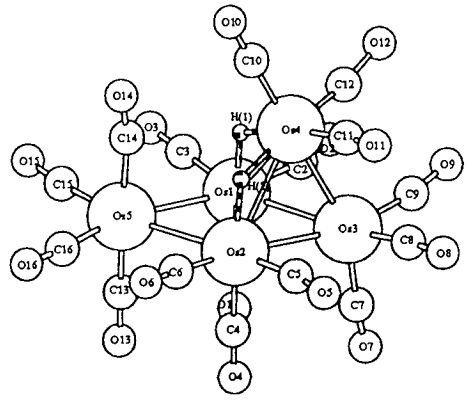




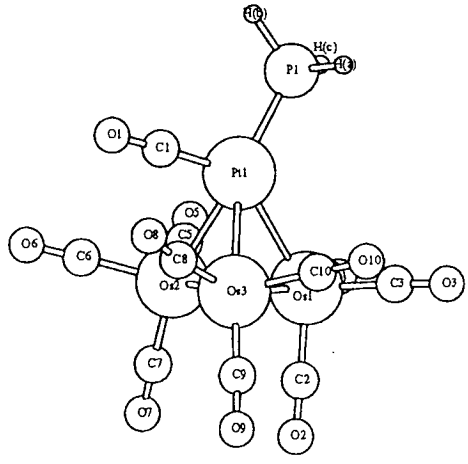
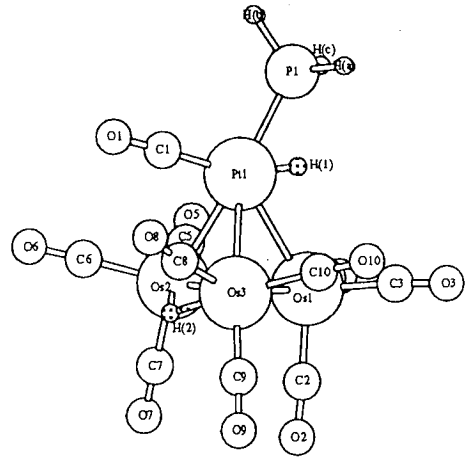
113. $[Os_4H_2(CO)_{12}]^{2-}$. (IMPOSC).

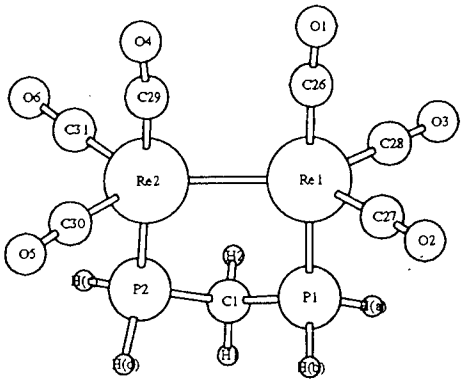
114. $Os_5H_2(CO)_{16}$. (FONGAQ).

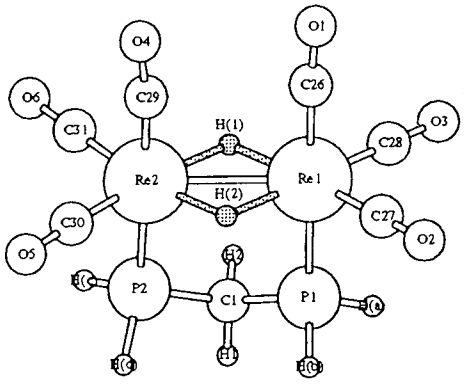



115. $Os_3PtH_2(CO)_{10}(PH_3)$, a model for $Os_3PtH_2(CO)_{10}\{P(C_6H_{11})_3\}$. (HXPTOS).

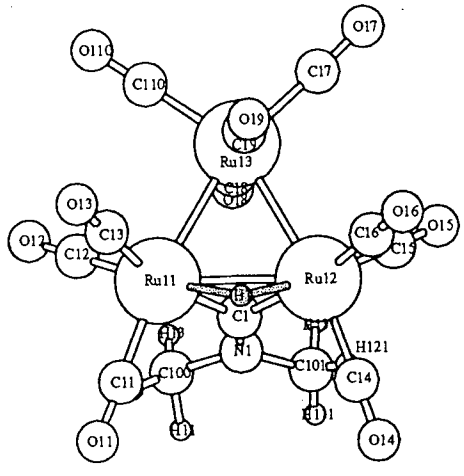



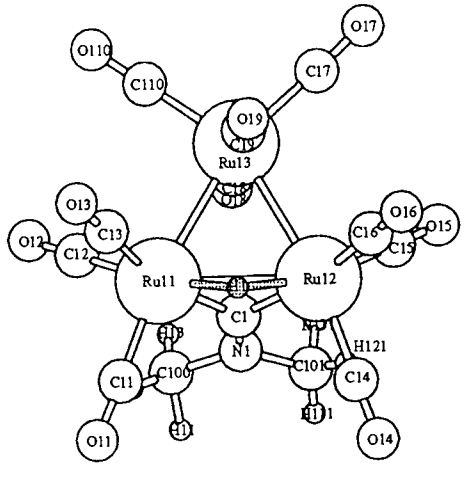
117. $Re_2H_2(CO)_6\{(PH_2)_2CH_2\}$, a model for $Re_2H_2(CO)_6\{(PPh_2)_2CH_2\}$. (PPECRE).



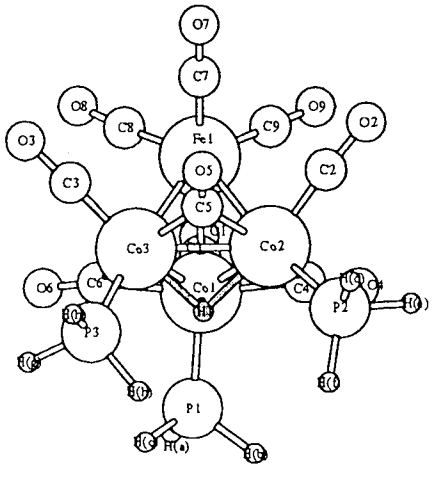


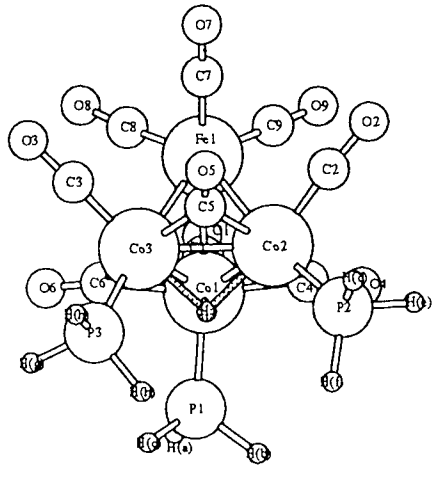
121. $Ru_3H(CO)_{10}(CNMe_2)$. (DCTRIM10).



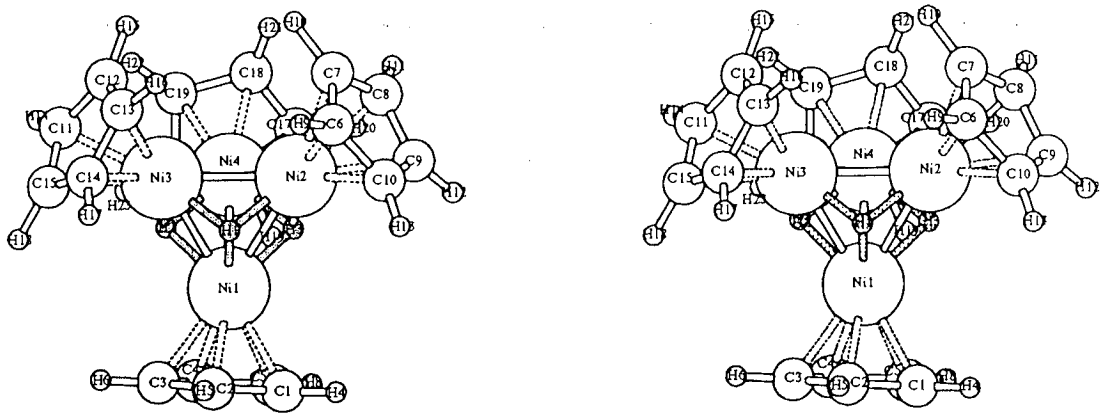


125. $FeCo_3H(CO)_9(PH_3)_3$, a model for $FeCo_3H(CO)_9\{P(OMe)_3\}_3$. (HMPCIC01).

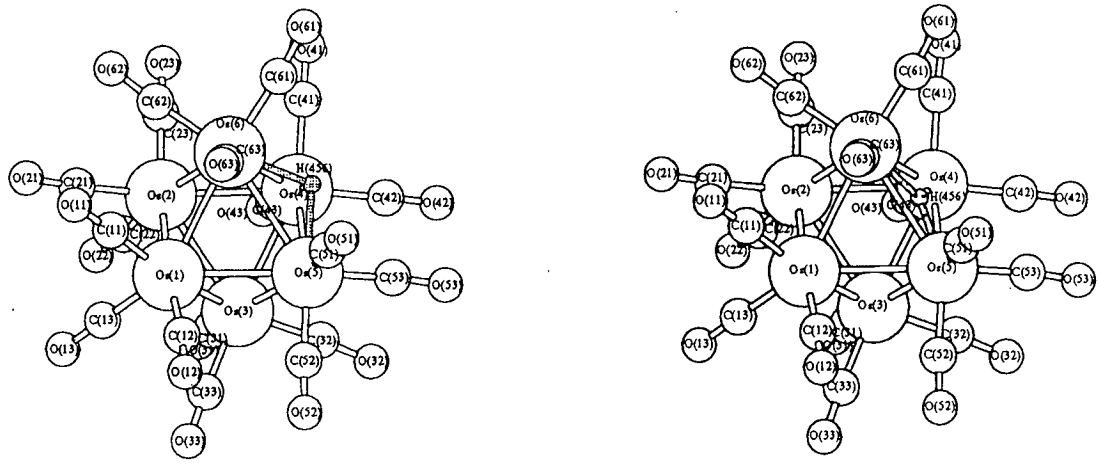




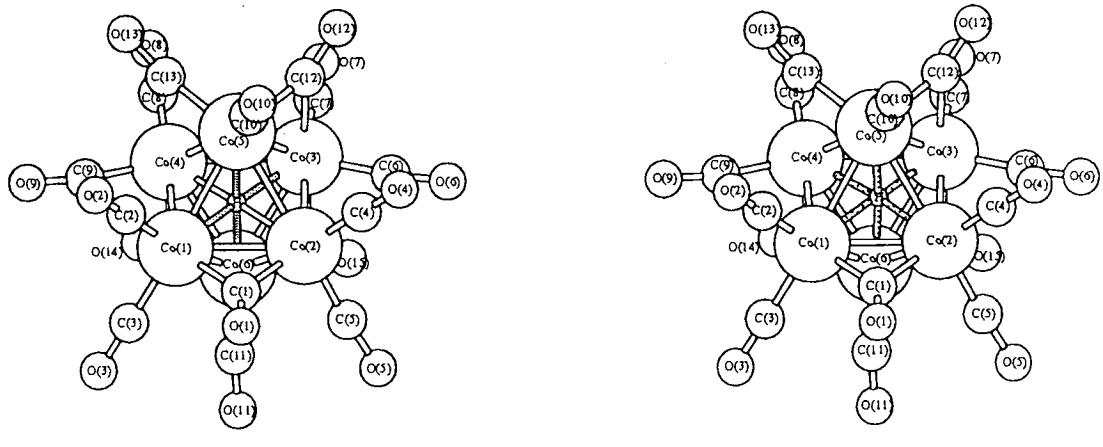
126. $Ni_4H_3(\eta^5\text{–}C_5H_5)_4$. (TCPNIH11).



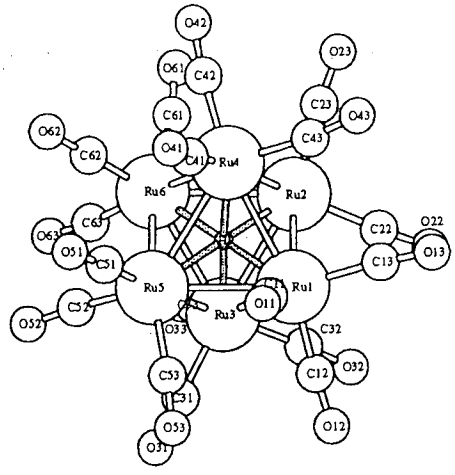
127. $[Os_6H(CO)_{18}]^-$. (FUCPUO01).

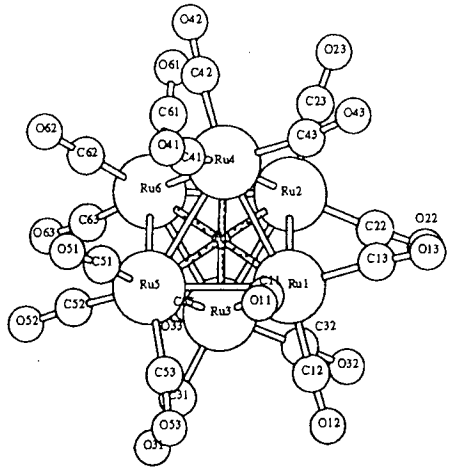


161. $[Co_6H(CO)_{15}]^-$. (PIMHCO10).

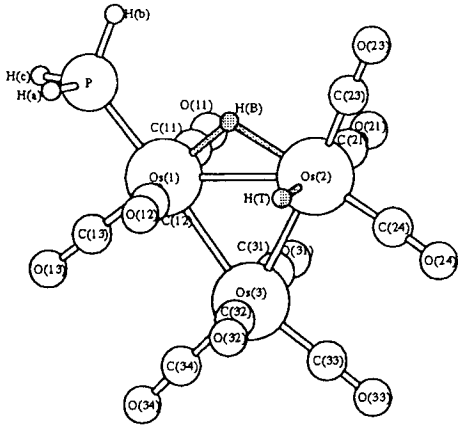


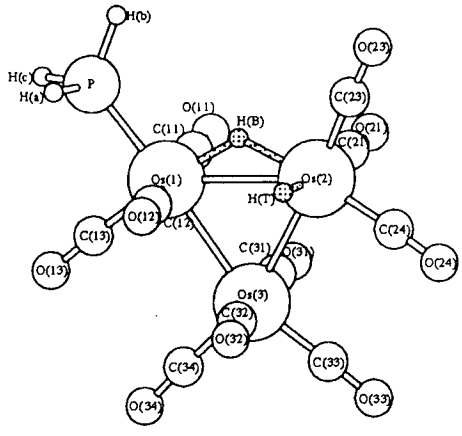
162. $[Ru_6H(CO)_{18}]^-$. (PAHCRU).





174. $Os_3H_2(CO)_{10}(PH_3)$, a model for $Os_3H_2(CO)_{10}(PPh_3)_3$.





Appendix E

The Locator: Raw Results

Tables E–1 to E–7 show the results of the Locator for each of the models listed in Appendix C.

The table headings are:

- *Model:* referring to the model list in Appendix C. A qualifying letter is included for those models which have more than three hydride ligands, for example model 167 B, to indicate which hydride ligands are optimised.

- *nH:* the number of hydride ligands in the model.

- *Rat'ion:* The success, or otherwise, of the automatic rationalisation procedure.

- *Automatic Rationalisation:* Under this main column heading are the results of those experiments for which the locator was able to rationalise the model.

  - *Distance Error* This quantity is described in Chapter 5. Figures for *H(1), H(2), H(3)* are tabulated in order of increasing absolute error.

    * *Initial Estimate:* The distance error(s) of the initial estimate of the hydride ligand position(s).

  - *Optimised Locations:* The distance error(s) of the hydride ligand position(s) once optimised.
  - #: The number of iterations required to locate the optimised positions.
  - The success, or otherwise, of the hydride location procedure.

- *User Rationalisation:* Under this main column heading are the results of those experiments for which the locator either was unable to rationalise the model, or for which this rationalisations was incorrect. The subcategories are the same as above.

The horizontal bold line splitting each of the tables indicates the divide between models determined by neutron diffraction (the top half) and those determined by X-ray diffraction (the bottom half).

Model	nH	Rat'ion	*Automatic Rationalisation*								*User Rationalisation*					
			Distance Error						#		Distance Error				#	
			Initial Estimate			*Optimised Locations*					*Initial Estimate*		*Optimised Locations*			
1	1	Correct	0.2499			0.0361			54	Pass						
2	1	Correct	3.0685			3.1733			86	Fail						
3	2	Correct	0.3583	0.7224		0.1435	0.2369		165	Pass						
4	3	Correct	0.4393	0.9599	1.2322	0.7075	1.1261	1.9953	363	Fail						
5	1	Correct	0.1286			0.0320			46	Pass						
6	2	Correct	1.0962	1.2205		0.1338	0.1388		220	Pass						
7 A	2	Correct	0.0867	0.1771		0.1221	0.1446		146	Pass						
7 B	2	Correct	0.9510	2.9264		0.8132	2.2435		414	Fail						
8 A	3	Correct	0.2482	0.3722	0.6941	0.1262	0.1920	0.3137	286	Pass						
8 B	3	Correct	0.7767	0.8231	0.8664	0.3008	0.3544	0.4100	315	Pass						
9 A	2	Correct	1.1946	1.9519		0.0914	0.1886		522	Pass						
9 B	3	Correct	1.2799	1.8878	1.9281	0.4464	0.9475	1.5488	494	Fail						
10 A	3	Correct	0.4708	0.7635	1.3362	0.1740	0.2700	0.3150	807	Pass						
10 B	2	Correct	0.7773	1.2019		1.0283	1.8087		216	Fail						
11 A	3	Correct	0.3985	0.6485	1.0981	0.1485	0.1963	0.3008	475	Pass						
11 C	1	Correct	0.0401			0.0401			37	Pass						
12	1	Correct	1.3082			1.8804			52	Fail						
13	2	Correct	0.4249	0.7594		0.1764	0.2561		266	Pass						
14	3	Correct	0.4058	0.6587	1.2179	0.1549	0.2192	0.2283	261	Pass						
15	2	Correct	0.9569	3.0512		0.9574	4.6943		253	Fail						
16	2	Correct	0.0539	0.1533		0.2016	0.4812		123	Pass						
17	2	Failed									0.9788	0.9788	15.2821	18.3383	42	Fail
164 B	2	Failed									1.0718	1.1215	0.1135	0.1894	237	Pass
166 C	2	Correct	0.0626	0.0626		0.2441	0.2475		164	Pass						
167 B	1	Failed									0.5953		0.0689		61	Pass
168 B	2	Correct	0.9870	1.8518		0.1667	0.1819		315	Pass						
169 B	2	Correct	0.9927	1.8573		0.2160	0.2278		366	Pass						
18	1	Correct	1.1775			0.0378			63	Pass						
19	1	Correct	0.4990			0.2437			52	Pass						
20	2	Correct	n/a	n/a		pass	fail		107	Fail						
21	2	Correct	0.7936	1.2118		0.1501	0.2191		255	Pass						
22	1	Correct	0.6012			0.3183			54	Pass						
23	1	Correct	0.9037			10.2082			20	Fail						
24	1	Correct	0.7619			0.2387			50	Pass						
25	1	Correct	3.3642			3.4531			66	Fail						
26	2	Correct	1.0941	2.7506		0.4290	0.7046		378	Pass						
27	2	Correct	0.7848	0.9720		0.0837	0.1711		151	Pass						
28	2	Incorrect	2.2939	3.1536		0.3913	4.6384		239	Pass	1.0287	1.2842	0.1991	0.3666	219	Pass
29	2	Correct	1.0050	2.9860		0.2818	0.3463		410	Pass						
30	1	Correct	0.3421			0.4207			50	Pass						
31	1	Correct	0.7407			0.7803			88	Fail						
32	2	Correct	1.0013	1.2518		0.1712	0.1800		222	Pass						
33	1	Correct	0.4765			0.1633			48	Pass						
34	3	Correct	1.1080	1.1219	1.2356	0.1249	0.2770	0.5409	706	Pass						
35	1	Failed									0.8383		0.4203		67	Pass
36	1	Correct	2.6568			0.1570			103	Pass						
37	1	Correct	1.5521			0.1906			70	Pass						
38	1	Correct	1.8791			1.1718			67	Fail						
39	2	Correct	0.4832	1.1001		0.2213	0.2811		173	Pass						
40	2	Correct	0.4145	0.4536		0.3249	0.6401		124	Pass						
41	2	Correct	0.7406	1.1143		0.2388	0.3048		148	Pass						
42	2	Correct	0.9367	1.7553		0.2180	0.3312		133	Pass						
43	3	Correct	0.3633	0.5285	0.8106	0.1839	0.3266	0.3865	294	Pass						
44	1	Failed									1.1783		0.2176		70	Pass
45	2	Correct	0.5460	1.6499		0.0175	1.0996		210	Fail						
46	1	Correct	0.3767			0.2881			82	Pass						
47	1	Correct	0.1623			0.1069			81	Pass						
48	1	Correct	0.5699			0.3776			55	Pass						
49	1	Correct	0.5538			1.0530			72	Fail						
50	1	Correct	0.7715			0.1604			71	Pass						
51	1	Correct	0.5837			0.0791			51	Pass						
52	1	Correct	0.3740			0.2001			69	Pass						
53	2	Correct	1.1068	1.2881		0.1603	0.2085		217	Pass						
54	2	Correct	1.3680	1.9785		0.7559	1.7770		191	Fail						
55	2	Failed									0.0894	0.1134	0.0923	0.1023	112	Pass
56	2	Failed									0.9055	1.4067	0.1640	0.2871	167	Pass
57	2	Failed									1.1600	2.8694	1.0401	3.5719	179	Fail
58	2	Failed									0.9922	2.0229	10.8681	77.2419	106	Fail
59	1	Correct	0.5433			0.1711			50	Pass						
60	1	Correct	0.4866			0.1640			84	Pass						
61	1	Correct	0.3308			0.3302			57	Pass						
62	2	Correct	1.1234	1.2099		1.2956	4.8068		408	Fail						
63	1	Correct	0.0794			0.1150			42	Pass						
64	2	Correct	1.1324	1.2568		1.3584	1.5520		345	Fail						

Table E–1: Models with Terminal Hydride Ligands.

Model	nH	Rat'ion	*Automatic Rationalisation*								*User Rationalisation*							
			Distance Error						#		Distance Error						#	
			Initial Estimate			*Optimised Locations*					*Initial Estimate*			*Optimised Locations*				
65	1	Correct	0.3851			0.2789			57	Pass								
66	1	Failed									0.7863			0.0086			51	Pass
67	2	Incorrect	1.2670	1.6516		0.0479	0.5758		169	Pass	1.2670	1.3345		0.0581	0.5556		485	Pass
68	2	Correct	1.1194	2.9713		0.3555	4.6605		212	Fail								
69	1	Correct	0.0546			0.1423			54	Pass								
70	1	Correct	0.4422			0.1660			55	Pass								
71	1	Failed									0.2458			0.3179			50	Pass
72	1	Correct	0.0245			0.0633			59	Pass								
73	1	Failed									0.0368			0.2623			73	Pass
74	1	Incorrect	1.1329			0.1961			66	Pass	0.0656			0.1993			64	Pass
75	1	Correct	0.0444			0.1821			61	Pass								
76	2	Correct	0.0648	0.1833		0.1596	0.1596		164	Pass								
77	2	Incorrect	0.8615	1.8998		0.1243	0.2248		374	Pass	0.0497	1.6310		0.1389	0.2241		286	Pass
78	2	Failed									0.1249	3.1305		0.2919	3.6867		341	Fail
79	3	Incorrect	1.2607	1.3287	2.2007	0.1064	0.1163	0.1399	1445	Pass	0.0653	0.0905	0.1258	0.1191	0.1306	0.1309	329	Pass
80	3	Incorrect	0.4565	0.4917	3.3293	0.1926	0.2838	1.7090	875	Fail	0.2117	0.2340	0.3368	0.2523	0.2596	0.3793	268	Pass
81	3	Failed									0.1047	0.2772	0.2772	0.3104	0.3863	0.4077	239	Pass
82 A	2	Failed									0.1067	0.1177		0.3373	0.3559		185	Pass
82 B	2	Failed									0.1012	0.1033		0.3119	0.3194		210	Pass
83	3	Incorrect	0.5105	1.6909	2.5731	0.2158	0.2870	0.3551	1230	Pass	0.2703	0.4684	0.7386	0.2004	0.2936	0.3541	325	Pass
84	3	Incorrect	0.1757	0.5543	1.9475	0.2881	0.3185	0.3968	336	Pass	0.1802	0.2357	0.3565	0.2987	0.3281	0.3834	392	Pass
85	2	Failed									0.6989	0.6989		0.2341	0.2391		170	Pass
86	3	Incorrect	1.3538	2.1602	2.6650	0.1124	0.1535	0.2689	925	Pass	0.1937	0.7905	1.0453	0.1037	0.1708	0.2298	454	Pass
87	1	Incorrect	1.6518			0.1954			61	Pass	0.0969			0.2035			48	Pass
88	1	Failed									0.0983			0.2670			70	Pass
89	1	Incorrect	2.8205			0.3167			169	Pass	0.0517			0.3281			61	Pass
90 A	2	Failed									0.1470	0.1873		0.7065	0.7228		125	Pass
90 B	2	Failed									0.1267	0.1748		0.6755	0.6955		155	Pass
91	1	Failed									0.0648			0.3656			87	Pass
92	1	Failed									0.0901			0.3735			57	Pass
93	2	Failed									0.1346	0.1444		0.0973	0.1041		101	Pass
94	1	Failed									0.1048			0.3443			55	Pass
164 A	2	Failed									0.1251	0.4373		0.1570	0.1640		200	Pass
166 A	2	Correct	0.1324	0.1324		0.2007	0.2024		111	Pass								
167 A	3	Failed									0.8533	1.0719	1.5075	0.1002	0.1388	0.1731	807	Pass
168 A	3	Correct	0.2451	0.8557	0.8834	0.1053	0.1968	0.3006	439	Pass								
169 A	3	Correct	0.4704	1.3040		0.0698	0.1749		614	Pass								
95	1	Failed									0.2957			0.2377			41	Pass
96	2	Failed									0.3910	0.4496		0.2247	0.2419		173	Pass
97	2	Failed									0.3509	1.0217		0.2212	0.3955		152	Pass
98	1	Failed									0.5455			0.1853			55	Pass
99	3	Correct	0.1335	0.2239	0.5017	0.1239	0.2136	0.3260	387	Pass								
100	1	Incorrect	1.6722			0.3229			63	Pass	0.2421			0.3099			53	Pass
101	1	Incorrect	0.4195			0.2888			59	Pass	0.4730			0.3254			45	Pass
102	1	Correct	0.2433			0.1507			71	Pass								
103	1	Incorrect	4.3052			4.1978			68	Fail	0.2505			0.1181			48	Pass
104	1	Incorrect	1.1754			0.1170			63	Pass	0.3529			0.1305			44	Pass
105	3	Incorrect	0.4314	0.5557	1.3164	0.1337	0.3559	0.3680	460	Pass	0.3565	0.4532	0.5860	0.1297	0.3446	0.3703	387	Pass
106	1	Failed									0.1013			0.2046			51	Pass
107	1	Failed									0.6390			0.2308			60	Pass
108	1	Failed									0.7525			0.1240			64	Pass
109	3	Correct	0.1397	0.5018	0.7659	0.1450	0.2310	0.2399	344	Pass								
110	1	Incorrect	n/a			yes			63	Pass	n/a			yes			51	Pass
111	2	Correct		n/a		pass	pass		125	Pass								
112	3	Incorrect	n/a	n/a	n/a	yes	yes	no	340	Fail	n/a	n/a	n/a	yes	yes	no	561	Fail
113	2	Failed									n/a	n/a		yes	yes		133	Pass
114	2	Incorrect	n/a	n/a		yes	no		316	Fail	n/a	n/a		yes	yes		195	Pass
115	2	Failed									1.2412	1.7033		0.5184	1.6411		418	Fail
116	1	Failed									0.7062			0.5225			63	Pass
117	2	Correct	n/a	n/a		yes	yes		80	Pass								
118	1	Failed									0.5359			0.2380			66	Pass
119	2	Failed									0.2935	0.3605		0.2332	0.6085		203	Pass
120	1	Incorrect	0.7550			0.0634			62	Pass	0.3214			0.0608			46	Pass
121	1	Correct	0.1527			0.2683			50	Pass								
122	1	Correct	0.3542			0.4938			46	Pass								
123	3	Incorrect	0.8517	0.8722	1.8552	0.1057	0.1410	0.1556	981	Pass	0.4883	0.4883	0.7803	0.1345	0.1376	0.1558	489	Pass

Table E–2: Models with Edge-Bridging Hydride Ligands.

Model	nH	Rat'ion	*Automatic Rationalisation*								*User Rationalisation*							
			Distance Error						#		Distance Error						#	
			Initial Estimate			*Optimised Locations*					*Initial Estimate*			*Optimised Locations*				
124 A	2	Failed									0.4335	0.4800		0.4758	0.4895		192	Pass
124 B	2	Failed									0.3199	0.3501		0.2004	0.2044		142	Pass
124 C	2	Failed									0.5979	0.6298		0.5038	0.5051		219	Pass
125	1	Incorrect	3.8775			2.8531			63	Fail	0.2226			0.1521			54	Pass
126	3	Failed									0.2933	1.3156	1.3730	0.0445	0.0453	0.2090	268	Pass
127	1	Incorrect	2.0462			0.7250			59	Pass	0.2065			0.7229			62	Pass
128 A	2	Failed									0.1091	0.1091		0.1685	0.1760		175	Pass
128 B	2	Failed									0.0632	0.6320		0.1616	0.1757		175	Pass
129	1	Failed									0.5849			0.7028			42	Pass
130	1	Failed									0.5441			0.3121			56	Pass
131	2	Failed									0.1510	0.4178		0.1710	0.2124		121	Pass
132	1	Incorrect	3.8370			2.8069			69	Fail	0.3114			0.0440			46	Pass
133	1	Incorrect	3.7611			2.6892			60	Fail	0.5312			0.3779			49	Pass
134	1	Incorrect	1.9340			0.6497			111	Pass	0.4027			0.6559			65	Fail
135	1	Incorrect	0.8927			0.1931			55	Pass	0.1119			0.1869			52	Pass
136	1	Incorrect	4.8194			4.7669			57	Fail	0.2108			0.1796			49	Pass
137	1	Incorrect	3.0117			2.9718			43	Fail	0.3051			0.0463			80	Pass
138	1	Failed									0.1082			0.1162			42	Pass
139	1	Failed									0.5731			0.4280			47	Pass
140	2	Failed									0.2596	0.2757		0.1360	0.1455		187	Pass
141	3	Incorrect	1.9096	2.0011	4.2713	0.0870	0.2318	1.7748	526	Fail	0.1843	0.2265	0.2759	0.2439	0.3107	0.3934	246	Pass
142	2	Incorrect	3.4106	3.8604		1.6595	8.0640		198	Fail	0.3116	0.4137		0.4017	0.5851		194	Pass
143	2	Incorrect	1.1028	3.2782		1.3275	1.8750		156	Fail	0.5763	0.7120		0.3503	0.4376		167	Pass
144	2	Incorrect	1.8768	4.3854		0.1757	1.9728		410	Fail	0.1281	0.1484		0.3393	0.4023		137	Pass
145	2	Correct	0.0569	0.4212		0.0715	0.0742		176	Pass								
146	1	Incorrect	0.9274			0.3263			58	Pass	0.5503			0.3224			41	Pass
147	1	Incorrect	1.2865			0.2496			74	Pass	0.0536			0.2435			58	Pass
148	1	Failed									0.2069			9.3119			15	Fail
149	2	Failed									0.3406	0.4148		0.1854	0.2218		253	Pass
150	1	Incorrect	2.4160			1.6715			58	Fail	0.3853			1.1542			69	Fail
151	2	Failed									0.3055	0.3348		5.1024	5.1334		29	Fail
152	1	Incorrect	0.7327			0.3765			55	Pass	0.1566			0.3782			59	Pass
153	1	Failed									0.2942			0.3934			60	Pass
154	1	Failed									0.2064			0.3032			51	Pass
155	2	Incorrect	1.3918	1.3918		0.1651	0.1825		284	Pass	0.2275	0.2275		0.1718	0.1938		127	Pass
156	1	Incorrect	3.9366			0.1514			148	Pass	0.0903			0.1582			51	Pass
157	1	Incorrect	0.4953			0.2858			64	Pass	0.1641			0.2916			43	Pass
158	1	Incorrect	3.8244			6.2159			12	Fail	0.2399			3.4572			15	Fail
159	2	Incorrect	3.2548	3.5398		1.4911	2.8529		238	Fail	0.2604	0.2652		0.3630	0.3779		108	Pass
160	2	Incorrect	0.8038	4.3600		0.4434	3.7351		192	Fail	0.1424	0.1912		0.4084	0.4894		146	Pass

***Table E–3:** Models with Face-Bridging Hydride Ligands.*

Model	nH	Rat'ion	*Automatic Rationalisation*								*User Rationalisation*							
			Distance Error						#		Distance Error						#	
			Initial Estimate			*Optimised Locations*					*Initial Estimate*			*Optimised Locations*				
161	1	Failed									0.0242			0.0331			48	Pass
162	1	Correct	0.0078			0.0308			34	Pass								
163	1	Failed									0.1771			3.0589			47	Fail

***Table E–4:** Models with Interstitial Hydride Ligands.*

Model	nH	Rat'ion	*Automatic Rationalisation*								*User Rationalisation*							
			Distance Error						#		Distance Error						#	
			Initial Estimate			*Optimised Locations*					*Initial Estimate*			*Optimised Locations*				
165	3	Incorrect	0.4696	1.0003	1.2795	0.1128	0.1172	0.1367	482	Pass	1.0013	1.2795	1.4504	0.1046	0.1188	0.1282	825	Pass
170	3	Failed									0.2324	0.2641	0.2698	0.2207	0.3161	0.5209	410	Pass
171	2	Incorrect	4.0885	4.3695		1.2765	20.0327		199	Fail	0.3194	0.3906		0.3249	0.5129		161	Pass
172	2	Failed									0.2073	0.2092		0.4165	0.9085		115	Pass
173	2	Failed									0.1556	0.3386		0.1622	0.3890		182	Pass
174	2	Incorrect	1.1910	1.5985		0.1584	0.3751		420	Pass	0.1765	0.4644		0.1216	0.3398		132	Pass
175	2	Incorrect	1.0929	1.2051		0.3155	0.4010		399	Pass	0.1672	0.4315		0.3228	0.3943		116	Pass
176	2	Failed									0.1912	0.9583		0.1460	0.4844		107	Pass
177	2	Failed									0.1478	0.3271		2.8145	3.2330		43	Fail
178	2	Failed									0.1724	0.1833		3.3109	4.1331		49	Fail

***Table E–5:** Models with Mixed Hydride Ligands.*

Model	nH	Rat'ion	*Automatic Rationalisation*						*User Rationalisation*					
			Distance Error				#		Distance Error				#	
			Initial Estimate		*Optimised Locations*				*Initial Estimate*		*Optimised Locations*			
179	2	Correct	0.4295	2.1849	0.4127	2.4188	446	Fail						
180	2	Correct	0.6414	0.9533	4.7556	5.3270	156	Fail						
181	2	Correct	0.7418	0.8281	0.8983	0.9267	252	Fail						
182	2	Failed							1.1491	2.5690	0.3889	2.1958	179	Fail
183	2	Correct	0.8855	1.6000	0.2302	3.0850	198	Fail						
184	2	Correct	0.3565	3.1941	0.1909	4.8361	346	Fail						
185	2	Correct	0.8223	2.9918	1.6645	4.8594	299	Fail						
186	2	Correct	0.5972	1.3855	3.5173	4.3188	458	Fail						

(a) Using the Biggest Appropriate Hole Method for a Starting Point

Model	nH	Rat'ion	*Automatic Rationalisation*						*User Rationalisation*					
			Distance Error				#		Distance Error				#	
			Initial Estimate		*Optimised Locations*				*Initial Estimate*		*Optimised Locations*			
179	2	Correct	0.4295	0.4295	6.6567	7.5544	265	Fail						
180	2	Correct	0.3229	0.6651	4.2230	4.9244	203	Fail						
181	2	Correct	0.7324	0.7381	0.9064	0.9250	143	Fail						
182	2	Failed							1.0918	1.7719	4.8470	5.3540	403	Fail
183	2	Correct	0.3606	0.9950	3.6536	4.2583	143	Fail						
184	2	Correct	0.3359	1.2451	0.5129	0.8186	177	Fail						
185	2	Correct	0.5059	1.5255	0.0637	0.7300	149	Fail						
186	2	Correct	0.4625	0.4643	5.7225	6.1729	147	Fail						

(b) Using the Projection Method for a Starting Point

Model	nH	Distance Error		#	
		Initial Estimate			
179	2	6.2684	6.8323	186	Fail
181	2	0.9254	0.9395	320	Fail
180	2	16.6133	16.973	176	Fail
186	2	4.8651	4.6803	224	Fail
183	2	4.0458	4.5923	350	Fail
185	2	0.0489	0.7122	131	Fail
182	2	17.5585	16.8436	220	Fail
184	2	0.2503	1.5622	260	Fail

(c) Using the Literature Coordinates for a Starting Point

Table E–6: *Models with Dihydrogen Hydride Ligands.*

Model	nH	Rat'ion	*Automatic Rationalisation*								*User Rationalisation*							
			Distance Error						#		Distance Error						#	
			Initial Estimate			*Optimised Locations*					*Initial Estimate*			*Optimised Locations*				
5	1	Correct	0.1038			0.0301			44	Pass								
6	2	Correct	1.0364	1.1558		5.9282	6.8794		162	Fail								
14	3	Correct	0.0831	1.8015	1.8257	*program failed*				Fail								
20	2	Correct	n/a	n/a		pass	fail		138	Fail								
72	1	Correct	0.0376			0.0801			61	Pass								
74	1	Incorrect	2.8926			3.2526			57	Fail	0.2435			0.1881			60	Pass
76	2	Correct	1.4375	1.4670		0.1567	0.1700		203	Pass								
77	2	Incorrect	2.8723	2.9278		2.9572	3.1982		164	Fail	0.3690	0.5920		0.1355	0.2365		152	Pass
78	2	Failed									0.5452	0.6100		0.1407	0.1435		166	Pass
80	3	Incorrect	2.6604	2.8728	2.9786	0.2987	5.0737	8.5013	247	Fail	0.3655	0.5760	0.7689	0.2454	0.2494	0.4024	276	Pass
83	3	Incorrect	3.0431	3.0611	3.2451	0.3208	0.8653	2.6038	1422	Fail	0.1276	0.1638	0.2282	0.2072	0.2824	0.3544	182	Pass
86	3	Incorrect	2.6321	2.9201	3.2902	0.1249	0.1579	0.5902	1174	Pass	0.2455	1.6245	1.7622	0.1271	0.2155	0.2366	1182	Pass
87	1	Incorrect	4.5733			4.7990			49	Fail	0.1463			0.1929			54	Pass
92	1	Failed									0.3611			0.2469			54	Pass
100	1	Incorrect	1.6722			0.3229			63	Pass	0.2416			0.3276			50	Pass
105	3	Incorrect	2.9745	3.0897	3.1231	3.0429	3.1557	3.1820	496	Fail	0.1078	0.2753	0.6030	0.1362	0.3298	0.3872	236	Pass
110	1	Incorrect	n/a			pass			63	Pass	n/a			pass			58	Pass
111	2	Correct	n/a	n/a		pass	fail		220	Fail								
112	3	Incorrect	n/a	n/a	n/a	fail	fail	fail	936	Fail	n/a	n/a	n/a	pass	pass	fail	1101	Fail
113	2	Failed									n/a	n/a		pass	pass		209	Pass
114	2	Incorrect	n/a	n/a		pass	fail		298	Fail	n/a	n/a		pass	pass		164	Pass
115	2	Failed									0.9181	1.3699		0.8525	1.7318		196	Fail
117	2	Correct	n/a	n/a		pass	pass		154	Pass								
121	1	Correct	0.1476			0.2720			54	Pass								
125	1	Incorrect	4.7873			4.7476			52	Fail	0.2226			0.1518			54	Pass
126	3	Failed									0.2920	0.2933	0.3198	0.0545	0.0590	0.0808	440	Pass
127	1	Incorrect	2.0462			0.7250			59	Pass	0.2065			0.7229			62	Pass
161	1	Failed									0.0242			0.1557			48	Pass
162	1	Correct	0.0078			0.0308			34	Pass								
174	2	Incorrect	2.9317	2.1262		13.0772	17.3002		52	Fail	0.2718	2.1262		0.3360	4.4671		170	Fail

Table E–7: Projection Results.

Appendix F

Comparison of Initial Estimate Methods

For the purpose of comparison between the two methods it was considered unnecessary to include all 178 models. Thus only those models used by Mitchell for his evaluation of CCCP, and Orpen in his evaluation of Hydex, are used.

Table F–1 shows the success rates for the two initial estimate methods. The projection method for the initial estimate of a hydride ligand position had only a modest rate of success. The projection method had only a 42% overall success rate, 43% for the neutron diffraction models and 40% for the X-ray diffraction models. The success rate improved when erroneous rationalisation was corrected by the user, affording a 77% success rate — 89% for the neutron diffraction models and 58% for the X-ray diffraction models.

By comparison, the *biggest appropriate hole method* was considerably more successful than the projection method. This had an overall success rate of 79% without user rationalisation, improving to 87% with appropriate user intervention at the rationalisation stage. This is a far more satisfactory situation than that for the projection method.

	Automatic Rationalisation						*User Rationalisation*					
	Neutron		*Xray*		*Total*		*Neutron*		*Xray*		*Total*	
Projection	6 / 14	(43%)	4 / 10	(40%)	10 / 24	(42%)	16 / 18	(89%)	7 / 12	(58%)	23 / 30	(77%)
Biggest Hole	12 / 14	(86%)	7 / 10	(70%)	19 / 24	(79%)	17 / 18	(94%)	9 / 12	(75%)	26 / 30	(87%)

Table F–1: A comparison between the two initial estimate methods with respect to their success rates.

	Automatic Rationalisation						*User Rationalisation*					
	Neutron		*Xray*		*Total*		*Neutron*		*Xray*		*Total*	
Projection	22	2.341 ±0.558	7	5.765 ±2.522	29	3.167 ±0.763	29	0.624 ±0.299	9	0.982 ±0.465	38	0.709 ±0.252
Biggest Hole	25	0.369 ±0.123	7	0.283 ±0.038	32	0.350 ±0.096	32	0.309 ±0.112	9	0.449 ±0.154	41	0.340 ±0.093

(a) considering all models.

	Automatic Rationalisation						*User Rationalisation*					
	Neutron		*Xray*		*Total*		*Neutron*		*Xray*		*Total*	
Projection	9	0.230 ±0.084	2	0.297 ±0.025	11	0.242 ±0.068	27	0.196 ±0.027	5	0.291 ±0.043	32	0.211 ±0.024
Biggest Hole	21	0.199 ±0.032	7	0.283 ±0.038	28	0.220 ±0.026	30	0.197 ±0.025	7	0.269 ±0.039	37	0.211 ±0.022

(b) considering only those models for which the Locator suceeded.

Table F–2: A comparison between the two initial estimate methods with respect to their accuracy.

	Automatic Rationalisation						*User Rationalisation*					
	Neutron		*Xray*		*Total*		*Neutron*		*Xray*		*Total*	
Projection	14	287 ±122	10	247 ±88	24	270 ±78	18	199 ±67	12	229 ±66	30	212 ±51
Biggest Hole	14	317 ±106	10	202 ±51	24	269 ±65	18	170 ±31	12	191 ±49	30	178 ±27

(a) considering all models.

	Automatic Rationalisation						*User Rationalisation*					
	Neutron		*Xray*		*Total*		*Neutron*		*Xray*		*Total*	
Projection	6	263 ±184	4	84 ±24	10	191 ±111	16	201 ±73	7	132 ±50	23	179 ±51
Biggest Hole	12	292 ±112	7	180 ±68	19	251 ±74	17	160 ±32	9	126 ±37	26	149 ±24

(b) considering only those models for which the Locator suceeded.

Table F–3: A comparison between the two initial estimate methods with respect to their efficiency.

Tables F–2 and F–3 show the average accuracy and efficiency for the two initial estimate methods respectively. It can be seen that the projection method offers no improvement over the biggest appropriate hole method with regards to accuracy or efficiency, when considering all models, or only those for which the Locator succeeded.

In conclusion it would seem that the biggest appropriate hole method is the superior of the two methods for estimating the position for non-interstitial hydride ligands. Since the topology of the hydride ligand(s) was never intended to be a necessary input of the Locator, but rather a supplement should the information be available, the projection method is considered to be unsatisfactory in providing estimated hydride position(s) for subsequent optimisation.

Appendix G

Publications

1. *"An Interactive Molecular Graphics Tool based on the X Window System"* was published in the Proceedings of The European X Window System Conference, London, November 1990.

2. *"Hydride Ligand Location in Complexes of the Copper Triad and Other Systems"* was published in the Proceedings of the International Conference of the Chemistry of the Copper & Zinc Triads, Edinburgh, July 1992.

The publishers accept that I retain the right to republish this paper and abstract in whole or in part.

AN INTERACTIVE MOLECULAR GRAPHICS TOOL BASED ON THE X WINDOW SYSTEM

by R. J. Pooley,* G. D. M. Ross,* D. A. Welch* and A. J. Welsh*

ABSTRACT

This paper describes the porting of a simple molecular graphics package, designed and implemented originally under SunCGI, to the X Window system. The result is an easy to use and interactive tool for the practising chemist, presenting the user with a fast and informative view of his/her target molecule. The object in the picture can be edited, rotated, translated, or aesthetically altered using different colour palettes and/or different parameters for particular sub-objects.

INTRODUCTION

Molecular graphics packages currently available to the practising chemist leave a lot to be desired in terms of efficiency and 'user-friendliness'. They tend to consist mainly of large non-interactive FORTRAN (refs 1–4) and/or command-line controlled (refs 5–7) drawing programs for mainframe computers, PC-based drawing programs (refs 8, 9) which produce unappealing stick or space-filling diagrams, or programs which require the prohibitively expensive Evans and Sutherland workstation. Since distributed computing with bitmapped screens is becoming the standard in all areas of scientific application, and tools such as the subject of this paper are not available on these machines, a molecular graphics editor was designed and implemented in-house.

REASONS FOR PORTING TO X

SunCGI (ref 10) was initially chosen since, although more machine-specific than the version of X Windows then available, it was faster (X10 was noticeably slower than SunCGI when the same demonstration program was tried on both systems) and documentation more accessible. With the release of the much faster and more powerful X11 and of SunOS 4.1 which no longer supports SunCGI (replacing it with SunGKS) this argument was no longer viable.

Comparing the two toolkits it appeared that there were several advantages to using X: Since X is rapidly becoming the industry standard its inherent portability enables any applications program to be used on many more machines; SunCGI is limited to Sun Workstations running SunOS 4.03 or lower. The Xtoolkit with the Athena widget set provides many tools and widgets: these must be written independently if using SunCGI. The use a single widget set throughout would give a more uniform look and afeel to the application: the CGI toolkit encourages no such uniformity, hence an interface less than satisfactory in aspects of HCI. The extra facilities provided by the Xtoolkit give two advantages: the same program can be implemented in a more efficient and customisable way by using predefined widgets and new features such as three-dimensional graphics and a viewport can be exploited.

*Department of Computer Science, University of Edinburgh, Kings Buildings, Edinburgh, EH9 3JZ

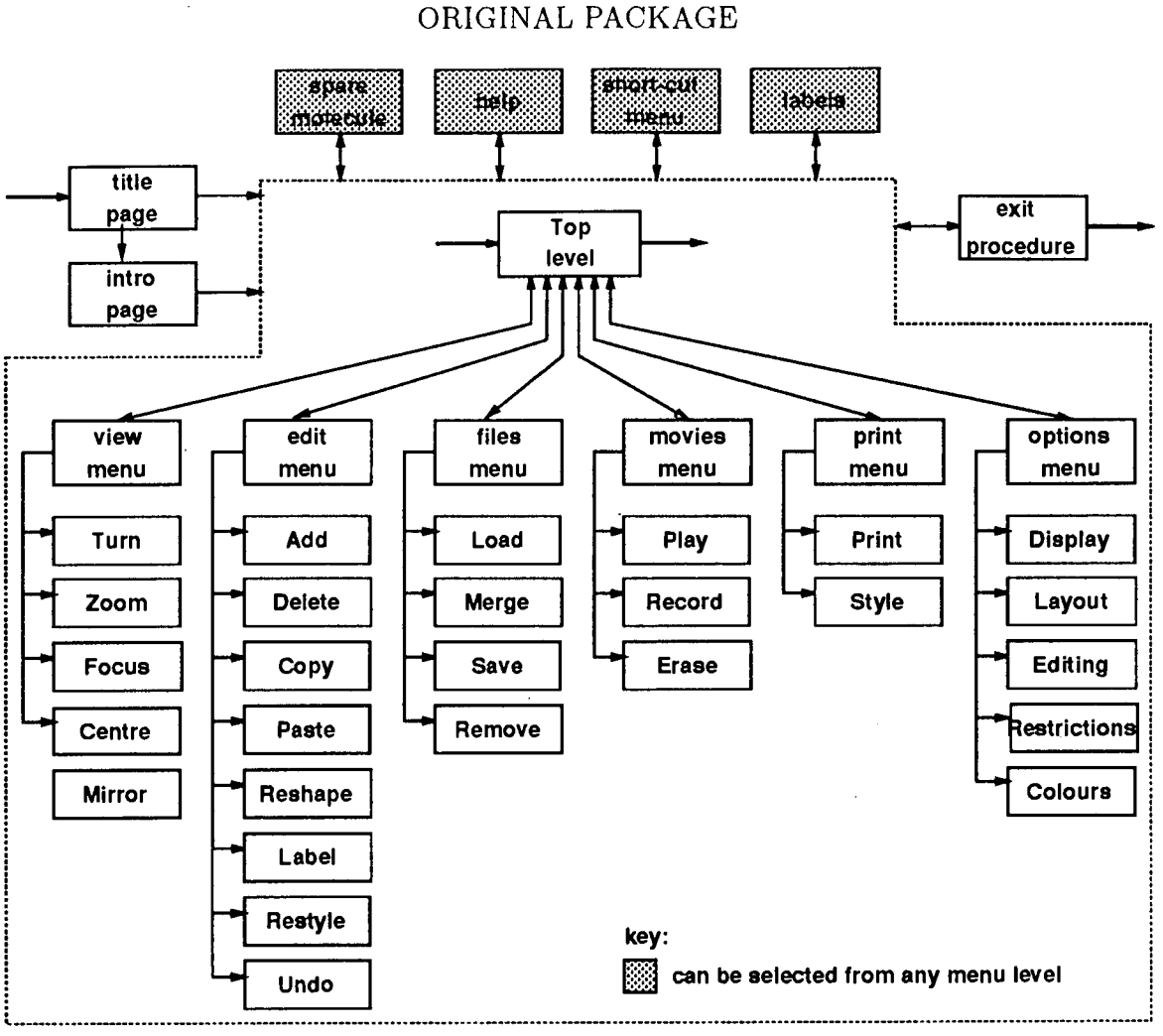


Figure 1: *A graphical representation of the hierarchical menu structure in the molecular editor, displaying all possible operations*

In the specification of the original system (ref 11), the package was required to be fully interactive (operations to be carried out in real time with a mouse by means of menu choices, buttons, sliders etc.) and intuitive (i.e. a user should be able to operate the program with only limited use of a manual or of the 'help' system). The CGI-based program allows the user to build up molecules from previously defined primitives (atoms and bonds) in turn defined in terms of CGI circles, polygons, rectangles and lines, either interactively or by loading crystallographic data. Figure 1 shows the menu structure of the final CGI implementation and includes the operations available to the user. This structure was used to build the X implementation. Features included the ability to build and manipulate molecules in real time using experimental or library data and store whole or partial molecules for future needs. The program had some user changeable knowledge of chemistry and informed the user investigating the effects of changing particular attributes when there were unlikely valences and bond lengths. The resultant molecule could be obtained in hard copy form, stored in a user library or incorporated as part of a "movie" — a prerecorded sequence of events. The system was also designed to be self-documenting.

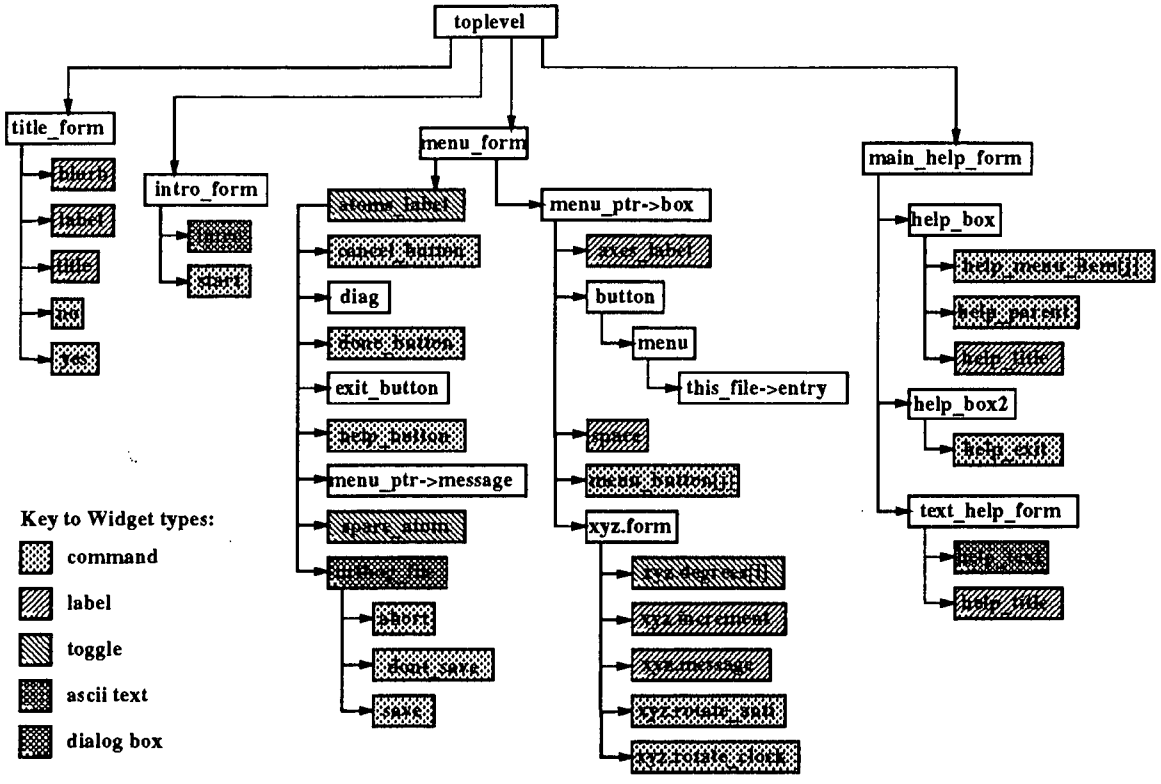


Figure 2: *A graphical representation of the hierarchical widget structure in the X program*

The new version of the package, now running under X, has the same *functional* specification as the original. Elements in the overall design, however, particularly with reference to HCI issues, have been altered: widgets such as buttons, pop-up menus and scrollbars have been used in the new implementation of the molecular graphics editor. Figure 2 shows the hierarchical widget map for the widgets used so far. Improvements in the design of the molecular graphics package can be illustrated by comparing figures 3 and 4 overleaf. Figure 3 shows the frame for rotation of the molecule from the CGI program: the frame consists of an instruction line across the top of the screen, a button icon for the spare molecule (denoted by a triatomic molecule) — a feature which allows the user to cut and paste between molecules, and a button icon (labelled 'L') to show atomic labels. The menu bar down the side consists of a set of axes to indicate orientation to the user, three circular sliders (for each of the XY, YZ and XZ planes) for the rotation of the molecule, a help button which initiates a help frame giving the users more detailed help if necessary and a quit button for termination when an action is complete.

The data for the molecular model for the CGI program is defined by an integer to indicate the number of atoms, a list of atoms with some associated parameters, an integer to indicate the number of bonds and a list of bonds with associated parameters. The parameters necessary for the atom definitions are five integers: X, Y and Z coordinates, radius and colour (the integer to colour mapping is listed in the manual)

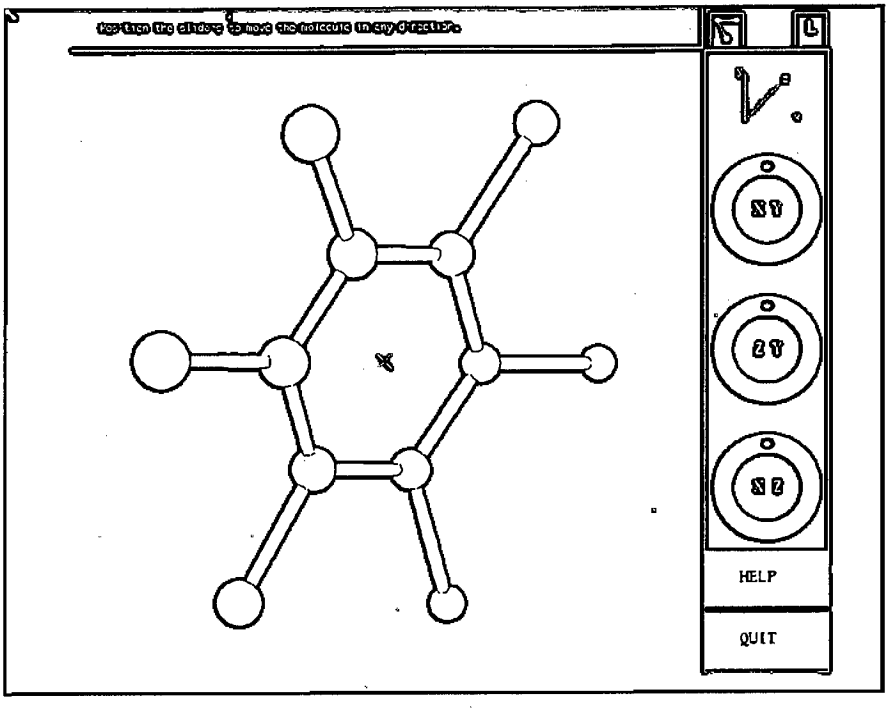


Figure 3: *An example frame of the CGI version, showing the 'turn' option*

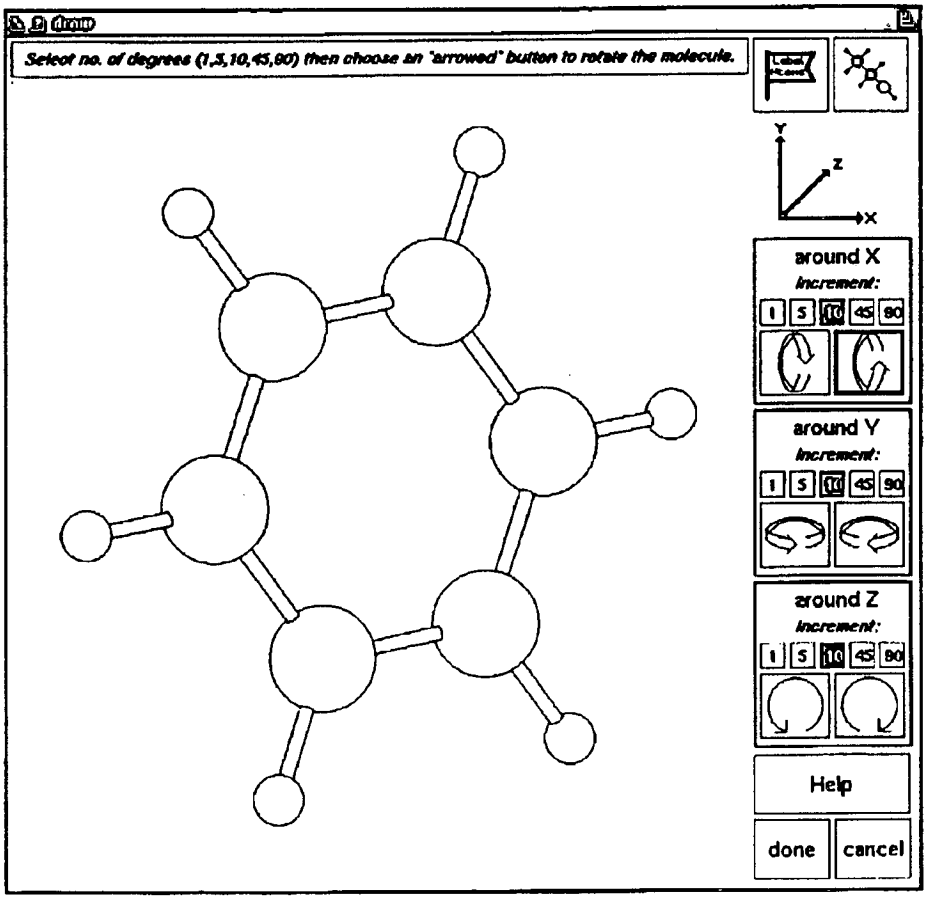


Figure 4: *An example frame of the X version, showing the 'rotate' option*

and a label (words). The bonds are defined by four integers: two to represent the atoms, where the value indicates the position in the atom list of the appropriate atom, a third for bond radius and a fourth for bond colour.

Comparing the interfaces for the two packages, figure 4 shows the frame for the X implementation. The X version has the equivalent instruction, atomic labels and spare molecule button icons (although redesigned to be more meaningful), axes, help button and quit button (renamed 'done' for the same reasons). The major difference is in the menu bar where sliders have been replaced with three mini-frames for rotation about the X, Y and Z axes respectively. Each mini-frame contains a label, a set of 'one of many' radio buttons to determine the extent of rotation (with a default of 10°) and two button icons, for rotation in each direction. Note that the 'around X' anticlockwise button icon is highlighted, showing that the button is currently selected.

The molecular model is much simpler to define for the X program and is based on the crystallographic data produced by the crystal structure solving program ShelX (refs 12, 13), a much used program in the chemistry community. There are three possible types of data specified in the molecular model file, identified by the keywords 'title', 'atom' and 'cell'. 'Title' is optional; 'cell' holds the cell parameters produced by ShelX, is again optional and is used for the scaling of the molecule. The atoms are specified by the keyword 'atom', two words, one for the atom type and one as an identifying label and three numbers for the X, Y, and Z coordinates. The relative size and conventional colour of the atoms are picked up from a database, although these may be changed by the user. The bonds are predicted (to a high success) by a routine which compares the interatomic distance with the sum of the radii of the atoms concerned; again these predictions may be overridden by the user.

IMPROVEMENTS TO THE INTERFACE

Most of the improvements made during the porting from CGI to X have made use of the extra facilities provided by Xlib, the X toolkit and the Athena widget set. In the original program the menu entries consisted of a series of filled rectangles placed at explicit positions on the screen; a procedure calculated within which rectangle the coordinates of a mouse button press fell and the appropriate routines were called. This system, although functional, fell short on some HCI aspects. Moving about the menu structure required several button-presses as some actions required traversing the menu structure down to a third level. The CGI program gave no indication of where the user was within the menu structure meaning that (s)he could become easily disoriented, perhaps forgetting to return to the top level before actions could be carried out. In the original system a short-cut menu was provided to alleviate this problem, but, it seemed, defeated the object of the structured menu system.

The X implementation made use of pull-down menus to reduce menu traversal and since the user is always returned to top level when an operation is completed then (s)he is unlikely to become disoriented. Buttons in the CGI implementation showed no indication of being selected or depressed so the user could never be sure if the button had actually been depressed when the resultant action took some time to be implemented. The Athena Widget set provides command (button) widgets complete with highlighting when the button is selected and then again when depressed, which

can carry out the above implementation more efficiently. The viewport facility was used for zooming in and out of specific parts of the molecule, and then scroll bars used for moving around, an improvement on the CGI version which allowed no movement around the magnified molecule. The above examples are just some of the improvements made to the molecular editor made possible with the facilities provided by the X toolkits and widget set.

FUTURE IMPROVEMENTS

The porting has yet to be fully implemented and there are many plans for new features and improvements. Colour has yet to be incorporated and, it is hoped, at some time in the future, to introduce specular reflection (perhaps with the position of the light sources movable by the user) to give a real three-dimensional effect to the molecule. One important feature planned is the ability to add predefined functional groups to specific atoms or to build molecules from scratch using atoms and these provided predefined functional groups. To determine the most likely positioning in space of the atoms involved it is hoped to provide an interactive minimisation routine taking into account the spatial aspects of the molecule and nearby atoms. At the moment only individual atoms may be added in two dimensions (the third dimension being arbitrarily determined by the program) with minimal user control over the bond length. In addition it is planned to add an enquiry feature whereby users can select bonds and/or atoms and be informed of bond length and angles, and of their feasibility, stereopsis for the stereo viewing of molecules, substantial chemical knowledge and use of chemical heuristics.

PROGRAMMING IN X

When this project was first started the use of InterViews (refs 14–16), was considered. InterViews is an object-oriented, two-dimensional structured graphics library which supports the composition of a graphical user interface from a set of interactive objects. Whilst InterViews is reported to be 'easy to use' there is a high initial learning barrier, especially for a naive user unfamiliar with object-oriented programming. In particular it was felt that the documentation provided was not sufficient for such a naive user.

Although there is still an initial learning barrier with the toolkit and the widget set it was felt that this this was substantially lower than was encountered with InterViews. Once that barrier had been overcome the additional power of the toolkit and the supplied building blocks of the widget set make it easy to achieve a good, pleasing and comfortable interface.

Working with primarily the Xtoolkit and the Athena widget set, and to a lesser extent Xlib, which was just used for the drawing area, was much simpler and documentation and expert help more readily available. Other widget sets such as Motif (ref 17) would have been just as appropriate but the Athena widget set was readily available with good documentation and Motif is not fully compatible with X11r4 yet. The fact that X is event driven was thought to be a potential problem, but since this is managed fully by the toolkit the only event handling done was the fielding of the exposure events for the Xlib drawing area. Since the intention was to port an existing package to X it was

attempted to map the structure of the original on to X. It was soon realised that this exercised a large constraint and that other tools and widgets, such as various types of buttons and menus, viewports, scrollbars, and composite widgets, such as forms and boxes, were more suitable. A Widget tree was drawn which left the user free to concentrate on the functional aspects of the application, without having to waste effort on the details of screen-handling and layout. An exercise which had started out purely as the porting of a molecular graphics package from one windowing system to another rapidly turned into an exercise of redesign.

It was found that expert help was needed to a degree, particularly at the start of the project with the setting up of the framework, although the example programs proved to be very useful. Once widget handling was understood expert help was primarily needed for occasional specific problems and general suggestions for more appropriate yet-to-be-discovered widgets to improve the implementation or interface.

OVERALL SYSTEM

The porting of the original SunCGI system to X was the first stage in the implementation of an intelligent chemical workbench — an integrated tool with a knowledge of chemistry to support the work of a practising chemist. Ultimately, as well as being used on its own to visualise molecular shapes, the program described is to be part of a larger, integrated collection of chemistry-related programs. The environment will consist of integrated programs with a knowledge of atomic parameters, and heuristics for calculating particular attributes such as bonds and bond lengths. Also to be included will be programs for molecular structure determination from X-ray data including refinement of the gross structure thus elucidated (refs 12, 13) and location of atoms (refs 18, 19) difficult to determine by other methods. The molecular drawing tool can be used to view the model at any stage of its synthesis, and therefore is a very real contribution to an applications area which has seen very little technology transfer.

References

[1] *ORTEP*, (Technical Report) ORNL-5138. C. K. Johnson. Oak Ridge National Laboratory, Tennessee, 1976.

[2] *PLUTO - program for Molecular and Crystal Structure Illustrations*, (technical report). W. D. S. Motherwell. University of Cambridge, 1976.

[3] R. Diamond. In **Computational Crystallography**, D. Sayre (editor), pp. 318–325. Oxford University Press, Oxford, 1982.

[4] T. J. O'Donnell and A. J. Olson. **ACM SIGGRAPH Computer Graphics**, Vol. 15, No. 3, pp. 133–142, August 1981.

[5] M. L. Connolly and A. J. Olson. **Computers and Chemistry**, Vol. 9, No. 1, pp. 1–6, 1985.

[6] T. A. Jones. In **Computational Crystallography**, D. Sayre (editor), pp. 318–325. Oxford University Press, Oxford, 1982.

[7] Insight^{TM} - *a trademark of* Biosym Technologies, Inc., 9605 Scranton Road, #101, San Diego, California.

[8] ChemText^{TM} - *a trademark of* Molecular Design Ltd., 10 Armstrong Mall, Southwood Summit Centre, Southwood, Nr Farnborough, Hampshire GU14 ONR.

[9] Sybyl^{TM} - *a trademark of* Tripos Associates Inc., 6548 Claytron Road, St. Louis, Missouri 63117.

[10] SunCGI^{TM} - *a trademark of* Sun Micosystems Plc.

[11] *Interactive Molecular Graphics Package*, Final year project report. A. E. Irvine. Department of Computer Science, University of Edinburgh, 1989.

[12] *Interviews: A C++ Graphical Interface Toolkit*, (unpublished). M. A. Linton, P. R. Calder, and J. M. Vlissides. Stanford University, Stanford, California 94305.

[13] *SHELX76 - a program for Crystal Structure Determination and Refinement*, (technical report). G. M. Sheldrick. University of Cambridge, England, 1976.

[14] *SHELX84 - a program for Crystal Structure Determination*, (technical report). G. M. Sheldrick. University of Göttingen, Federal Republic of Germany, 1984.

[15] *The Design and Implementation of InterViews.* M. A. Linton and P. R. Calder. In **Proceedings of the USENIX C++ Workshop**, pp. 256–267, November 1987.

[16] *Applying Object-Oriented Design to Structured Graphics*, (unpublished). J. Vlissides and M. A. Linton. Stanford University, Stanford, California 94305.

[17] Motif^{TM} - *a trademark of* Open Software Foundation, Eleven Cambridge Centre, Cambridge, Massachusetts 02142.

[18] G. F. Mitchell. **Hydride Location in Transition Metal Clusters**. PhD thesis, Department of Chemistry, University of Edinburgh, September 1988.

[19] *ICON - Program for Extended Huckel Molecular Orbital Calculations*, Program no. 344. J. Howell, A. Rossi, D. Wallace, K. Haraki, and R. Hoffmann. Quantum Chemistry Program Exchange, University of Indiana, 1977.

HYDRIDE LIGAND LOCATION IN COMPLEXES OF THE COPPER TRIAD AND OTHER SYSTEMS

G. F. Mitchell,[†] R. J. Pooley,[‡] A. J. Welch,[†] D. A. Welch,[‡] A. J. Welsh.[‡]

[†]*Department of Chemistry,* [‡]*Department of Computer Science, University of Edinburgh, UK.*

1 INTRODUCTION

X-ray diffraction is generally considered inadequate in locating metal-bound hydride ligands, especially for second or third row transition metals, and particularly in the cases of transition metal cluster compounds. Neutron diffraction provides the ideal solution, but such experiments are far from routine.

Accordingly a number of methods have been developed in which empirical considerations are used to suggest likely sites for hydride ligands given the non-hydrogen molecular skeleton established by an X-ray diffraction experiment. Amongst the best of these are the so-called "potential energy" methods[?,?] in which hydride ligand sites are predicted to be those in which repulsions between hydride ligands and non-hydride ligands are *minimised*. By their very nature, however, these methods fail to correctly predict hydride ligand sites in coordinatively unsaturated species, polyhydride species and interstitial hydride clusters, and for molecules such as these it is preferable to utilise a method which alternatively *maximises* attractions between the metal(s) and hydride ligand(s) present. Methods based on molecular orbital (MO) approaches are therefore appropriate.

The H ligand position in $HFe(CO)_4Mo(CO)_5$ has been successfully optimised by *ab initio* MO calculations[?] and working at this level of calculation on real molecules is clearly the ultimate objective. In the medium term, however, hydride ligand prediction *via* modified extended Hückel MO (MEHMO) calculations may be useful. EHMO calculations, which are very quick and easy to run, are generally good for probing angular interactions between atoms or fragments, but very poor at estimating optimum radial distances as no (electrostatic) repulsion terms are included. Thus in MEHMO calculations a repulsive term WB(r) is added to the EHMO-calculated energy E_{EHMO} to give an overall potential W*(r). Thus for a polyatomic species:

$$W*(r) = \sum W_B(r) + W_{EHMO}(r)$$

$W_B(r)$ is a measure of the interaction of nucleus A (charge Z_A) with "free atom" B, and depends on r, Z_A, and the number and Slater exponent (ζ) of the valence electrons of B, conventionally taken as the more electronegative atom. In metal hydrides H is B, and $W_H(r)$ is calculated as:

$$W_H(r) = e^{-2\zeta r}(\zeta + r^{-1})$$

2 IMPLEMENTATION

A MEHMO program (the Locator) has been written which incorporates evaluation of $W_H(r)$ in the established EHMO program ICON8. Input is the molecular skeleton established by an X-ray diffraction study (with appropriate ligand simplification) and the number of hydride ligands to be optimised. An initial routine uses this information to estimate initial position(s) for the hydride ligand(s), which are then optimised by the numerical analysis method of simplex minimisation. Thus far the program is at the stage of being tested against a number of established metal hydride topologies, but ultimately could be used predictively.

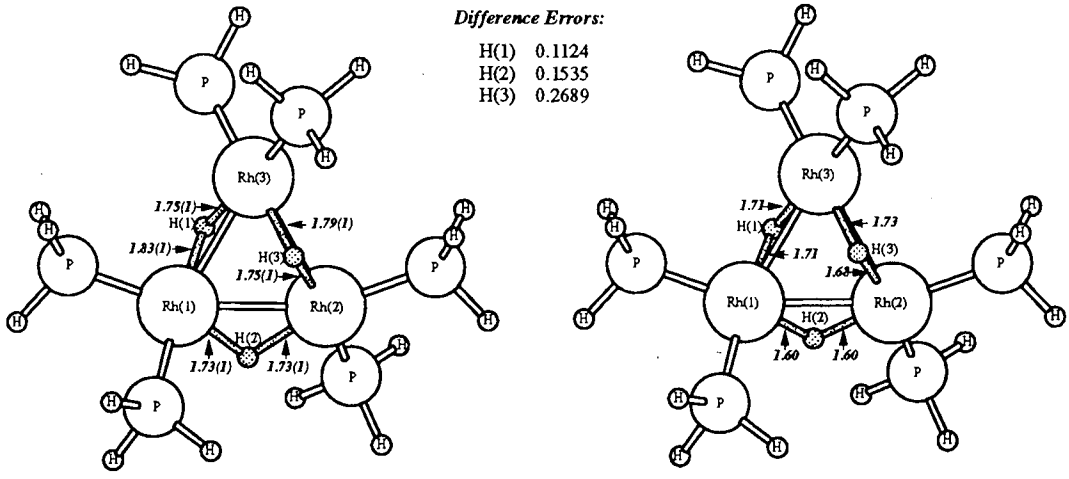
The program has an user-friendly graphical interface and is controlled by buttons and pull-down menus. A ball-and-stick representation of the molecule is displayed at all stages. Whenever a more energetically stable position for the hydride ligand is encountered the display is updated, allowing the user to monitor the progress of the experiment.

3 RESULTS

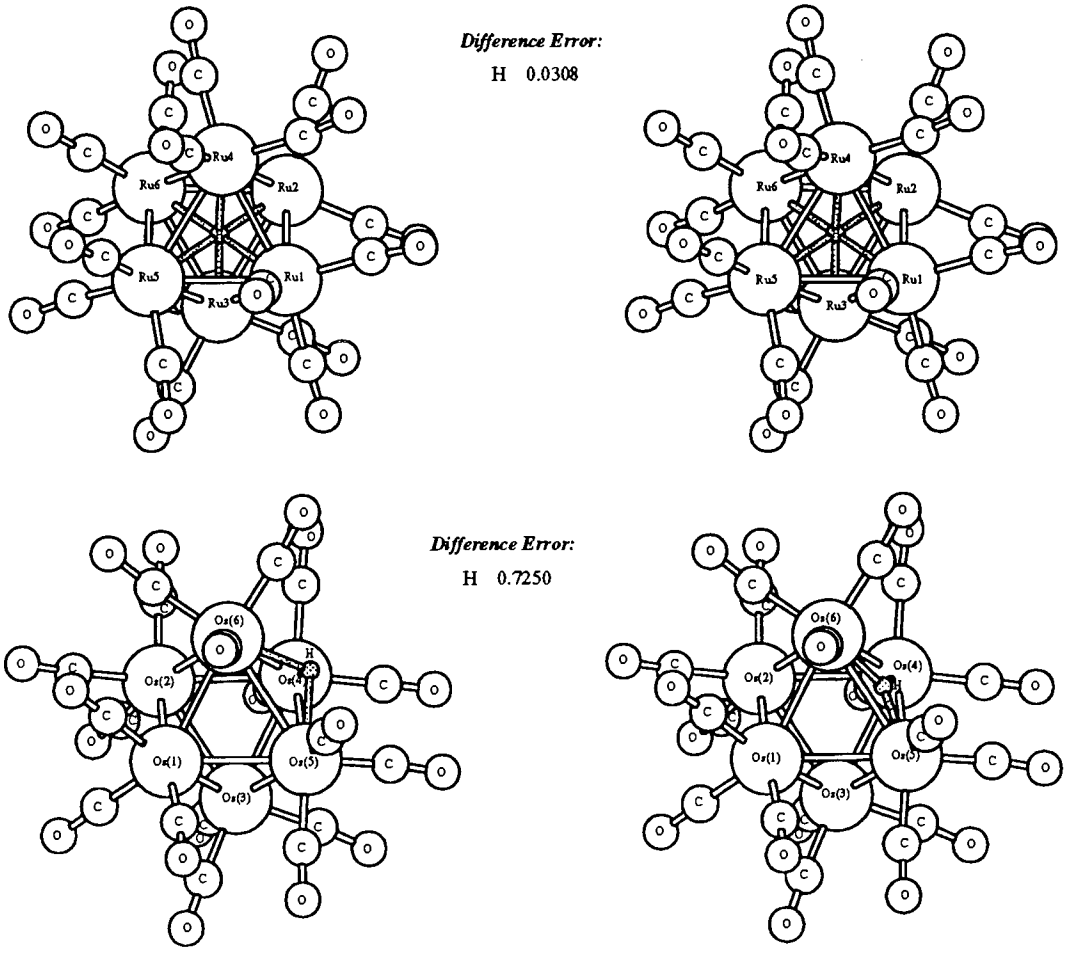
The following figures show the potential of the MEHMO method in finding hydride ligand positions. Each left hand diagram shows the molecular geometry including hydride ligand as established by an alternative method, usually a neutron diffraction study. On the right is plotted the same molecular skeleton together with the hydride ligand(s) as optimised by the program. The figures are annotated with M–H bond lengths in Å, and "difference errors" (the distance from the original neutron or X-ray position to that predicted by the Locator), also in Å.

4 CONCLUSIONS

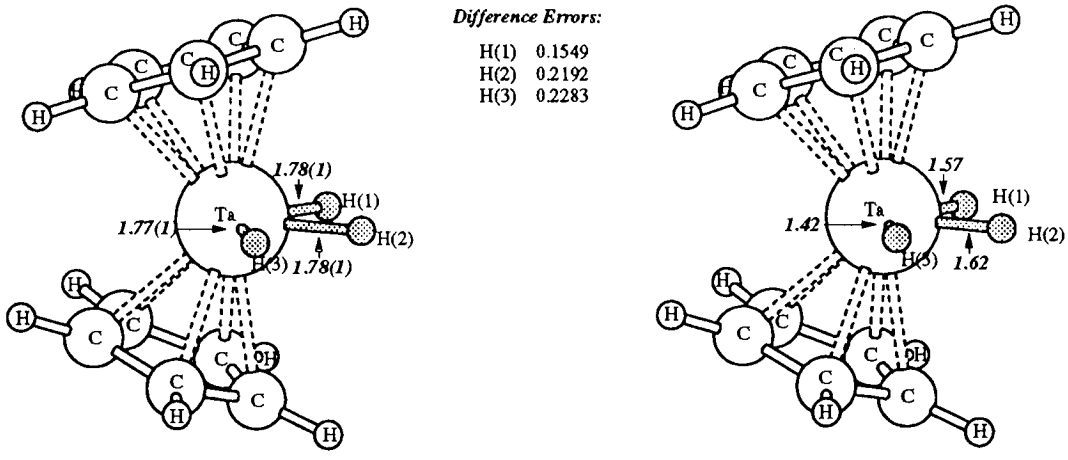
Agreement is generally good, and in particular it is notable that the MEHMO method works well for a number of types of metal hydrides (polyhydride, coordinatively unsaturated, interstitial) for which the potential energy method fails. Furthermore, asymmetry in edge-bridging hydride ligands is often correctly predicted.



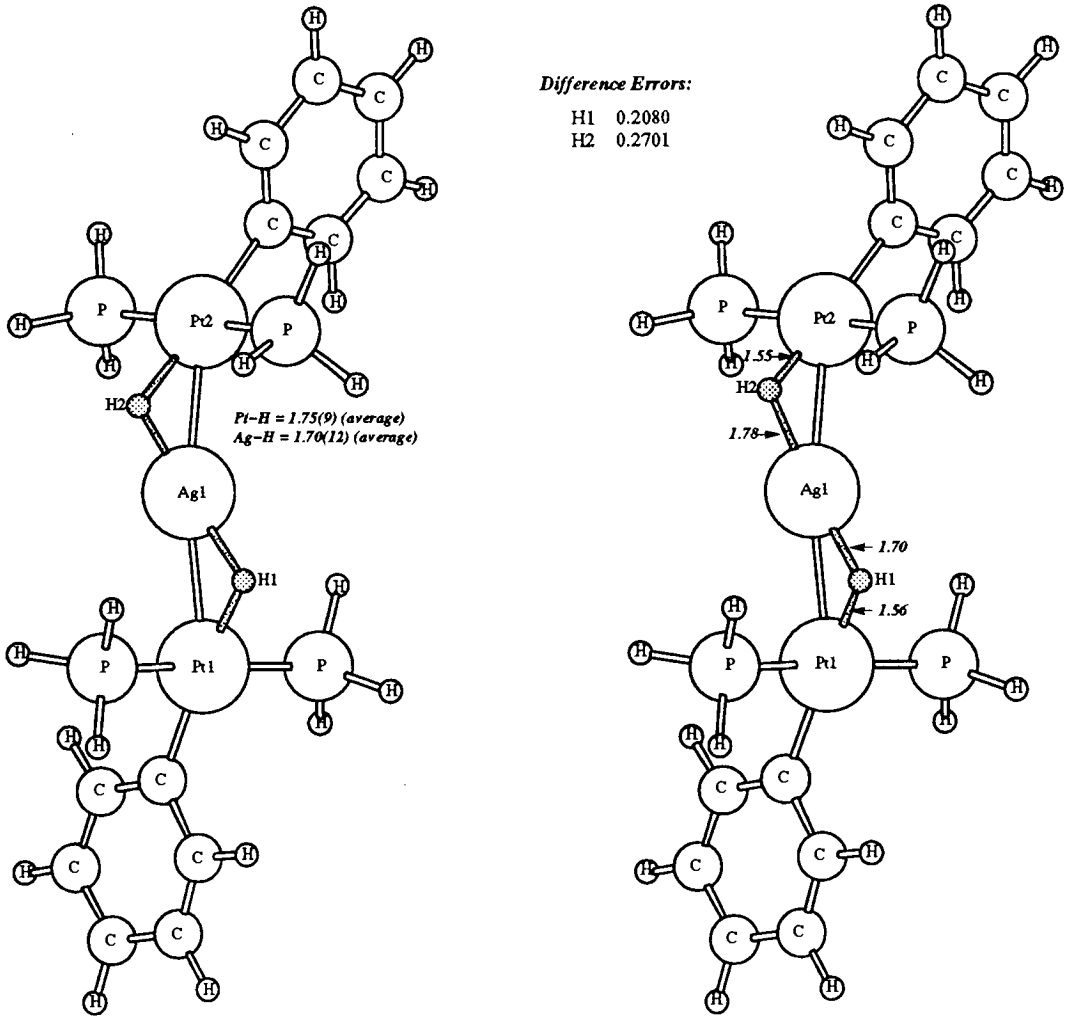
Example 1: *$Rh_3H_3\{P(OMe)_3\}_6$,* [?] *modelled by $Rh_3H_3(PH_3)_6$, is a coordinatively unsaturated metal hydride.*



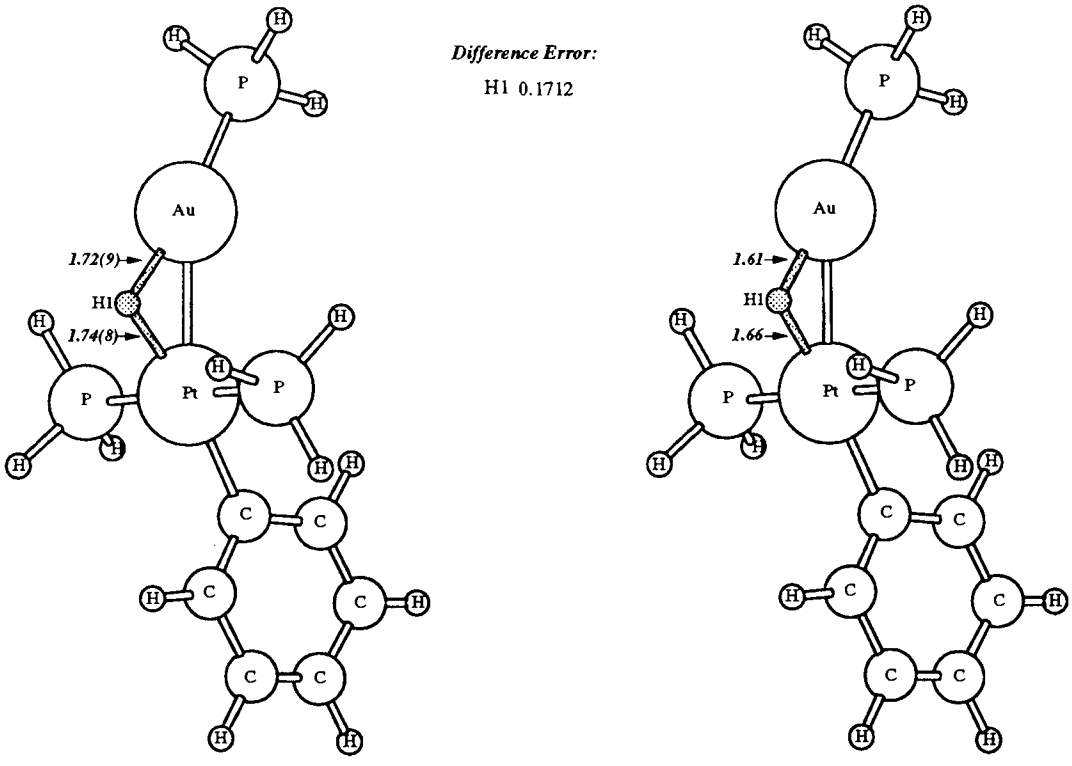
Examples 2 and 3: *$[Ru_6H(CO)_{18}]^-$* [?,?,?] *and $[Os_6H(CO)_{18}]^-$* [?,?] *are chemical analogues, but have topologically different hydride ligands — interstitial and face-capping respectively.*



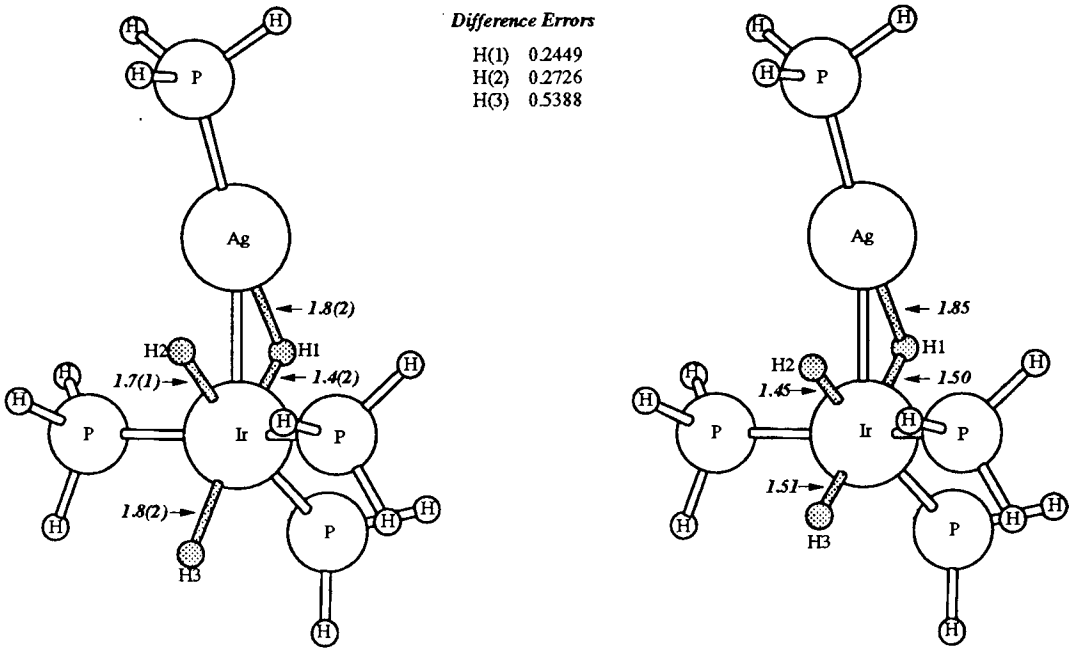
Example 4: *$TaH_3(\eta_5\text{-}C_5H_5)_2$? is a polyhydride.*



Example 5: *$[\{((PH_3)P)_2(C_6H_5)PtH\}_2Ag]^+$,? modelled by $[\{(PH_3)_2(C_6H_5)\text{-}PtH\}_2Ag]^+$, is an example of a compound with a group 11 metal atom involved in two hydride bridges.*



Example 6: *In* $[((PH_3)P)Au(\mu_2\text{-}H)Pt(PEt_3)_2\text{-}(C_6H_5)]^+$,[?] *modelled by* $[((PH_3))AuHPt(PH_3)_2(C_6H_5)]^+$, *both metal atoms are coordinatively unsaturated.*



Example 7: $[((PH_3))AgH_3Ir(PH_3)_3]^+$,[?] *has two terminal and one bridging hydride ligand, all correctly located. The asymmetry in the* $Ag\text{-}\mu H\text{-}Ir$ *system is correctly reproduced.*

5 REFERENCES

[1] G. Ciani, D. Giusto, M. Manassero, and A. Albinati. *J. Chem. Soc., Dalton Trans.*, 1976, 1943.

[2] A. G. Orpen. *J. Chem. Soc., Dalton Trans.*, 1980, 2509.

[3] C. F. Halpin and M. B. Hall. *J. Amer. Chem. Soc.*, 1986, **108**, 1695.

[4] R. K. Brown, J. M. Williams, A. J. Sivak, and E. L. Muetterties. *Inorg. Chem.*, 1980, **19**, 370.

[5] C. R. Eady, P. F. Jackson, B. F. G. Johnson, J. Lewis, M. C. Malatesta, M. McPartlin, and W. J. H. Nelson. *J. Chem. Soc., Dalton Trans.*, 1980, 383.

[6] P. F. Jackson, B. F. G. Johnson, J. Lewis, P. R. Raithby, M. McPartlin, W. J. H. Nelson, K. D. Rouse, J. Allibon, and S. A. Mason. *J. Chem. Soc., Chem. Commun.*, 1980, 295.

[7] C. R. Eady, B. F. G. Johnson, J. Lewis, M. C. Malatesta, P. Machin, and M. McPartlin. *J. Chem. Soc., Chem. Commun.*, 1976, 954.

[8] A. G. Orpen and T. F. Koetzle. *Acta Cryst.*, 1987, **C43**, 2084.

[9] M. McPartlin, C. R. Eady, B. F. G. Johnson, and J. Lewis. *J. Chem. Soc., Chem. Commun.*, 1976, 883.

[10] R. D. Wilson, T. F. Koetzle, D. W. Hart, Å. Kvick, D. L. Tipton, and R. Bau. *J. Amer. Chem. Soc.*, March 1977, **99**, 1775.

[11] A. Albinati, F. Demartin, L. M. Venanzi, and M. K. Wolfer. *Angew. Chem., Int. Ed. Engl.*, 1988, **27**, 563.

[12] A. Albinati, H. Lehner, L. M. Venanzi, and M. Wolfer. *Inorg. Chem.*, 1987, **26**, 3933.

[13] A. Albinati, C. Anklin, P. Janser, H. Lehner, D. Matt, P. S. Pregosin, and L. M. Venanzi. *Inorg. Chem.*, 1989, **28**, 1105.

Bibliography

[1] *Solvent-Accessible Surfaces of Proteins and Nucleic Acids.* M. L. Connolly. **Science**, Vol. 221 (1983), pp. 709–713.

[2] *Analytical Molecular Surface Calculation.* M. L. Connolly. **Journal of Applied Crystallography**, Vol. 16 (1983), pp. 548–558.

[3] *The Storage and Retrieval of Chemical Structures.* F. H. Allen and M. F. Lynch. **Chemistry in Britain**, Vol. 25 (1989), pp. 1101–1104.

[4] *A Graphical Approach to Reaction Retrieval in Organic Chemistry.* J. Boother. **Chemistry in Britain**, Vol. 21 (1985), pp. 68–69.

[5] *Linking Structures and Data.* D. Bawden. **Chemistry in Britain**, Vol. 25 (1989), pp. 1107–1108.

[6] *Computer Aids to Synthesis Planning.* A. P. Johnson. **Chemistry in Britain**, Vol. 21 (1985), pp. 59–67.

[7] *Chemical Structures Online.* P. Rhodes. **Chemistry in Britain**, Vol. 21 (1985), pp. 53–58.

[8] *Microcomputers and Information Systems.* W. G. Town. **Chemistry in Britain**, Vol. 25 (1989), pp. 1118–1120.

[9] *Chemical Information Systems.* P. Willet. **Chemistry in Britain**, Vol. 25 (1989), p. 1099.

[10] *Information Retrieval.* P. Willet. **Computing**, 1989, p. 35.

[11] *Nitrosyl Complexes of Rhenium. Part 2. Crystal and Molecular Structure of Dihydridonitrosyltris(triphenylphosphate)rhenium^{I}-Benzene(2/1).* G. Ciani, D. Giusto, M. Manassero, and A. Albinati. **Journal of the Chemical Society (Dalton Transactions)**, 1976, pp. 1943–1947.

[12] *Indirect Location of Hydride Ligands in Metal Cluster Complexes.* A. G. Orpen. **Journal of the Chemical Society (Dalton Transactions)**, 1980, pp. 2509–2516.

[13] *Hydride Location in Transition Metal Clusters.* G. F. Mitchell. In **Theoretical and Experimental Studies on Charge Deformation Densities and Hydrogen Bridges in Borane Anions and Related Species**, PhD thesis, Department of Chemistry, University of Edinburgh, 1988.

[14] *GRANNY, a Companion to GRAMPS for the Real Time Manipulation of Macromolecular Models.* M. L. Connolly and A. J. Olson. **Computers and Chemistry**, Vol. 9 (1985), pp. 1–6.

[15] *Frodo: A Graphics Fitting Program for Macromolecules.* T. A. Jones. In **Computational Crystallography**, D. Sayre (editor), pp. 303–317. Oxford University Press, 1982.

[16] *GRAMPS — A Graphics Language Interpreter for Real-Time Interactive Three-Dimensional Picture Editing and Animation.* T. J. O'Donnell and A. J. Olson. **ACM SIGGRAPH Computer Graphics**, Vol. 15 (1981), pp. 133–142.

[17] *Survey of Dialogue Systems and Literature on Dialogue Design*, Technical Report CS/HCI/8701/01. J. L. Alty, J. Mullen, and G. Weir. Scottish Human-Computer Interaction Centre, Strathclyde University, 1986.

[18] *Building User Interfaces by Direct Manipulation*, Technical Report. L. Cardelli. DEC SRC, 1987.

[19] *The Human Factors of Graphic Interaction Tasks and Techniques*, Technical Report TR-81-3. J. D. Foley, V. L. Wallace, and P. Chan. Department of Computer Science, University of Kansas, Lawrence, Kansas 66045, 1981.

[20] *Direct Manipulation Interfaces.* E. L. Hutchins, J. D. Hollan, and D. A. Norman. In **User Centered System Design: New Perspectives on Human-Computer Interaction**, D. Norman and S. Draper (editors), chapter 5, pp. 87–124. Lawrence Erlbaum Associates, 1986.

[21] *A Computer Graphics System for Polynomials.* B. Shneiderman. **The Mathematics Teacher**, 1974, pp. 111–113.

[22] *Information Systems and User Resistance: Theory and Practice.* R. Hirschheim and M. Newman. **The Computer Journal**, Vol. 31 (1988), pp. 398–408.

[23] *The Behavioural Side of MIS: Some Aspects of the People Problem.* G. W. Dickson and J. K. Simmons. **Business Horizons**, 1970, pp. 59–71.

[24] *Behavioural Reactions to the Introduction of a Management Information System at the U. S. Post Office: Some Empirical Observations.* G. W. Dickson, J. K. Simmons, and J. C. Anderson. In **Computers and Management**, D. Sanders (editor), pp. 410–422. McGraw-Hill, 1974.

[25] *Hostility in Organization Change.* L. Fried. **Journal of Systems Management**, 1972.

[26] D. Sanders. **Computers and Management**. McGraw-Hill, second edition, 1974.

[27] *The Organisational Validity of Management Information Systems.* M. L. Markus and D. Robey. **Human Relations**, Vol. 36 (1983), pp. 203–226.

[28] *Prototype Systems for Accounting, Information and Control.* M. J. Earl. **Accounting, Organisations and Society**, Vol. 3 (1978), pp. 161–170.

[29] *An Investigation of User-Led System Design: Rational and Political Perspectives.* C. R. Franz and D. Robey. **Communications of the ACM**, Vol. 27 (1984), pp. 1202–1209.

[30] *User Involvement and MIS Success: A Review of Research.* B. Ives and M. H. Olson. **Management Science**, Vol. 30 (1984), pp. 586–603.

[31] *User Reactions and the Management of Information Services.* H. C. Lucas. **Management Informatics**, Vol. 2 (1973), pp. 165–172.

[32] *User Engineering Principles for Interactive Systems.* W. J. Hansen. In **Proceedings of the Fall Joint Computer Conference**, Vol. 39, pp. 523–532, 1971. AFIPS Press.

[33] *Are 'Human Factors' Human?* N. Bjørn-Andersen. **The Computer Journal**, Vol. 31 (1988), pp. 386–390.

[34] *Architectural Design: An Analogy.* K. Hooper. In **User Centered System Design: New Perspectives on Human-Computer Interaction**, D. A. Norman and S. W. Draper (editors), chapter 1, pp. 9–23. Lawrence Erlbaum Associates, 1986.

[35] I. Sommerville. **Software Engineering**. Addison-Wesley, 3rd edition, 1989.

[36] P. Booth. **An Introduction to Human-Computer Interaction**. Lawrence Erlbaum Associates, 1990.

[37] *The Design of 'Idiot-Proof' Interactive Programs.* A. I. Wasserman. In **Proceedings of the National Computer Conference**, Vol. 42, pp. M34–M38, 1973. AFIPS Press.

[38] *Some Experience in Interactive System Development and Application.* B. R. Gaines and P. V. Facey. **Proceedings of the IEEE**, Vol. 63 (1975), pp. 894–911.

[39] *Dialog Engineering.* B. R. Gaines and M. L. G. Shaw. In **Designing for Human-Computer Communication**, M. E. Sime & M. J. Coombs (editor), chapter 1, pp. 23–53. Academic Press, 1983.

[40] *The Design of Interactive Procedures for Man-Machine Communication.* T. C. S. Kennedy. **International Journal of Man-Machine Studies,** Vol. 6 (1974), pp. 309–334.

[41] *Formal Grammer and Human Factors Design of an Interactive Graphics System.* P. Reisner. **IEEE Transactions on Software Engineering,** Vol. SE-7 (1981), pp. 229–240.

[42] *Task-Action Grammars: A Model of the Mental Representation of Task Languages.* S. J. Payne and T. R. G. Green. **Human-Computer Interaction,** Vol. 2 (1986), pp. 93–133.

[43] *Conceptual Consistency in the User Interface: Effects on User Performance.* W. A. Kellog. In **Human-Computer Interaction — INTERACT'87. Amsterdam,** H-J. Bullinger and B. Shackel (editors). Elsevier, 1987.

[44] *The Case Against User Interface Consistency.* J. Grudin. **Communications of the ACM,** Vol. 32 (1989), pp. 1164–1173.

[45] *The Machine Inside the Machine: Users' Models of Pocket Calculators.* R. M. Young. **International Journal of Man-Machine Studies,** Vol. 15 (1981), pp. 51–85.

[46] B. Shneiderman. **Designing the User Interface: Strategies for Effective Human-Computer Interaction.** Addison-Wesley, 1987.

[47] *Dialogue Determination.* H. Thimbleby. **International Journal of Man-Machine Studies,** Vol. 13 (1980), pp. 295–304.

[48] O. Jones. **Introduction to the X Window System.** Prentice-Hall, 1989.

[49] T. O'Reilly (editor). **The X Window System in a Nutshell,** Vol. 5 of *The X Window System Series.* O'Reilly and Associates, 1990.

[50] *The X Window System.* D. Pountain. **BYTE,** 1989, pp. 353–361.

[51] V. Quercia and T. O'Reilly. **X Window System User's Guide for X11 R3 and R4,** Vol. 3 of *The X Window System Series.* O'Reilly and Associates, 3rd edition, 1990.

[52] *Guide to the X Window System Version 11 Release 4,* Technical Report CSR-5-90. G. D. M. Ross. Department of Computer Science, University of Edinburgh, West Mains Road, Edinburgh, 1990.

[53] *The X Window System.* R. W. Scheifler. **ACM Transactions on Graphics,** Vol. 5 (1986), pp. 79–109.

[54] J. Gettys, R. W. Scheifler, and R. Newman. **Xlib — C Language X Interface. MIT X Consortium Standard. X Version 11, Release 4.** Massachussetts Institute of Technology, Cambridge, Massachussetts, 1989.

[55] R. W. Scheifler and J. Gettys. **X Window System: The Complete Reference to Xlib, X Protocol.** Digital Press, 2nd edition, 1990.

[56] P. J. Asente and R. R. Swick. **X Window System Toolkit: the Complete Programmer's Guide and Specification.** Digital Press, 1990.

[57] J. McCormack, P. Asente, and R. R. Swick. **X Toolkit Inrinsics — C Language X Interface. MIT X Consortium Standard. X Version 11, Release 4.** Massachussetts Institute of Technology, Cambridge, Massachussetts, 1988.

[58] A. Nye and T. O'Reilly. **X Toolkit Intrinsics Programming Manual Second Edition for X11 Release 4,** Vol. 4 of *The X Window System Series.* O'Reilly and Associates, 2nd edition, 1990.

[59] T. O'Reilly (editor). **X Toolkit Intrinsics Reference Manual Second Edition for X11 Release 4,** Vol. 5 of *The X Window System Series.* O'Reilly and Associates, 2nd edition, 1990.

[60] C. D. Peterson. **Athena Widget Set — C Language X Interface. X Window System. X Version 11, Release 4.** Massachussetts Institute of Technology, Cambridge, Massachussetts, 1989.

[61] *Visualisation in Scientific Computing (ViSC): Definitions, Domain and Recommendations.* B. H. McCormick, T. A. DeFanti, and M. D. Brown (eds). **ACM SIGGRAPH Computer Graphics,** Vol. 21 (1987), pp. 1–E9.

[62] *Visualisation: Expanding Scientific and Engineering Research Opportunities.* T. A. DeFanti, M. D. Brown, and B. H. McCormick. **Computer (IEEE Computer Science),** 1989, pp. 12–25.

[63] *Visualisation in Scientific Computing.* G. M. Neilson. **Computer (IEEE Computer Science),** 1989, pp. 10–11.

[64] *Scientific Visualisation at Research Laboratories.* L. J. Rosenblum. **Computer (IEEE Computer Science),** 1989, pp. 68–101.

[65] R. W. Hamming. **Numerical Methods for Scientists and Engineers.** McGraw-Hill, 1962.

[66] *Interactive Visualisation of 3D Medical Data.* H. Fuchs, M. Levoy, and S. M. Pizer. **Computer (IEEE Computer Science),** Vol. 22 (1989), pp. 46–51.

[67] *Influence of Visual Technology on the Evolution of Language Environments.* A. L. Ambler and M. M. Burnett. **Computer (IEEE Computer Science)**, Vol. 22 (1989), pp. 9–22.

[68] *Representation and Display of Vector Field Topology in Fluid Flow Data Sets.* J. Helman and L. Hesselink. **Computer (IEEE Computer Science)**, Vol. 22 (1989), pp. 27–36.

[69] *Visualising Large Data Sets in the Earth Sciences.* W. Hibbard and D. Santek. **Computer (IEEE Computer Science)**, Vol. 22 (1989), pp. 53–57.

[70] *The Role of Visualisation in the Simulation of Quantum Electronic Transport in Semiconductors.* N. C. Klucksdahl, A. M. Kriman, and D. K. Ferry. **Computer (IEEE Computer Science)**, Vol. 22 (1989), pp. 60–65.

[71] *Graphical Configuration Programming.* J. Kramer, J. Magee, and K. Ng. **Computer (IEEE Computer Science)**, Vol. 22 (1989), pp. 53–65.

[72] *Visualising Performance Debugging.* T. Lehr, Z. Segall, D. F. Vrsalovic, E. Caplan, A. L. Chung, and C. E. Fineman. **Computer (IEEE Computer Science)**, Vol. 22 (1989), pp. 38–51.

[73] *Acquisition and Representation of 2D and 3D Data from Turbulent Flows and Flames.* M. B. Long, K. Lyons, and J. K. Lam. **Computer (IEEE Computer Science)**, Vol. 22 (1989), pp. 39–45.

[74] *A Functional View of Proteins.* A. J. Olson and D. S. Goodsell. **IEEE Computer Graphics and Applications**, 1991, pp. 15–17.

[75] *A Declarative Approach to Visualising Concurrent Computations.* G. C. Roman and K. C. Cox. **Computer (IEEE Computer Science)**, Vol. 22 (1989), pp. 25–36.

[76] *PROTEUS: Graphics Software for Proteins.* H. Dayringer, A. Tramantano, and R. Fletterick. In **Computer Graphics and Molecular Modelling**, R. Fletterick and M. Zoller (editors), pp. 3–8. Cold Spring Harbour Laboratory, 1985.

[77] *Bilder: an Attractive Graphics Program for Biopolymers.* R. Diamond. In **Computational Crystallography**, D. Sayre (editor), pp. 318–325. Oxford University Press, 1982.

[78] *HYDRA: Current and Future Developments.* R. E. Hubbard. In **Computer Graphics and Molecular Modelling**, R. Fletterick and M. Zoller (editors), pp. 9–11. Cold Spring Harbour Laboratory, 1985.

[79] *Parallel Algorithms for Molecular Graphics*, Final Year Project Report. R. Sayle. Imperial College of Science, Technology & Medicine, 1990.

[80] *RasMol: A Program for Fast, Realistic Rendering of Molecular Structures with Shadows.* R. Sayle and A. Bissell. In **Proceedings of the 10th Eurographics UK Conference**, pp. 125–138. University of Edinburgh, 1992.

[81] *Computer Generated Molecular Modeling.* S. D. Stewart. **Simulation**, Vol. 47 (1986), pp. 18–23.

[82] *Molecular Graphics for the Medicinal Chemist.* J. G. Vinter. **Chemistry in Britain**, Vol. 21 (1985), pp. 32–38.

[83] *Computer Graphics as an Aid to Drug Design.* C. H. Hassall. **Chemistry in Britain**, Vol. 21 (1985), pp. 39–46.

[84] *Computer Graphics in Solid-State and Surface Chemistry.* S. Ramdas and J. M. Thomas. **Chemistry in Britain**, Vol. 21 (1985), pp. 49–52.

[85] *Macromolecular Graphics.* A. J. Olson and D. S. Goodsell. In **Applications of Computer Graphics to Molecular Modeling (Course Notes 44)**, pp. 23–44. SIGGRAPH, 1992.

[86] *Interactive Molecular Graphics Package*, Final Year Project Report. A. E. Irvine. Department of Computer Science, University of Edinburgh, 1989.

[87] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes. **Computer Graphics: Principles and Practice**. Addison-Wesley, 2nd edition, 1990.

[88] *ORTEP*, Technical Report ORNL-5138. C. K. Johnson. Oak Ridge National Laboratory, Tennessee, 1976.

[89] *Novel Reactions of Metal Carbonyl Cluster Compounds.* R. D. Adams and I. T. Horváth. **Progress in Inorganic Chemistry**, Vol. 33 (1985), pp. 127–181.

[90] *Homo- and Heteronuclear Cluster Compounds of Gold.* K. P. Hall and D. M. P. Mingos. **Progress in Inorganic Chemistry**, Vol. 32 (1984), pp. 237–323.

[91] *The Hydrido-Transition Metal Cluster Complexes.* A. P. Humphries and H. D. Kaesz. **Progress in Inorganic Chemistry**, Vol. 25 (1977), pp. 145–222.

[92] F. A. Cotton and G. Wilkinson. **Advanced Inorganic Chemistry**. John Wiley and Sons Ltd., 5th edition, 1988.

[93] *Making Molecules by Numbers.* D. Rouvray. **New Scientist**, 1991, pp. 22–26.

[94] *Structures of Transition-Metal Hydride Complexes.* R. Bau, R. G. Teller, S. W. Kirtley, and T. F. Koetzle. **Accounts of Chemical Research**, Vol. 12 (1979), pp. 176–183.

[95] *The Structure of Metal Cluster Compounds.* P. R. Raithby. In **Transition Metal Clusters**, B. F. G. Johnson (editor), chapter 2, pp. 5–192. John Wiley and Sons Ltd., 1980.

[96] J. Warmold. **Diffraction Methods**. Oxford University Press, 1973.

[97] *Quantum Mechanical Prediction of Hydride Locations in Transition-Metal Systems.* C. F. Halpin and M. B. Hall. **Journal of the American Chemical Society**, Vol. 108 (1986), pp. 1695–1696.

[98] *A High-Precision Neutron Diffraction Investigation of the Prototype Biscyclopentadienyl Transition Metal Hydride* $(\eta^5\text{-}C_5H_5)_2MoH_2$ *and Resolution of the Structure in Terms of Modern Bonding Theory.* A. J. Shultz, K. L. Stearly, J. M. Williams, R. Mink, and G. D. Stucky. **Inorganic Chemistry**, Vol. 16 (1977), pp. 3303–3306.

[99] *Neutron Diffraction Studies of Tetrahedral Cluster Transition Metal Hydride Complexes:* $HFeCo_3(CO)_9(P(OCH_3)_3)_3$ *and* $H_3Ni_4(C_5H_5)_4$. T. F. Koetzle, R. K. McMullan, R. Bau, D. W. Hart, R. G. Teller, D. L. Tipton, and R. D. Wilson. In **Transition Metal Hydrides**, R. Bau (editor), chapter 5, pp. 61–72. American Chemical Society, 1978.

[100] *Location of Terminal Hydride Ligands in Transition Metal Hydrides.* J. A. Ibers. In **Transition Metal Hydrides**, R. Bau (editor), chapter 3, pp. 26–35. American Chemical Society, 1978.

[101] *X-Ray and Neutron Diffraction Study of 1,1,1,2,2,2,3,3,3-Nonacarbonyl-1,2;1,-3-di-*μ*-hydrido-*μ_3*-sulphido-triangulo-triosmium.* B. F. G. Johnson, J. Lewis, D. Pippard, P. R. Raithby, G. M. Sheldrick, and K. D. Rouse. **Journal of the Chemical Society (Dalton Transactions)**, 1979, pp. 616–618.

[102] *Structural Studies on Polynuclear Osmium Carbonyl Hydrides. Part 10. Molecular Geometry and Comments on the Location of the* μ_2*-Hydride Ligand in* $(\mu\text{-}H)Os_3W(CO)_{12}(\eta^5\text{-}C_5H_5)$. M. R. Churchill and F. J. Hollander. **Inorganic Chemistry**, Vol. 18 (1979), pp. 843–848.

[103] *Structural Studies on Polynuclear Osmium Carbonyl Hydrides. Part 3. Crystal Structure and Molecular Geometry of* $H_2Os_3(CO)_{10}(PPh_3)$*, Including Direct Location of Both Terminal and* μ_2*-Bridging Hydride Ligands.* M. R. Churchill and B. G. DeBoer. **Inorganic Chemistry**, Vol. 16 (1977), pp. 2397–2403.

[104] *Improved Syntheses of the Hexanuclear Clusters of* $[Ru_6(CO)_{18}]^{2-}$, $[HRu_6(CO)_{18}]^-$, *and* $H_2Ru_6(CO)_{18}$. *The X-ray Analysis of* $[HRu_6(CO)_{18}]^-$, *a Polynuclear Carbonyl Containing an Interstitial Hydrogen Ligand.* C. R. Eady, P. F. Jackson, B. F. G. Johnson, J. Lewis, M. C. Malatesta, M. McPartlin, and W. J. H. Nelson. **Journal of the Chemical Society (Dalton Transactions)**, 1980, pp. 383–392.

[105] *Mixed-Metal Gold Clusters. Synthesis and X-Ray Structural Characterization of* $[Au_5Re(H)_4(PPh_3)_7](PF_6)_2$ *and* $[Au_3Rh(H)(CO)(PPh_3)_5](PF_6)$. P. D. Boyle, B. J. Johnson, A. Buehler, and L. H. Pignolet. **Inorganic Chemistry**, Vol. 25 (1986), pp. 5–7.

[106] *Coordinately Unsaturated Metal Clusters. A Neutron Diffraction Study of the Crystal and Molecular Structure of the Trimeric Rhodium Complex* $\{HRh[P(OCH_3)_3]_2\}_3$. R. K. Brown, J. M. Williams, A. J. Sivak, and E. L. Muetterties. **Inorganic Chemistry**, Vol. 19 (1980), pp. 370–374.

[107] *Synthesis and X-Ray Analysis of the tetrahydrido-dianion* $[H_4Os_{10}(CO)_{24}]^{2-}$, *the First Non-carbido Decaosmium Cluster.* D. Braga, J. Lewis, B. F. G. Johnson, M. McPartlin, W. J. H. Nelson, and M. D. Vargas. **Journal of the Chemical Society (Chemical Communications)**, 1983, pp. 241–243.

[108] *Direct Location of the Interstitial Hydride Ligand in* $[HRu_6(CO)_{18}]^-$ *by both X-Ray and Neutron Analyses of* $[Ph_4As][HRu_6(CO)_{18}]$. P. F. Jackson, B. F. G. Johnson, J. Lewis, P. R. Raithby, M. McPartlin, W. J. H. Nelson, K. D. Rouse, J. Allibon, and S. A. Mason. **Journal of the Chemical Society (Chemical Communications)**, 1980, pp. 295–297.

[109] *Crystal and Molecular Structure of Tristriphenylphosphine Rhodium Carbonyl Hydride.* S. J. La Placa and J. A. Ibers. **Acta Crystallographica**, Vol. 18 (1965), pp. 511–519.

[110] N. Pavel, C. Quagliata, and N. Scarcelli. **Zeitschrift fuer Kristallographie**, Vol. 144 (1976), p. 64.

[111] *Could One Have Five? Synthesis and Structure of* $Re_2H_4(PMe_2Ph)_4[P(OCH_2)_3CEt]_2$ *and Its Conjugate Acid.* M. A. Green, J. C. Huffman, and K. G. Caulton. **Journal of the American Chemical Society**, Vol. 104 (1982), pp. 2319–2320.

[112] *Molecular Metal Clusters.* E. L. Muetterties. **Science**, Vol. 196 (1977), pp. 839–848.

[113] *Boron Hydrides: LCAO–MO and Resonance Studies.* R. Hoffmann and W. N. Lipscomb. **Journal of Chemical Physics**, Vol. 37 (1962), pp. 2872–2883.

[114] *An Extended Hückel Theory. I. Hydrocarbons.* R. Hoffmann. **Journal of Chemical Physics**, Vol. 39 (1963), pp. 1397–1412.

[115] *Theoretical Organometallic Chemistry.* R. Hoffmann. **Science**, Vol. 211 (1981), pp. 995–1002.

[116] *Description of Diatomic Molecules using One Electron Configuration Energies with Two-Body Interactions.* A. B. Anderson and R. Hoffmann. **Journal of Chemical Physics**, Vol. 60 (1974), pp. 4271–4273.

[117] *Perpendicular and Parallel Acetylene Complexes.* D. M. Hoffman, R. Hoffmann, and C. R. Fisel. **Journal of the American Chemical Society**, Vol. 104 (1982), pp. 3858–3875.

[118] *Transition-Metal Complexed Olefins: How their Reactivity toward a Nucleophile Relates to their Electronic Structure.* O. Eisenstein and R. Hoffmann. **Journal of the American Chemical Society**, Vol. 103 (1981), pp. 4308–4320.

[119] *Deformations from Octahedral Geometry in d^4 Transition-Metal Complexes.* P. Kubáček and R. Hoffmann. **Journal of the American Chemical Society**, Vol. 103 (1981), pp. 4320–4332.

[120] *Metalloporphyrins with Unusual Geometries. 1. Mono-, Di-, Triatom-Bridged Porphyrin Dimers.* K. Tatsumi and R. Hoffmann. **Journal of the American Chemical Society**, Vol. 103 (1981), pp. 3328–3341.

[121] *ICON — Program for Extended Hückel Molecular Orbital Calculations,* Program no. 344. J. Howell, A. Rossi, D. Wallace, K. Haraki, and R. Hoffmann. Quantum Chemistry Program Exchange, University of Indiana, 1977.

[122] F. P. Preparata and M. I. Shamos. **Computational Geometry: An Introduction**. Springer-Verlag, 2nd edition, 1985.

[123] *Crystal and Molecular Structure of an Iron Dihydrogen Metallocarborane, [2,3-$(CH_3)_2C_2B_4H_4]_2FeH_2$. Steric Aspects of the Oxidative Fusion Process.* J. R. Pipal and R. N. Grimes. **Inorganic Chemistry**, Vol. 18 (1979), pp. 263–267.

[124] *Some Reactions of Dodecacarbonyltetrahydridotetraosmium with Olefins; The Molecular Structure of 1,1,1,2,2,3,3,3,4,4,4-Undecacarbonyl-1,2-μ-(1′-σ,1′–2′-η-cyclohexyl)-tri-μ-hydrido-tetrahedro-tetraosmium.* S. Bhaduri, B. F. G. Johnson, J. W. Kelland, J. Lewis, P. R. Raithby, S. Rehani, G. M. Sheldrick, K. Wong, and M. McPartlin. **Journal of the Chemical Society (Dalton Transactions)**, 1979, pp. 562–568.

[125] *Synthesis of Cluster Compounds via Nucleophilic Metal Complexes; X-Ray Crystal Structures of $[Os_3Pt(\mu\text{-}H)_2(CO)_{10}\{(C_6H_{11})_3P\}]$ and $[Os_2Pt_2(\mu\text{-}H)_2(CO)_8(Ph_3P)_2]$.* L. J. Farrugia, J. A. K. Howard, P. Mitrprachachon, J. L. Spencer, F. G. A. Stone, and P. Woodward. **Journal of the Chemical Society (Chemical Communications)**, 1978, pp. 260–262.

[126] *The Crystal and Molecular Structure of Trans-dichloro-mer-tris(dimethylphenyl-phosphine)hydridoiridium(III) by X-Ray and Neutron Diffraction.* G. B. Robertson and P. A. Tucker. **Australian Journal of Chemistry**, Vol. 40 (1987), pp. 1043–1052.

[127] *Concept of the H(δ+)=H(δ-) Interaction — a Low-Temperature Neutron Diffraction Study of cis-$[IrH(OH)(PMe_3)_4]PF_6$.* R. C. Stevens, R. Bau, D. Milstein,

O. Blum, and T. F. Koetzle. **Journal of the Chemical Society (Dalton Transactions)**, 1990, pp. 1429–1432.

[128] *Dihydridobis(triethylsilyl)pentamethylcyclopentadienyliridium(V) — Neutron Diffraction and Photoelectron Spectroscopic Studies.* J. S. Ricci, Jr., T. F. Koetzle, M.-J. Fernandez, P. M. Maitlis, and J. C. Green. **Journal of Organometallic Chemistry**, Vol. 299 (1986), pp. 383–389.

[129] *Structural and Spectroscopic Characterization of Iridium Trihydride Complexes — Evidence for Proton-Proton Exchange Coupling.* D. M. Heinekey, J. M. Millar, T. F. Koetzle, N. G. Payne, and K. W. Zilm. **Journal of the American Chemical Society**, Vol. 112 (1990), pp. 909–919.

[130] *Nature of the Metal–Hydrogen Bond in Transition Metal–Hydrogen Complexes: Neutron and X-Ray Diffraction Studies of β-Pentacarbonylmanganese Hydride.* S. J. La Placa, W. C. Hamilton, J. A. Ibers, and A. Davison. **Inorganic Chemistry**, Vol. 8 (1969), pp. 1928–1935.

[131] *Neutron and X-Ray Diffraction Studies on Tris(dimethylphenylphosphine)osmiumtetrahydride.* D. W. Hart, R. Bau, and T. F. Koetzle. **Journal of the American Chemical Society**, Vol. 99 (1977), pp. 7557–7564.

[132] *Crystal and Molecular Structure of Bis(diisopropylphenylphosphine)hexahydridoosmium, $[OsH_6(PC_{12}H_{19})_2]$ — Single-Crystal Neutron Diffraction Study At 20-K.* J. A. K. Howard, O. Johnson, T. F. Koetzle, and J. L. Spencer. **Inorganic Chemistry**, Vol. 26 (1987), pp. 2930–2933.

[133] *Pentahydridorhenium — Crystal and Molecular Structure of $ReH_5(PMePh_2)_3$.* T. J. Emge, T. F. Koetzle, J. W. Bruno, and K. G. Caulton. **Inorganic Chemistry**, Vol. 23 (1984), pp. 4012–4017.

[134] *Low-Temperature Neutron Diffraction Study of $[ReH_5(PPhIpr_2)_2(SiHPh_2)_2]$ and Low-Temperature X-Ray Diffraction Study of $[ReH_5(PCyp_3)_2(SiH_2Ph)_2]$.* J. A. K. Howard, P. A. Keller, T. Vogt, A. L. Taylor, N. D. Dix, and J. L. Spencer. **Acta Crystallographica**, Vol. B48 (1992), pp. 438–444.

[135] *Classical Versus Non-Classical Polyhydride Complexes — Single-Crystal Neutron Diffraction Study of 1,2-Bis(diphenylphosphino)ethaneheptahydridorhenium(VII) $[ReH_7(Dppe)]$ using a Position-Sensitive Detector.* J. A. K. Howard, S. A. Mason, O. Johnson, I. C. Diamond, S. Crennell, P. A. Keller, and J. L. Spencer. **Journal of the Chemical Society (Chemical Communications)**, 1988, pp. 1502–1503.

[136] *Location of the Elusive Hydride Ligand in $HRh[P(C_6H_5)_3]_4$ via a Neutron Diffraction Analysis.* M. R. McLean, R. C. Stevens, R. Bau, and T. F. Koetzle. **Inorganica Chimica Acta**, Vol. 166 (1989), pp. 173–175.

[137] *Synthesis, X-Ray, and Low-Temperature Neutron Diffraction Study of a Rhodium(V) Complex — Dihydridobis(triethylsilyl)-pentamethylcyclopentadienyl-*

rhodium. M.-J. Fernandez, P. M. Bailey, P. O. Bentz, J. S. Ricci, T. F. Koetzle, and P. M. Maitlis. **Journal of the American Chemical Society**, Vol. 106 (1984), pp. 5458–5463.

[138] *X-ray and Neutron Diffraction Studies on $H_3Nb(C_5H_5)_2$ and $H_3Ta(C_5H_5)_2$.* R. D. Wilson, T. F. Koetzle, D. W. Hart, Å. Kvick, D. L. Tipton, and R. Bau. **Journal of the American Chemical Society**, Vol. 99 (1977), pp. 1775–1781.

[139] *Synthesis, Characterization, Low-Temperature (50-K) Neutron Diffraction Structure, and Reactivity of Chlorodihydridotetrakis(trimethylphosphine)tantalum(III).* M. L. Luetkens, Jr., M. D. Hopkins, A. J. Schultz, J. M. Williams, C. K. Fair, F. K. Ross, J. C. Huffman, and A. P. Sattelberger. **Inorganic Chemistry**, Vol. 26 (1987), pp. 2430–2436.

[140] *Formation of Carboxylate Complexes from the Reactions of CO_2 with Ethylene Complexes of Molybdenum and Tungsten — X-Ray and Neutron Diffraction Studies.* R. Alvarez, E. Carmona, A. Galindo, E. Gutierrez, J. M. Marin, A. Monge, M. L. Poveda, C. Ruiz, and J. M. Savariault. **Organometallics**, Vol. 8 (1989), pp. 2430–2439.

[141] *Terminal Zinc-Hydrogen Bonding. X-ray and Neutron Diffraction Studies of the (2-Dimethylamino-N-methylethylamido)hydridozinc Dimer.* N. A. Bell, P. T. Moseley, H. M. M. Shearer, and C. B. Spencer. **Acta Crystallographica**, Vol. B36 (1980), p. 2950.

[142] *Trans-hydridocarbonyltris(triphenylphosphine)cobalt(I), $C_{55}H_{46}COOP_3$.* D. C. Moody and R. R. Ryan. **Crystal Structure Communications**, Vol. 10 (1981), pp. 129–132.

[143] *Synthesis and Crystal Structure of an Iron Nitrosyl Carbonyl Hydride.* M. Cygler, F. R. Ahmed, A. Forgues, and J. L. A. Roustan. **Inorganic Chemistry**, Vol. 22 (1983), pp. 1026–1030.

[144] *An Attractive Cis-Effect of Hydride on Neighbor Ligands — Experimental and Theoretical-Studies on the Structure and Intramolecular Rearrangements of $Fe(H)_2(\eta_2\text{-}H_2)(PEtPh_2)_3$.* L. S. Van Der Sluys, J. Eckert, O. Eisenstein, J. H. Hall, J. C. Huffman, S. A. Jackson, T. F. Koetzle, G. J. Kubas, P. J. Vergamini, and K. G. Caulton. **Journal of the American Chemical Society**, Vol. 112 (1990), pp. 4831–4841.

[145] *Insertion of Iridium into the C–H Bonds of Alkenes — the π-Complex cannot be an Intermediate.* P. O. Stoutland and R. G. Bergman. **Journal of the American Chemical Society**, Vol. 107 (1985), pp. 4581–4582.

[146] *Oxidative Addition of Unactivated Epoxides to Iridium(I) Complexes — Formation of Stable Cis-hydridoformylmethyl and Hydridacylmethyl Complexes.* D. Milstein and J. C. Calabrese. **Journal of the American Chemical Society**, Vol. 104 (1982), pp. 3773–3774.

[147] *Rational Design in Homogeneous Catalysis — $Ir^{(I)}$-Catalyzed Addition of Aniline to Norbornylene via N-H Activation.* A. L. Casalnuovo, J. C. Calabrese, and D. Milstein. **Journal of the American Chemical Society**, Vol. 110 (1988), pp. 6738–6744.

[148] *Formation, Structures, and Reactivity of Cis-hydroxy Hydrides, Cis-methoxy Hydrides, and Cis-mercaptoiridium Hydrides — Oxidative Addition of Water to $Ir^{(I)}$.* D. Milstein, J. C. Calabrese, and I. D. Williams. **Journal of the American Chemical Society**, Vol. 108 (1986), pp. 6387–6389.

[149] *Boryliridium and Boraethyliridium Complexes fac-[$IrH_2(PMe_3)_3(BRR')$] and fac-[$IrH(PMe_3)_3(\eta^2$-$CH_2BHRR')$].* R. T. Baker, D. W. Ovenall, J. C. Calabrese, S. A. Westcott, N. J. Taylor, I. D. Williams, and T. B. Marder. **Journal of the American Chemical Society**, Vol. 112 (1990), pp. 9399–9400.

[150] B. D. Alexander, B. J. Johnson, S. M. Johnson, P. D. Boyle, N. C. Kann, A. M. Mueting, and L. H. Pignolet. **Inorganic Chemistry**, Vol. 26 (1987), p. 3506.

[151] *Unsaturated Bis(phosphido)-bridged Heterobimetallic Polyhydrides via Dihydrogen Activation.* R. T. Baker, T. E. Glassman, D. W. Ovenall, and J. C. Calabrese. **Israel Journal of Chemistry**, Vol. 31 (1991), pp. 33–53.

[152] *Structural Determination of [$(C_5H_5)_2Mo(H)CO$][$(C_5H_5)Mo(CO)_3$].* M. A. Adams, K. Folting, J. C. Huffman, and K. G. Caulton. **Inorganic Chemistry**, Vol. 18 (1979), pp. 3020–3023.

[153] *Reactions of Co-ordinated Ligands. Part 44. The Reactivity of the Carbyne Complex [$Mo(\equiv CCH_2Bu^t)\{P(OMe)_3\}_2(\eta$-$C_5H_5)$] towards Protonic Acids, Diphenylphosphine, and the π-Acid Tetracyanoethylene; Formation and Crystal Structures of [$MoH(\equiv CCH_2Bu^t)\{P(OMe)_3\}_2(\eta$-$C_5H_5)$]-[$BF_4$], [$Mo\{\sigma$-$OC(O)CF_3\}^-\{\eta^2$-$HOC(O)CF_3\}\{P(OMe)_3\}_2(\eta$-$C_5H_5)$], and [$Mo(\equiv CCH_2Bu^t)\{\eta^2$-$C_2(CN)_4\}^-\{P(OMe)_3\}(\eta$-$C_5H_5)$].* M. Bottrill, M. Green, A. G. Orpen, D. R. Saunders, and I. D. Williams. **Journal of the Chemical Society (Dalton Transactions)**, 1989, pp. 511–518.

[154] *Binuclear and Mononuclear Di-η^5-cyclopentadienylniobium Chemistry — a New Insight — Crystal and Molecular Structure of [Cp_2NbH_2]-[Na.Benzo-15-Crown-5]$^+$.* D. A. Lemenovskii, I. E. Nifant'ev, I. F. Urazowski, E. G. Perevalova, T. V. Timofeeva, Yu. L. Slovokhotov, and Yu. T. Struchkov. **Journal of Organometallic Chemistry**, Vol. 342 (1988), pp. 31–44.

[155] *Facile Olefin Hydrogenation with an Osmium Dihydrogen Complex.* T. J. Johnson, J. C. Huffman, K. G. Caulton, S. A. Jackson, and O. Eisenstein. **Organometallics**, Vol. 8 (1989), pp. 2073–2074.

[156] *A Lipophilic Salt of a Transition-Metal Polyhydride — $KOsH_3(PMe_2Ph)_3$.* J. C. Huffman, M. A. Green, S. L. Kaiser, and K. G. Caulton. **Journal of the American Chemical Society**, Vol. 107 (1985), pp. 5111–5115.

[157] *Synthesis and Reactions of Dinuclear Palladium Complexes containing Methyls and Hydride on Adjacent Palladium Centers — Reductive Elimination and Carbonylation Reactions.* S. J. Young, B. Kellenberger, J. H. Reibenspies, S. E. Himmel, M. Manning, O. P. Anderson, and J. K. Stille. **Journal of the American Chemical Society**, Vol. 110 (1988), pp. 5744–5753.

[158] *Activation of the C-H Bonds of Benzene by Reaction with [Bis(dicyclohexylphosphino)ethane]platinum(0), Generated by the Thermolysis of cis-[Bis(dicyclohexylphosphino)ethane]hydridoneopentyl-platinum(II).* M. Hackett, J. A. Ibers, and G. M. Whitesides. **Journal of the American Chemical Society**, Vol. 110 (1988), pp. 1436–1448.

[159] *Novel Platinum Complexes enable Selective C-H and C-Si Activation of Tetramethylsilane at Room-Temperature.* P. Hofmann, H. Heiss, P. Neiteler, G. Muller, and J. Lachmann. **Angewandte Chemie — International Edition in English**, Vol. 29 (1990), pp. 880–882.

[160] *Synthesis and Structural Characterization of a Mixed-Ring Rhenocene Hydride Complex, $(\eta^5\text{-}C_5Me_5)(\eta^5\text{-}C_5H_5)ReH$.* R. A. Paciello, P. Kiprof, E. Herdtweck, and W. A. Herrmann. **Inorganic Chemistry**, Vol. 28 (1989), pp. 2890–2893.

[161] *Preparation, Dynamic Behavior, and C-H and C-C Cleavage Reactions of $(\eta^4\text{-}C_5H_6)Re(PPh_3)_2H_3$ — Structures of $(\eta^4\text{-}C_5H_6)Re(PPh_3)_2H_3$, $CpRe(PPh_3)_2H_2$, and $CpRe(PPh_3)H_4$.* W. D. Jones and J. A. Maguire. **Organometallics**, Vol. 6 (1987), pp. 1301–1311.

[162] *Photoinitiated Intramolecular Hydrogen Transfer from Rhenium Polyhydrides to C_8 Cyclopolyolefins.* M. C. L. Trimarchi, M. A. Green, J. C. Huffman, and K. G. Caulton. **Organometallics**, Vol. 4 (1985), pp. 514–519.

[163] *Multiple Bonds between Main-Group Elements and Transition-Metals .66. The Reactivity of Hexahydrido(η^5-pentamethylcyclopentadienyl)rhenium(VII), $(\eta^5\text{-}C_5Me_5)ReH_6$.* W. A. Herrmann, H. G. Theiler, E. Herdtweck, and P. Kiprof. **Journal of Organometallic Chemistry**, Vol. 367 (1989), pp. 291–311.

[164] *Reactions of the Polyhydride Complex $ReH_7(PPh_3)_2$ with Hydroxypyridine and Mercaptopyridine Ligands — Formation of Hydrido Complexes of Rhenium(III), Rhenium(IV), and Rhenium(V) and the Characterization of 8-Coordinate Isomers in the Solid-State and in Solution.* M. Leeaphon, P. E. Fanwick, and R. A. Walton. **Inorganic Chemistry**, Vol. 30 (1991), pp. 4986–4995.

[165] *Platinum(0) Activation of Rhenium Polyhydride Complexes: Synthesis and X-Ray Crystal Structure of the Hydrido-ethylene Complex*

$[ReH_3(C_2H_4)_2(PPhPr^i{}_2)_2]$. N. J. Hazel, J. A. K. Howard, and J. L. Spencer. **Journal of the Chemical Society (Chemical Communications)**, 1984, pp. 1663–1664.

[166] *Unsaturated, PCy_2-Bridged Re-M Heterobimetallics (M = Rh, Ir, Pd — Cy = Cyclohexyl) — Metal Metal Bond Isomerism, Reversible P–H Bond Activation, and Cooperative Reactivity.* R. T. Baker, J. C. Calabrese, and T. E. Glassman. **Organometallics**, Vol. 7 (1988), pp. 1889–1891.

[167] *Synthesis and Structure of Rhodium Hydride Dimethylphosphite Complexes.* E. E. Nifant'ev, T. S. Kukhareva, M. Yu. Antipin, and Yu. T. Struchkov. **Zh.Obshch.Khim.(some Russian Thing)**, Vol. 55 (1985), pp. 298–304.

[168] *Dynamic Properties of Phosphinehydridorhodium(I) Complexes and the Structure of Tris(triisopropylphosphine)hydridorhodium(I) $RhH(P(i\text{-}Pr)_3)_3$.* T. Yoshida, D. L. Thorn, T. Okano, S. Otsuka, and J. A. Ibers. **Journal of the American Chemical Society**, Vol. 102 (1980), p. 6451.

[169] *Azarhoda-closo-dodecaborane.* H. P. Hansen, J. Muller, U. Englert, and P. Paetzold. **Angewandte Chemie — International Edition in English**, Vol. 30 (1991), pp. 1377–1379.

[170] *Reactions of Catecholatoborane with Phosphinorhodium Complexes: Molecular Structures of $[RhHCl(Bcat)(PPr^i{}_3)_2]$ and $[((Pr^i{}_2PCH_2)_2)Rh((\eta^6\text{-}cat)Bcat)]$ (cat = 1,2-$O_2C_6H_4$).* S. A. Westcott, N. J. Taylor, T. B. Marder, R. T. Baker, N. J. Jones, and J. C. Calabrese. **Journal of the Chemical Society (Chemical Communications)**, 1991, pp. 304–305.

[171] I. P. Kovalev, K. V. Evdakov, M. G. Vinogradov, A. I. Yanovskii, and Yu. T. Struchkov. **(Metalloorg.Khim.(Organometallic Chem.in USSR))**, Vol. 2 (1989), p. 673.

[172] *C-H rather than O-H Activation — Synthesis and Molecular Structure of a Cationic Cis-hydrido-σ-acetylide Complex of Rhodium.* T. B. Marder, D. Zargarian, J. C. Calabrese, T. H. Herskovitz, and D. Milstein. **Journal of the Chemical Society (Chemical Communications)**, 1987, pp. 1484–1485.

[173] *Synthesis and Structures of Rhodium Isocyanide Complexes Containing an η^2-hydrotris(3,5-dimethylpyrazolyl)borate Ligand.* W. D. Jones and E. T. Hessell. **Inorganic Chemistry**, Vol. 30 (1991), pp. 778–783.

[174] *Hydration and Reduction of Carbon Dioxide by Rhodium Hydride Compounds, Preparation and Reactions of Rhodium Bicarbonate and Formate Complexes and the Molecular Structure of $RhH_2(O_2CHO)(P(i\text{-}Pr)_3)_2$.* T. Yoshida, D. L. Thorn, T. Okano, J. A. Ibers, and S. Otsuka. **Journal of the American Chemical Society**, Vol. 101 (1979), pp. 4212–4221.

[175] *Structural Determinations of Chlorobis(triisopropylphosphine)rhodium Compounds in the +2 and +3 Oxidation-States.* R. L. Harlow, D. L. Thorn, R. T. Baker, and N. L. Jones. **Inorganic Chemistry**, Vol. 31 (1992), pp. 993–997.

[176] *Preparations and Reactions of Some Hydridodinitrogentrialkylphosphine Complexes of Rhodium(I). The Structure of a Dinitrogen-Bridged Rhodium(I) Dimer, $[RhH(P(i\text{-}Pr)_3)_2](\mu\text{-}N_2)$.* T. Yoshida, T. Okano, D. L. Thorn, T. H. Tulip, S. Otsuka, and J. A. Ibers. **Journal of Organometallic Chemistry**, Vol. 181 (1979), pp. 183–201.

[177] *Reactions of Dihydrido(bicarbonato)bis(triisopropylphosphine)rhodium(III) with Alkynes — Formation of $Rh_2H_2(O_2CO)(PHC{=}CPH)(P(i\text{-}Pr)_3)_3$ and the Stereoselective Hydrogenation of Alkynes to Trans Olefins.* T. Yoshida, W. J. Youngs, T. Sakaeda, T. Ueda, S. Otsuka, and J. A. Ibers. **Journal of the American Chemical Society**, Vol. 105 (1983), pp. 6273–6278.

[178] *Multiple Routes to the Formation of a Phosphine-bridged Trirhodium Dihydride Cation, $[Rh_3(\mu\text{-}(Ph_2PCH_2)_2PPh)_2(H)_2(\mu\text{-}CL)_2(CO)_2]$.* A. L. Balch, J. C. Linehan, and M. M. Olmstead. **Inorganic Chemistry**, Vol. 25 (1986), pp. 3937–3942.

[179] *Zinc-containing Heterometallic Hydride Compounds — Observation of Terminal and Bridging $ZnN(SiMe_3)_2$ Groups.* R. L. Geerts, J. C. Huffman, D. E. Westerberg, K. Folting, and K. G. Caulton. **New Journal of Chemistry**, Vol. 12 (1988), pp. 455–465.

[180] *Oligophosphane Ligands .28. C-H Activation of Benzene and Alkylated Arenes by (Chelate-phosphane)ruthenium(0) Fragments.* L. Dahlenburg and K. M. Frosin. **Chemische Berichte**, Vol. 121 (1988), pp. 865–869.

[181] *Structure of [1,3-bis(diphenylphosphino)propane](η^5-cyclopentadienyl)hydridoruthenium(II).* S. A. Litster, A. D. Redhouse, and S. J. Simpson. **Acta Crystallographica**, Vol. C48 (1992), pp. 1661–1663.

[182] *Interaction of Trialkylaluminum Reagents with Metal-bound Ethylene and Carbon-Monoxide — the Molecular Structure of $(\eta^5\text{-}C_5Me_5)_2Ta(H)(C_2H_4\text{-}AlEt_3)$.* C. McDade, V. C. Gibson, B. D. Santarsiero, and J. E. Bercaw. **Organometallics**, Vol. 7 (1988), pp. 1–7.

[183] *Isolable Paramagnetic Hydride Complexes of Niobium(IV) and Tantalum(IV).* M. L. Luetkens, Jr., W. L. Elcesser, J. C. Huffman, and A. P. Sattelberger. **Journal of the Chemical Society (Chemical Communications)**, 1983, pp. 1072–1074.

[184] *Ir, Multinuclear-NMR, and Structural Studies on $[WH(CO)_2(NO)(PR_3)_2]$ — Cis-Influence of Phosphorus Ligands on Hydride Character.* A. A. H. van der Zeijden, C. Sontag, H. W. Bosch, V. Shklover, H. Berke, D. Nanz, and W. von Philipsborn. **Helvetica Chimica Acta**, Vol. 74 (1991), pp. 1194–1204.

[185] *The Crystal Structure of $[WH_2Cl_2(PMe_2Ph)_4]$ and its Dehydrochlorination to generate a Reactive Metal Center.* H. Dadkhah, N. Kashef, R. L. Richards, D. L. Hughes, and A. J. L. Pombeiro. **Journal of Organometallic Chemistry**, Vol. 255 (1983), p. C1.

[186] *Low-Temperature (20-K) Neutron Diffraction Study of $[K(Crypt\text{-}222)]^+[Cr_2(CO)_{10}(\mu\text{-}H)]^-$ — Influence of the Lattice Environment and the K Ion on the Anions Metal-Carbonyl Framework and the Ordered Cr–H–Cr Bond.* J. L. Petersen, R. K. Brown, and J. M. Williams. **Inorganic Chemistry**, Vol. 20 (1981), pp. 158–165.

[187] *A Comparison of Room-Temperature Single-Crystal Neutron and X-Ray Diffraction Studies of the Bis(triphenylphosphine)iminium Salt of the $[Cr_2(CO)_{10}(\mu\text{-}H)]^-$ Monoanion.* J. L. Petersen, P. L. Johnson, J. O'Connor, L. F. Dahl, and J. M. Williams. **Inorganic Chemistry**, Vol. 17 (1978), pp. 3460–3469.

[188] *A Closed 3-Center Carbon-Hydrogen-Metal Interaction — a Neutron Diffraction Study of $HFe_4(\eta^2\text{-}CH)(CO)_{12}$.* M. A. Beno, J. M. Williams, M. Tachikawa, and E. L. Muetterties. **Journal of the American Chemical Society**, Vol. 103 (1981), pp. 1485–1492.

[189] *The Crystal and Molecular Structure of Chloromer-tris(dimethylphenylphosphine)-cis-dihydridoiridium(III) by X-Ray and Neutron Diffraction.* G. B. Robertson and P. A. Tucker. **Australian Journal of Chemistry**, Vol. 41 (1988), pp. 641–650.

[190] *X-Ray and Neutron Diffraction Study of $[P(CH_2C_6H_5)(C_6H_5)_3]^+[HIr_4(CO)_{11}]^-$.* R. Bau, M. Y. Chiang, C.-Y. Wei, L. Garlaschelli, S. Martinengo, and T. F. Koetzle. **Inorganic Chemistry**, Vol. 23 (1984), pp. 4758–4762.

[191] *A Long Si–H Bond or a Short Si–H Nonbond — Neutron Diffraction Study of $(\eta^5\text{-}CH_3C_5H_4)(CO)_2(H)MnSiF(C_6H_5)_2$.* U. Schubert, K. Ackermann, and B. Worle. **Journal of the American Chemical Society**, Vol. 104 (1982), pp. 7378–7380.

[192] *Low-Temperature Neutron Diffraction Study of $HMn_2Re(CO)_{14}$ and Studies of a Metal Metal Exchange Equilibrium that Converts $HMn_2Re(CO)_{14}$ into $HMnRe_2(CO)_{14}$.* R. M. Bullock, L. Brammer, A. J. Schultz, A. Albinati, and T. F. Koetzle. **Journal of the American Chemical Society**, Vol. 114 (1992), pp. 5125–5130.

[193] *Neutron Diffraction Studies of the Metal–Hydrogen–Metal Bond. I. The Symmetric, Bent, Three-Center, Two-electron Molybdenum–Hydrogen–Molybdenum Bond in μ-Hydrido-μ-dimethylphosphido-bis(η^5-cyclopentadienyldicarbonylmolybdenum), $Mo_2(\eta^5\text{-}C_5H_5)_2(CO)_4(\mu_2\text{-}H)(\mu_2\text{-}P(CH_3)_2)$.* J. L. Petersen, L. F. Dahl, and J. M. Williams. **Journal of the American Chemical Society**, Vol. 96 (1974), pp. 6610–6620.

[194] *A Neutron Diffraction Crystallographic Study of the Tetramethylammonium Salt of the Hexachlorobis(μ-chloro)(μ-hydrido)dimolybdenum(III) Ion, $[Mo_2Cl_8H]^{3-}$.* F. A. Cotton, P. C. W. Leung, W. J. Roth, A. J. Schultz, and J. M. Williams. **Journal of the American Chemical Society**, Vol. 106 (1984), pp. 117–120.

[195] *Decacarbonyl-μ-hydrido-μ-vinyl-triangulo-triosmium: a Combined X-Ray and Neutron Diffraction Study.* A. G. Orpen, D. Pippard, G. M. Sheldrick, and K. D. Rouse. **Acta Crystallographica**, Vol. B34 (1978), pp. 2466–2472.

[196] *Neutron Diffraction Structure-Analysis of $[(\mu\text{-}H)(\mu\text{-}NCHCF_3)Os_3(CO)_{10}]$ at 20 K.* Z. Dawoodi, M. J. Mays, and A. G. Orpen. **Journal of Organometallic Chemistry**, Vol. 219 (1981), pp. 251–257.

[197] *Interaction of Hydrogen and Hydrocarbons with Transition Metals. Neutron Diffraction Study of Di-μ-hydrido-decacarbonyltriosmium, $(\mu\text{-}H)_2Os_3(CO)_{10}$ Containing a Four-Centre, Four-Electron H_2Os_2 Bond.* R. W. Broach and J. M. Williams. **Inorganic Chemistry**, Vol. 18 (1979), pp. 314–318.

[198] *Bridging Hydrides in $[H_2Os_3(CO)_{10}]$ and $[(C_2H_3)HOs_3(CO)_{10}]$: a Combined X-ray–Neutron Diffraction Study.* A. G. Orpen, A. V. Rivera, E. G. Bryan, D. Pippard, G. M. Sheldrick, and K. D. Rouse. **Journal of the Chemical Society (Chemical Communications)**, 1978, pp. 723–724.

[199] *Interaction of Hydrogen and Hydrocarbons with Transition Metals. Neutron Diffraction Study of the Crystal and Molecular Structure of $(\mu\text{-}H)_2Os_3(CO)_{10}(\mu\text{-}CH_2)$.* A. J. Shultz, J. M. Williams, R. B. Calvert, J. R. Shapley, and G. D. Stucky. **Inorganic Chemistry**, Vol. 18 (1979), pp. 319–323.

[200] *Structural and Vibrational Analysis of Nonacarbonyltri-μ-hydrido-μ_3-methylidyne-triangulo-triosmium — X-Ray and Neutron Diffraction Studies.* A. G. Orpen and T. F. Koetzle. **Acta Crystallographica**, Vol. B40 (1984), pp. 606–612.

[201] *Combined X-ray and Neutron Diffraction Study of $[Os_4(CO)_{11}H_3(CHCHPh)]$.* B. F. G. Johnson, J. Lewis, A. G. Orpen, P. R. Raithby, and K. D. Rouse. **Journal of the Chemical Society (Dalton Transactions)**, 1981, pp. 788–792.

[202] *Joint X-Ray and Neutron Diffraction Refinement of the Crystal and Molecular Structure of Tri-μ-hydrido-μ-iodo-cyclo-tetrakis(tricarbonylosmium)(4 Os–Os).* B. F. G. Johnson, J. Lewis, P. R. Raithby, K. Wong, and K. D. Rouse. **Journal of the Chemical Society (Dalton Transactions)**, Vol. 7 (1980), pp. 1248–1250.

[203] *X-Ray and Neutron Diffraction Analysis of the Structures of $H_3Re_3(CO)_{11}[P(C_6H_5)_3]$ and $H_4Os_4(CO)_{11}[P(OCH_3)_3]$.* C.-Y. Wei, L. Garlaschelli, R. Bau, and T. F. Koetzle. **Journal of Organometallic Chemistry**, Vol. 213 (1981), pp. 63–78.

[204] *Joint X-Ray and Neutron Diffraction Structure Analysis of* $[(\mu\text{-}H)_3Re_3(CO)_8\{(EtO)_2POP(OEt)_2\}_2]$. D. W. Prest, M. J. Mays, A. G. Orpen, and R. K. McMullan. **Journal of Organometallic Chemistry**, Vol. 240 (1982), pp. 395–406.

[205] *Crystal Structure of* $(HRh[P(O\text{-}i\text{-}C_3H_7)]_2)_2$ *— A Neutron Diffraction Study.* R. G. Teller, J. M. Williams, T. F. Koetzle, R. R. Burch, R. M. Gavin, and E. L. Muetterties. **Inorganic Chemistry**, Vol. 20 (1981), pp. 1806–1811.

[206] *Neutron-diffraction Study of a Symmetrical Ruthenium–Hydrogen–Ruthenium Bond in Nonacarbonylhydrido(t-butylethynyl)-triangulo-triruthenium.* M. Catti, G. Gervasio, and S. A. Mason. **Journal of the Chemical Society (Dalton Transactions)**, 1977, pp. 2260–2264.

[207] *X-Ray and Neutron Diffraction Studies of the Hydridotriruthenium Cluster Complex* $(\mu\text{-}H)Ru_3(CO)_7(\mu\text{-}As(C_6H_5)CH_2As(C_6H_5)_2)((C_6H_5)_2AsCH_2As(C_6H_5)_2).CH_2Cl_2$ *— an Example of Dibridged Metal-Metal Bond* $M(\mu\text{-}H)(\mu\text{-}X)M$, *Involving a Heavy Ridgehead Atom.* N. Lugan, J.-M. Savariault, G. Lavigne, and J.-J. Bonnet. **Journal of Crystallographic and Spectroscopic Research**, Vol. 13 (1983), pp. 389–405.

[208] *Preparation and Structural Characterization of* $[(Ph_3P)_2N][HFeRu_3(CO)_{13}]$ *and* $[(Ph_3P)_2N][HFe_2Ru_2(CO)_{13}]$. F. Takusagawa, A. Fumagalli, T. F. Koetzle, G. R. Steinmetz, R. P. Rosen, W. L. Gladfelter, G. L. Geoffroy, M. A. Bruck, and R. Bau. **Inorganic Chemistry**, Vol. 20 (1981), pp. 3823–3828.

[209] *X-ray and Neutron Diffraction Studies on* $[Ru_4(CO)_8(\mu\text{-}H)_4\{P(OCH_3)_3\}_4]$ *at 293 and 20K: Characterisation of the Vibrational Behaviour of Two-co-ordinate Hydrogen atoms.* A. G. Orpen and R. K. McMullan. **Journal of the Chemical Society (Dalton Transactions)**, 1983, pp. 463–474. C20 H40 O20 P4 Ru4.

[210] *X-ray and Neutron Diffraction Studies on μ-Hydrido-nonacarbonyl(nitrosyl)ditungsten* $[HW_2(CO)_9(NO)]$. *Evidence for a Closed Three-center Metal-Hydrogen-Metal Bond.* J. P. Olsen, T. F. Koetzle, S. W. Kirtley, M. A. Andrews, D. L. Tipton, and R. Bau. **Journal of the American Chemical Society**, Vol. 96 (1974), pp. 6621–6627. two pages missing.

[211] *X-Ray and Neutron Diffraction Studies on* $HW_2(CO)_8(NO)(P(OCH_3)_3)$. *A Compound with a Slightly Asymmetric Hydrogen Bridge Bond.* R. A. Love, H. B. Chin, T. F. Koetzle, S. W. Kirtley, B. R. Whittlesey, and R. Bau. **Journal of the American Chemical Society**, Vol. 98 (1976), pp. 4491–4498.

[212] *Improved Synthesis, Chemistry, and X-ray and Neutron Diffraction Studies of* $[(Ph_3P)_2N]^+{}_2[H_2W_2(CO)_8]^{2-}$. C.-Y. Wei, M. W. Marks, R. Bau, S. W. Kirtley, D. E. Bisson, M. E. Henderson, and T. F. Koetzle. **Inorganic Chemistry**, Vol. 21 (1982), pp. 2556–2565.

[213] *A Neutron Diffraction Study of* $[Ph_4P]^+[HW_2(CO)_{10}]^-$. D. W. Hart, R. Bau, and T. F. Koetzle. **Organometallics**, Vol. 4 (1985), pp. 1590–1594.

[214] M. Green, A. G. Orpen, I. D. Salter, and F. G. A. Stone. **Journal of the Chemical Society (Chemical Communications)**, 1982, p. 813.

[215] *Mixed Transition-Metal Gold Cluster Hydrides.* B. D. Alexander, B. J. Johnson, S. M. Johnson, A. L. Casalnuovo, and L. H. Pignolet. **Journal of the American Chemical Society**, Vol. 108 (1986), pp. 4409–4417.

[216] *New Bimetallic Cobalt(II) Complexes of Chelated, Bridged Phosphido Ligands.* L. Chen, D. J. Kountz, and D. W. Meek. **Organometallics**, Vol. 4 (1985), pp. 598–601.

[217] *Synthesis Characterization , and X-ray Structure of Confacial-Bioctahedral Iron(II)– and Cobalt(II)–Hydrido Complexes with Tridentate Tripod Ligands Containing Phosphorus and Arsenic as the Donor Atoms.* P. Dapporto, S. Midollini, and L. Sacconi. **Inorganic Chemistry**, Vol. 14 (1975), pp. 1643–1650.

[218] *Deprotonierung und Substitutionsreaktionen Dreikerniger Eisencluster $Fe_3(CO)_9(H)(SR)$.* A. Winter, L. Zsolnai, and G. Huttner. **Chemische Berichte**, Vol. 115 (1982), pp. 1286–1304.

[219] *Mössbauer and X-Ray Diffraction Studies on $HFe_3(CO)_9(SR)(R = i\text{-}C_3H_7$ and $t\text{-}C_4H_9)$. Example of Noncorrelation of M–M Distance with M–H–M Bonding.* R. Bau, B. Don, R. Greatrex, R. J. Haines, R. A. Love, and R. D. Wilson. **Inorganic Chemistry**, Vol. 14 (1975), pp. 3021–3025.

[220] *Crystal and Molecular Structures of $HFe_3(CH_3C{=}NH)(CO)_9$ and $HFe_3(N{=}CHCH_3)(CO)_9$. Isomeric Cluster Complexes Derived from the Reduction of Acetonitrile.* M. A. Andrews, G. van Buskirk, C. B. Knobler, and H. D. Kaesz. **Journal of the American Chemical Society**, Vol. 101 (1979), pp. 7245–7254.

[221] *Metal Clusters with Exposed and Low-Coordinate Nitride Nitrogen Atoms.* M. Tachikawa, J. Stein, E. L. Muetterties, R. G. Teller, M. A. Beno, E. Gebert, and J. M. Williams. **Journal of the American Chemical Society**, Vol. 102 (1980), pp. 6648–6649.

[222] *Mild, Reversible C-H Oxidative Addition Reductive Elimination in a Dinuclear Iridium Complex.* W. D. McGhee and R. G. Bergman. **Journal of the American Chemical Society**, Vol. 108 (1986), pp. 5621–5622.

[223] *Location of Hydrogen Atoms in $H_3Mn_3(CO)_{12}$. A Crystal Structure Determination.* S. W. Kirtley, J. P. Olsen, and R. Bau. **Journal of the American Chemical Society**, Vol. 95 (1973), pp. 4532–4536.

[224] *Synthesis and Structure of a Dimolybdenum Alkylidyne-hydride.* T. W. Coffindaffer, I. P. Rothwell, and J. C. Huffman. **Journal of the Chemical Society (Chemical Communications)**, 1983, pp. 1249–1251.

[225] *The Lewis Acidity of Nickel(0) .5. Synthesis and Structure of* $[Na(Tmeda)_2]^+[H\text{-}Ni_2(C_2H_4)_4]^{(-)}$. K. R. Porschke, W. Kleimann, G. Wilke, K. H. Claus, and C. Kruger. **Angewandte Chemie — International Edition in English**, Vol. 22 (1983), pp. 991–992.

[226] *Electron-Density Distributions in Metalloorganic Compounds — Changing Structural and Bonding Relationships in Dimeric Metalloorganic Nickel Hydrides with Ion-Pair Relationships to the Main Group-Metals Sodium and Lithium.* R. Goddard, C. Kruger, K. R. Porschke, and G. Wilke. **Journal of Organometallic Chemistry**, Vol. 308 (1986), pp. 85–103.

[227] *Photochemistry of Transition-Metal Polyhydrides — Dimerization of* $OsH_4(PMe_2Ph)_3$ *Following Photo-Dissociation of Dihydrogen.* M. A. Green, J. C. Huffman, and K. G. Caulton. **Journal of Organometallic Chemistry**, Vol. 243 (1983), p. C78.

[228] *Bis(triphenylphosphine)iminium* μ_2*–Hydrido–*μ_3*–sulphido–nonacarbonyltriosmate.* B. F. G. Johnson, J. Lewis, D. Pippard, and P. R. Raithby. **Acta Crystallographica**, Vol. 34 (1978), pp. 3767–3770.

[229] *Nonacarbonyl-di-*μ*-hydrido-(triphenylphosphine)-triangulo-triosmium.* R. E. Benfield, B. F. G. Johnson, J. Lewis, P. R. Raithby, C. Zuccaro, and K. Henrick. **Acta Crystallographica**, Vol. B35 (1979), pp. 2210–2212.

[230] *The Reduction of Triosmium Dodecacarbonyl by Sodium Borohydride. The Preparation and X-Ray Structure of the Cluster Anion* $[H_2Os_4(CO)_{12}]^{2-}$. B. F. G. Johnson, J. Lewis, P. R. Raithby, G. M. Sheldrick, and G. Suss. **Journal of Organometallic Chemistry**, Vol. 162 (1978), pp. 179–187.

[231] *Di-*μ*-hydrido-hexadecacarbonylpentaosmium.* J. J. Guy and G. M. Sheldrick. **Acta Crystallographica**, Vol. B34 (1978), pp. 1725–1727.

[232] *Chemistry of Di- and Tri-Metal Complexes with Bridging Carbene or Carbyne Ligands. Part 26. Reactions of the Compounds* $[PtW\{\mu\text{-}\sigma\text{:}\eta^3\text{-}CH(C_6H_4Me\text{-}4)\}(CO)_2(PR_3)_2(\eta\text{-}C_5H_5)][BF_4]$ $(PR_3 = PMe_3, PMe_2Ph,$ *or* $PMePh_2)$ *and* $[PtW\{\mu\text{-}\sigma\text{:}\eta^3\text{-}C(Me)C_6H_4Me\text{-}4\}(CO)_2(PMe_3)_2(\eta\text{-}C_5H_5)][SO_3CF_3]$ *with Nucleophiles; X-Ray Crystal Structure of* $[PtW(\mu\text{-}H)\text{-}\{\mu\text{-}CH(C_6H_4Me\text{-}4)\}(CO)_2(PMe)_2(\eta\text{-}C_5H_5)]$. J. C. Jeffery, I. Moore, and F. G. A. Stone. **Journal of the Chemical Society (Dalton Transactions)**, 1984, pp. 1571–1580.

[233] *Structure and Reactivity of a Dimetallic Rhenium Hydride Complex with a Rhenium–Rhenium Double Bond: X-Ray Crystal Structures of* $[Re_2(CO)_6\text{-}(Ph_2PCH_2PPh_2)H_2]$ *and* $[Re_2(CO)_6(Ph_2PCH_2PPh_2)\text{-}(NCHMe)H]$. M. J. Mays, D. W. Prest, and P. R. Raithby. **Journal of the Chemical Society (Chemical Communications)**, 1980, pp. 171–173.

[234] *Coordinately Unsaturated Clusters — a Novel Catalytic Reaction.* R. R. Burch, A. J. Shusterman, E. L. Muetterties, R. G. Teller, and J. M. Williams. **Journal of the American Chemical Society**, Vol. 105 (1983), pp. 3546–3556.

[235] *Novel Pyrazolylborato-ruthenium and Pyrazolyl-ruthenium Complexes — The Crystal and Molecular Structures of* $[(H(\mu\text{-}H)B(3,5\text{-}Me_2pz)_2)Ru(\eta^4\text{-}C_8H_{12})H]$ *and* $[(\eta^4\text{-}C_8H_{12})_2Ru_2H(\mu\text{-}N{=}C(Me)pz)(\mu\text{-}pz)(\mu\text{-}H)]$ *Confirmation of the 1st Examples of the Ru=H-B Agostic Interaction and the Bridging Amidine Ligand.* M. O. Albers, S. F. A. Crosby, D. C. Liles, D. J. Robinson, A. Shaver, and E. Singleton. **Organometallics**, Vol. 6 (1987), pp. 2014–2017.

[236] *A Novel Bonding Scheme of an Allen System in a Transition Metal Cluster Complex Molecular Structure of* μ_3*-*η*-(1-Methyl-3-ethyl-3-allenyl)-*μ*-hydrido-triangulo-tris(tricarbonylruthenium)(3Ru-Ru).* G. Gervasio, D. Osella, and M. Valle. **Inorganic Chemistry**, Vol. 15 (1976), pp. 1221–1224.

[237] *Crystal Structure and Molecular Geometry of* $HRu_3(CO)_{10}(C{=}NMe_2)$*, Including the Direct Location of the* μ_2*-Bridging Hydride Ligand.* M. R. Churchill, B. G. DeBoer, and F. J. Rotella. **Inorganic Chemistry**, Vol. 15 (1976), pp. 1843–1853.

[238] *Synthesis, Spectroscopic Characterization and Crystal and Molecular Structure of* $HRu_3(OCN(CH_3))(CO)_{10}$*.* R. Szostak, C. E. Strouse, and H. D. Kaesz. **Journal of Organometallic Chemistry**, Vol. 191 (1980), pp. 243–260.

[239] *Crystal Structure of Nonacarbonyl-*μ_3*-ethylidyne-tri-*μ*-hydrido-tri-ruthenium.* G. M. Sheldrick and J. P. Yesinowski. **Journal of the Chemical Society (Dalton Transactions)**, 1975, pp. 873–876.

[240] *Neutron Diffraction Structure-Analysis of a Hexanuclear Copper Hydrido Complex,* $H_6Cu_6[P(Para\text{-}Tolyl)_3]_6$ *— an Unexpected Finding.* R. C. Stevens, M. R. McLean, R. Bau, and T. F. Koetzle. **Journal of the American Chemical Society**, Vol. 111 (1989), pp. 3472–3473.

[241] *A Neutron Diffraction Study of* $HFeCo_3(CO)_9(P(OCH_3)_3)_3$*, a Metal Cluster Complex with a Triply Bridging Hydride Ligand.* R. G. Teller, R. D. Wilson, R. K. McMullan, T. F. Koetzle, and R. Bau. **Journal of the American Chemical Society**, Vol. 100 (1978), pp. 3071–3077.

[242] *Die Struktur des Nickel-Clusters* $H_3(C_5H_5)_4Ni_4$*.* G. Huttner and H. Lorenz. **Chemische Berichte**, Vol. 107 (1974), pp. 996–1008.

[243] *Low-Temperature X-ray and Neutron Diffraction Studies of Tetra-n-butylammoniumoctadecacarbonyl-*μ_3*-hydrido-octahedro-hexaosmium.* A. G. Orpen and T. F. Koetzle. **Acta Crystallographica**, Vol. C43 (1987), pp. 2084–2088.

[244] *X-Ray Structures of the Hexanuclear Cluster Complexes* $[Os_6(CO)_{18}]^{2-}$*,* $[HOs_6(CO)_{18}]^-$*, and* $[H_2Os_6(CO)_{18}]$*.* M. McPartlin, C. R. Eady, B. F. G. Johnson,

and J. Lewis. **Journal of the Chemical Society (Chemical Communications)**, 1976, pp. 883–885.

[245] *Low-Temperature Neutron Diffraction Analysis of* $[Rh_4(\eta^5\text{-}C_5Me_5)_4H_4][BF_4]_2$. J. S. Ricci, T. F. Koetzle, R. J. Goodfellow, P. Espinet, and P. M. Maitlis. **Inorganic Chemistry**, Vol. 23 (1984), pp. 1828–1831.

[246] *Nonacarbonyl-*μ_3*-hydrido-tris(*μ*-diphenylphosphido)-tetrahedro-tetrarhenium, a 2-fold Unsaturated 56-Valence Electron Cluster and a Related Gold Rhenium Compound.* H. J. Haupt, U. Florke, and H. Schnieder. **Acta Crystallograph -ica**, Vol. C47 (1991), pp. 2304–2307.

[247] *Synthesis of Pentanuclear Clusters Derived from* $[Ru_3RhH_2(CO)_{10}(PPh_3)(\mu\text{-}COMe)]$ *— Crystal Structures of* $[Ru_3Rh_2H_2(CO)_{12}(PPh_3)_2]$ *and* $[Ru_3RhAuH(CO)_{10}(PPh_3)_2\text{-}(\mu_3\text{-}COMe)]$. J. Evans, P. M. Stroud, and M. Webster. **Journal of the Chemical Society (Dalton Transactions)**, 1991, pp. 1017–1023.

[248] *Borane — THF as a Dual Reagent. Preparation and Characterization of the Metal-Rich Cobaltaborane* $(\mu_3\text{-}H)_2[(\eta^5\text{-}C_5H_5)Co_4B_2H_2$. Jiang Feilong, T. P. Fehlner, and A. L. Rheingold. **Journal of the American Chemical Society**, Vol. 109 (1987), pp. 1860–1861.

[249] *Structure of the Disubstituted Triphenyl Phosphine Derivative of* $HFeCO_3(CO)_{12}$. E. Iiskola and T. Pakkanen. **Acta Chemica Scandinavica**, Vol. A38 (1984), pp. 731–734.

[250] ^{59}*Co NMR-Study of Cluster Reactions — Solvent and Ligand Effects in Mixed-Metal, Tetrahedral* MCo_3 *(M = Fe, Ru) Carbonyl Clusters — Crystal Structure of* $FeCo3(\mu_3\text{-}H)(\mu\text{-}CO)_3(CO)_8(PPh_2H)$. P. Braunstein, J. Rose, P. Granger, J. Raya, S. E. Bouaoud, and D. Grandjean. **Organometallics**, Vol. 10 (1991), pp. 3686–3693.

[251] *Direct Metal Exchange in Germylidyne Tricobalt Clusters.* P. Gusbeth and H. Vahrenkamp. **Chemische Berichte**, Vol. 118 (1985), pp. 1770–1781.

[252] *Structural Trends in Phosphine-Substituted* $[RuCO_xRh_{3-x}H(CO)_{12}]$ *(X = 0-3) Clusters — Crystal Structures of* $[RuCo_3H(CO)_{11}(PR_3)]$ $(PR_3 = PMe_3$ *or* $PMe_2PH)$, $[RuCo_2RhH(CO)_{11}(PMe_2Ph)]$, $[RuCc_xRh_{3-x}H(CO)_{10}(PPh_3)_2]$ *(X = 1-2), and* $[RuRh_3H(CO)_{11}(PR_3)]$ $(PR_3 = PMe_3$ *or* $PMe_2Ph)$. J. Pursiainen, M. Ahlgren, T. A. Pakkanen, and J. Valkonen. **Journal of the Chemical Society (Dalton Transactions)**, 1990, pp. 1147–1153.

[253] *Dimethyl Sulfide Substituted Mixed-Metal Clusters — Synthesis, Structure, and Characterization of* $HRuCo_3(CO)_{11}(SMe_2)$ *and* $[HRuRh_3(CO)_9]_2[SMe_2]_3$. S. Rossi, J. Pursiainen, M. Ahlgren, and T. A. Pakkanen. **Organometallics**, Vol. 9 (1990), pp. 475–479.

[254] *Heteronuclear Complexes with Metal-Metal Bonds .6. Small Heteronuclear Clusters with Copper and $HCr(CO)_5$ Fragments - Structures of $[(N_2)Cu(\mu_3\text{-}H)Cr_2(CO)_{10}]$, ($N_2$=2$NH_3$, Bpy, 1,2-Diaminobenzene), $[\{(Py)Cu(\mu_3\text{-}H)Cr(CO)_5\}_2]$ and $[Cu(PPh_3)_3][Cr_2(\mu\text{-}H)(CO)_{10}]$ as well as the Structure of $[Zn(NH_3)_4][Cr_2(\mu\text{-}H)(CO)_{10}]_2$.* P. Klufers and U. Wilhelm. **Journal of Organometallic Chemistry**, Vol. 421 (1991), pp. 39–54.

[255] *The Heteronuclear Cluster Chemistry of the Group 1B Metals .1. Structural Similarities and Differences among Mixed-Metal Cluster Compounds containing Copper, Silver, or Gold Atoms Ligated by Phosphines — X-Ray Crystal Structures of $[CuRu_4(\mu_3\text{-}H)_3(CO)_{12}(PMePh_2)]$ and $[CuRu_3(CO)_9(C_2But)(PPh_3)]$.* R. A. Brice, S. C. Pearse, I. D. Salter, and K. Henrick. **Journal of the Chemical Society (Dalton Transactions)**, 1986, pp. 2181–2192.

[256] *The Heteronuclear Cluster Chemistry of the Group-1B Metals .2. Synthesis, X-Ray Crystal Structures, and Dynamic Behavior of the Bimetallic Hexanuclear Group-1B Metal Cluster Compounds $[Cu_2Ru_4H_2(CO)_{12}(PPh_3)_2]$, $[Au_2Ru_4H_2(CO)_{12}(PPh_3)_2]$.* M. J. Freeman, A. G. Orpen, and I. D. Salter. **Journal of the Chemical Society (Dalton Transactions)**, 1987, pp. 379–390.

[257] *The Heteronuclear Cluster Chemistry of the Group 1B Metals .6. Synthesis, Structures, and Dynamic Behavior of the Bimetallic Hexanuclear Group-1B Metal Cluster Compounds $[M_2Ru_4(\mu_3\text{-}H)_2(\mu\text{-}Ph_2P(CH_2)NPPh_2)(CO)_{12}]$ (M=Cu or Ag, N=1-6) — X-Ray Crystal Structures of $[Cu_2Ru_4(\mu_3\text{-}H)_2(\mu\text{-}Ph_2P(CH_2)NPPh_2)(CO)_{12}]$ (N = 2, 3, or 5).* S. S. D. Brown, I. D. Salter, and L. Toupet. **Journal of the Chemical Society (Dalton Transactions)**, 1988, pp. 757–767.

[258] *A Novel Bidentate Bonding Mode for a Triarylphosphine Ligand — The Synthesis and X-Ray Structure-Analysis of the Mixed-Metal Cluster Compound $[Cu_2Ru_4(\mu_3\text{-}H)_2(CO)_{12}(\mu\text{-}P(CH_2Ph)_2(\eta^2\text{-}CH_2Ph))]$.* P. J. McCarthy, M. McPartlin, H. R. Powell, and I. D. Salter. **Journal of the Chemical Society (Chemical Communications)**, 1989, pp. 395–396.

[259] *Mononuclear and Binuclear σ-Aryl Iron-Lithium Hydrides — Synthesis and Molecular Structure.* T. A. Bazhenova, L. M. Kachapina, A. E. Shilov, M. Yu. Antipin, and Yu. T. Struchkov. **Journal of Organometallic Chemistry**, Vol. 428 (1992), pp. 107–123.

[260] *Synthesis of Bimetallic Fe-Ni Carbonyl Clusters — Crystal Structure of $[N(CH_3)_3CH_2Ph][Fe_3Ni(CO)_8(\mu\text{-}CO)_4(\mu_3\text{-}H)]$.* A. Ceriotti, P. Chini, A. Fumagalli, T. F. Koetzle, G. Longoni, and F. Takusagawa. **Inorganic Chemistry**, Vol. 23 (1984), pp. 1363–1368.

[261] *Heteronuclear Metal-Complexes with Bridging Methoxymethylidyne Ligands — X-Ray Crystal Structures of $[AuRu_3(\mu_2\text{-}COMe)(CO)_{10}(PPh_3)]$ and $[Fe_3Pt(\mu_3\text{-}H)(\mu_3\text{-}COMe)(CO)_{10}(PPh_3)]$.* M. Green, K. A. Mead, R. M. Mills, I. D.

Salter, F. G. A. Stone, and P. Woodward. **Journal of the Chemical Society (Chemical Communications)**, 1982, pp. 51–53.

[262] *The Reaction of $Hg(O_2CCF_3)_2$ with $[Os_6(CO)_{18}]$ — Synthesis and Crystal Structures of 2 New Products — $[HOs_5Hg_2(CO)_{15}(O_2CCF_3)_3]$ and Cis-$[Os_4(CO)_{12}(O_2CCF_3)_2]$.* M. P. Diebold, B. F. G. Johnson, J. Lewis, V. P. Saharan, M. McPartlin, and H. R. Powell. **Journal of Organometallic Chemistry**, Vol. 405 (1991), p. C25.

[263] *Metal Atoms in the Synthesis of Metal-Clusters .2. Reactions of Substituted Cyclopentadienes and Nickelocene with Nickel and Iron Atoms.* J. J. Schneider, R. Goddard, C. Kruger, S. Werner, and B. Metz. **Chemische Berichte**, Vol. 124 (1991), pp. 301–308.

[264] *Structural Studies on Polynuclear Osmium Carbonyl Hydrides .35. Coordination and Scission of Alkynes on a Tungsten Triosmium Framework — Crystal Structures of $(\eta^5\text{-}C_5H_5)WOs_3(CO)_{10}[\mu_3\text{-}\eta^2\text{-}C_2(C_6H_4CH_3)_2]H$ and $(\eta^5\text{-}C_5H_5)WOs_3(CO)_{10}[\mu_3\text{-}\eta^2\text{-}C_2(CO_2C_2H_5)_2]H$ — Species with an Alkyne Ligand Bound to Os_3 and WOs_2 Faces of Tetrahedral WOs_3 Clusters.* J. T. Park, J. R. Shapley, C. Bueno, J. W. Ziller, and M. R. Churchill. **Organometallics**, Vol. 7 (1988), pp. 2307–2316.

[265] *A New Type of Bridging Carbonyl Ligand — $\mu_3\text{-}\eta^2$, and not one but 2.* R. D. Adams, Zhaoyang Li, Jau-Ching Lii, and Wengan Wu. **Journal of the American Chemical Society**, Vol. 114 (1992), pp. 4918–4920.

[266] U. Florke and H. J. Haupt. **Zeitschrift fuer Kristallographie**, Vol. 193 (1990), p. 309.

[267] U. Florke and H. J. Haupt. **Zeitschrift fuer Kristallographie**, Vol. 192 (1990), p. 300.

[268] *A Cluster with the Triply Facially Bridging Trimethylenemethane Ligand — Structure and Bonding of $[(Rh(Cod))_3(\mu_3\text{-}H)(\mu_3\text{-}C(CH_2)_3)]$.* G. E. Herberich, U. Englert, L. Wesemann, and P. Hofmann. **Angewandte Chemie — International Edition in English**, Vol. 30 (1991), pp. 313–315.

[269] *Synthesis and Characterization of $HRuRh_3(CO)_{12}$ — Crystal Structures of $HRuRh_3(CO)_{12}$, $HRuRh_3(CO)_{10}(PPh_3)_2$ and $HRuCo_3(CO)_{10}(PPh_3)_2$.* J. Pursiainen, T. A. Pakkanen, and J. Jaaskelainen. **Journal of Organometallic Chemistry**, Vol. 290 (1985), pp. 85–98.

[270] *Pentadienyl Ligands in Cluster Complexes — Synthesis and Structural Characterization of $Ru_5(CO)_{12}(\mu_4\text{-}S)(\mu\text{-}H)(\mu\text{-}1,5\text{-}Me_2C_5H_5)$ and $Ru_6(CO)_{15}(\mu_4\text{-}S)(\mu\text{-}H)(\mu\text{-}1,5\text{-}Me_2C_5H_5)$.* R. D. Adams, J. E. Babin, M. Tasi, and T. A. Wolfe. **Journal of the American Chemical Society**, Vol. 110 (1988), pp. 7093–7097.

[271] *Formation and Characterization of an 84-Electron Hexaruthenium Cluster — Spontaneous Decarbonylation of $H_2Ru_6(CO)_{18}$ to form $H_2Ru_6(CO)_{17}$.* D. A.

M^{c}Carthy, J. A. Krause, and S. G. Shore. **Journal of the American Chemical Society**, Vol. 112 (1990), pp. 8587–8589.

[272] *The Activation of Hydrogen by a High Nuclearity Metal Cluster Complex — the Synthesis and Structural Characterization of* $Ru_6(CO)_{15}(\mu\text{-}CO)(\mu_3\text{-}H)_2(\mu_4\text{-}S)$. R. D. Adams, J. E. Babin, M. Tasi, and T. A. Wolfe. **New Journal of Chemistry**, Vol. 12 (1988), pp. 481–486.

[273] *An Interstitial Hydrogen Atom in a Hexanuclear Metal Cluster: X-Ray and Neutron Diffraction Analysis of* $[(Ph_3P)_2N]^+[HCo_6(CO)_{15}]^-$. D. W. Hart, R. G. Teller, C. Y. Wei, R. Bau, G. Longoni, S. Campanella, P. Chini, and T. F. Koetzle. **Journal of the American Chemical Society**, Vol. 103 (1981), pp. 1458–1466.

[274] *Six-Coordinate Hydrogen: Structure of* $[(Ph_3P)_2N]^+[HCo_6(CO)_{15}]^-$ *by Single-Crystal Neutron Diffraction Analysis.* D. W. Hart, R. G. Teller, C. Y. Wei, R. Bau, G. Longoni, S. Campanella, P. Chini, and T. F. Koetzle. **Angewandte Chemie — International Edition in English**, Vol. 18 (1979), pp. 80–81.

[275] $[HRu_6(CO)_{18}]^-$*: a Ruthenium Anion Having an Interstitial H-Ligand. X-ray Crystal Structures of Two Modifications.* C. R. Eady, B. F. G. Johnson, J. Lewis, M. C. Malatesta, P. Machin, and M. M^{c}Partlin. **Journal of the Chemical Society (Chemical Communications)**, 1976, pp. 954–946.

[276] *Cluster Synthesis. Part 22. Synthesis and Structural Characterization of* $Ru_7(CO)_{19}(\mu\text{-}CNMe_2)(\mu_6\text{-}H)$ *Containing an Interstitial Hydride Ligand and* $Ru_5(CO)_{12}(PMe_2Ph)_2(\mu_4\text{-}\eta^2\text{-}CNMe_2)(\mu\text{-}H)$ *containing a Quadruply Bridging (Dimethylamino)carbyne Ligand.* R. D. Adams, J. E. Babin, and J. T. Tanner. **Organometallics**, Vol. 7 (1988), pp. 2027–2033.

[277] *Gold and Silver Hydrides: Synthesis of Heterobimetallic Ir–M (M = Au, Ag) Complexes and X-ray Crystal Structures of* $[(PPh_3)Au(\mu\text{-}H)IrH_2(PPh_3)_3](BF_4)$ *and* $[(PPh_3)Ag(\mu\text{-}H)IrH_2(PPh_3)_3](CF_3SO_3)$. A. Albinati, C. Anklin, P. Janser, H. Lehner, D. Matt, P. S. Pregosin, and L. M. Venanzi. **Inorganic Chemistry**, Vol. 28 (1989), pp. 1105–1111.

[278] *Neutron Diffraction Analysis of* $[H_3Pt_2(Ph_2PCH_2CH_2PPh_2)_2]^+[BPh_4]^-$ *— An Unusual Example of an Unsymmetrical* $[P_2Pt(\mu\text{-}H)_2PtP_2H]$ *Core.* M. Y. Chiang, R. Bau, G. Minghetti, A. L. Bandini, G. Banditelli, and T. F. Koetzle. **Inorganic Chemistry**, Vol. 23 (1984), pp. 122–124.

[279] *A Quadruply Hydrogen-Bridged Metal-Metal Bond. The Neutron Diffraction Analysis of* $H_8Re_2(PEt_2Ph)_4$. R. Bau, W. E. Carroll, R. G. Teller, and T. F. Koetzle. **Journal of the American Chemical Society**, Vol. 99 (1977), pp. 3872–3874.

[280] *Hydridic Atom Arrangement and Bonding in the Molecular Structure of* $(C_2H_5)_4N[Re_2(\mu\text{-}H)_3H_6(CH_3C(CH_2P(C_6H_5)_2)_3)]^-.CH_3CN$ *— a Complex con-*

taining 10-Coordinated Rhenium. S. C. Abrahams, A. P. Ginsberg, T. F. Koetzle, P. Marsh, and C. R. Sprinkle. **Inorganic Chemistry**, Vol. 25 (1986), pp. 2500–2510.

[281] *Transition Metal-Hydrogen-Alkali Metal Bonds — Synthesis and Crystal Structures of [K(18-Crown-6)][W(PMe$_3$)$_3$H$_5$], [Na(15-Crown-5)][W(PMe$_3$)$_3$H$_5$] and [(W(PMe$_3$)$_3$H$_5$Li)$_4$] and Related Studies.* A. Berry, M. L. H. Green, J. A. Bandy, and K. Prout. **Journal of the Chemical Society (Dalton Transactions)**, 1991, pp. 2185–2206.

[282] *Reactivity and Catalytic Activity of Heteronuclear Clusters .1. Fluxional Decapping of the Heterometallic Moiety in $H_3Ru_4(CO)_{12}AuPR_3$, $H_3Ru_4(CO)_{12}CuPR_3$ and the Crystal Structure of $H_3Ru_4(CO)_{12}AuPPh_3$.* J. Evans, A. C. Street, and M. Webster. **Organometallics**, Vol. 6 (1987), pp. 794–798.

[283] *Control of the Metal Core Geometries and Dynamic Behavior of Heteronuclear Group 1B Metal Cluster Compounds using Bidentate Phosphine-Ligands — X-Ray Crystal Structures of [$Au_2Ru_4(\mu$-H)(μ_3-H)(μ-$Ph_2PCH_2PPh_2$)$(CO)_{12}$] and [$Ag_2Ru_4(\mu_3$-H)$_2$-(μ-$Ph_2PCH_2PPh_2$)$(CO)_{12}$].* P. A. Bates, S. S. D. Brown, A. J. Dent, M. B. Hursthouse, G. F. M. Kitchen, A. G. Orpen, I. D. Salter, and V. Sik. **Journal of the Chemical Society (Chemical Communications)**, 1986, pp. 600–602.

[284] *The Heteronuclear Cluster Chemistry of the Group 1B Metals. Part 11. Effect of the Nature of the Bidentate Diphosphine Ligand on the Metal Framework Structures of the Gold Heteronuclear Cluster Compounds [$Au_2Ru_4(\mu_3$-H)(μ-H)(μ-$Ph_2P(CH_2)_nPPh_2$)$(CO)_{12}$] (n = 1-6). X-Ray Crystal Structures of [$Au_2Ru_4(\mu_3$-H)(μ-H)(μ-$Ph_2P(CH_2)_nPPh_2$)$(CO)_{12}$] (n=1 or 2).* S. S. D. Brown, I. D. Salter, A. J. Dent, G. F. M. Kitchen, A. G. Orpen, P. A. Bates, and M. B. Hursthouse. **Journal of the Chemical Society (Dalton Transactions)**, 1989, pp. 1227–1236.

[285] *Basic Cluster Reactions .3. Hetero Site Reactivities in $Ru_2Co_2(CO)_{13}$.* E. Roland and H. Vahrenkamp. **Organometallics**, Vol. 2 (1983), pp. 183–184.

[286] *Synthesis, Crystal Structure, and Fluxional Behavior of $Re_3H_2(CO)_{10}(\eta^5$-$C_7H_9)$.* T. Beringhelli, G. Ciani, G. D'Alfonso, P. Romiti, A. Sironi, and M. Freni. **Inorganic Chemistry**, Vol. 23 (1984), pp. 2849–2852.

[287] *Solution and Crystal Structures of the Hydridoruthenium Raft Clusters $H_2Ru_6(CO)_{15}(L)(C_6H_4O)$ (L=CO, P(OMe)$_3$).* S. Bhaduri, K. Sharma, H. Khwaja, and P. G. Jones. **Journal of Organometallic Chemistry**, Vol. 412 (1991), pp. 169–176.

[288] *Isolation and X-Ray Structure Determination of $H_2Ru_6(CO)_{16}(C_6H_4O)$, a Hydridoruthenium Raft Cluster.* S. Bhaduri, K. Sharma, and P. G. Jones. **Journal of the Chemical Society (Chemical Communications)**, 1987, pp. 1769–1770.

[289] *Single-Crystal X-Ray and Neutron Diffraction Studies of an η^2-Dihydrogen Transition-Metal Complex — Trans-[Fe(η^2-H_2)(H)($PPh_2CH_2CH_2PPh_2$)$_2$]BPh_4.* J. S. Ricci, T. F. Koetzle, M. T. Bautista, T. M. Hofstede, R. H. Morris, and J. F. Sawyer. **Journal of the American Chemical Society**, Vol. 111 (1989), pp. 8823–8827.

[290] *Low-Temperature Neutron Diffraction Study of Cis(tri-para-tolylphosphine)-heptahydridorhenium(VII), [$ReH_7(P(C_6H_4CH_3)_3)_2$] — a Polyhydride Complex with a Short H=H Contact that Models an Intermediate Stage in the Oxidative Addition of Dihydrogen at Transition-Metal Centers.* L. Brammer, J. A. K. Howard, O. Johnson, T. F. Koetzle, J. L. Spencer, and A. M. Stringer. **Journal of the Chemical Society (Chemical Communications)**, 1991, pp. 241–243.

[291] *The Preparation and Structure of a Dinuclear η^2-H_2 Complex (P–N)(η^2-H_2)Ru(μ-Cl)$_2$(μ-H)Ru(H)(PPh_3)$_2$, P–N = Fe(η-C_5H_3(CHMeNMe$_2$)P(i-Pr)$_2$-1,2)(η-C_5H_5).* C. Hampton, W. R. Cullen, B. R. James, and J.-P. Charland. **Journal of the American Chemical Society**, Vol. 110 (1988), pp. 6918–6919.

[292] *Solution and Solid-State Conformational Isomers of the Molecular Dihydrogen Complex $ReCl(H_2)(PMePh_2)_4$ — does it contain an Asymmetric Molecular Dihydrogen Ligand.* F. A. Cotton and R. L. Luck. **Inorganic Chemistry**, Vol. 30 (1991), pp. 767–774.

[293] *Preparation, X-Ray Molecular Structure, and Electronic-Structure of the 1st 16-Electron Dihydrogen Complexes $RuH(H_2)X(PCy_3)_2$.* B. Chaudret, G. Chung, O. Eisenstein, S. A. Jackson, F. J. Lahoz, and J. A. Lopez. **Journal of the American Chemical Society**, Vol. 113 (1991), pp. 2314–2316.

[294] *Characterization of the 1st Examples of Isolable Molecular-Hydrogen Complexes, $Mo(CO)_3(PCy_3)_2(H_2)$, $W(CO)_3(PCy_3)_2(H_2)$, $Mo(CO)_3(\pi$-$PR_3)_2(H_2)$, $W(CO)_3(\pi$-$PR_3)_2(H_2)$ — Evidence for a Side-On Bonded H_2 Ligand.* G. J. Kubas, R. R. Ryan, B. I. Swanson, P. J. Vergamini, and H. J. Wasserman. **Journal of the American Chemical Society**, Vol. 106 (1984), pp. 451–452.