An Evaluation of the Multimedia Personal Computer as an Assessment Tool

Francis Bryan Bell B.Sc. (Hons.) App. Chem.

A thesis submitted to the University of Huddersfield
in partial fulfilment of the requirements for
the degree of Doctor of Philosophy

March 1997

University of Huddersfield

Department of Chemical and Biological Sciences

Abstract

The aim of this research was to investigate and evaluate how computerised assessment could be progressed using current advances in technology (1993-1996) such as the increased computing power and storage capacity, and multimedia capabilities. This research would be conducted with respect to a typical chemistry degree course.

After a detailed literature survey it was concluded that the major computerised assessment technique was the multiple-choice format or a hybrid of it. A typical degree course was analysed in order to determine the learning outcomes and if these could be assessed by computerisation. Of the assessment methods used in a degree course, it was found that 57% of these could be computerised to some extent. These assessment methods were found to be objective in nature.

This research focused on investigating the feasibility of producing computerised assessment methods that in effect used the computer as an *electronic notepad* and an on-line tutor. Students would be able to input each step of their solution to a problem. The student would be assessed on that step of their solution and not only on the end answer.

It was found that computerised assessment could be extended beyond the multiple-choice format to assessing students' process skills in problem solving. Students can input each stage of their solution to a problem and receive feedback in real-time as to whether or not their methodology is correct.

Several areas have been investigated: practical chemistry, in particular apparatus assembly; physical chemistry, mathematical problem solving; organic chemistry, reaction mechanisms; chemical information and retrieval. During these investigations several prototypes of the applications were developed. These prototypes were developed using Visual Basic 3.0 Professional for Windows 3.1. These prototypes serve as the data that support this research.

List of Contents

List of Tables and Figures

Preface

It is intended to review the work that has been undertaken previously in computerised assessment. This will ensure that work carried out in this research does not duplicate work that has already been done. It will also highlight the achievements and that which is already known about the subject. Most importantly it will serve as the starting point from which this research can be developed.

Background theory will also be detailed to provide a better understanding fundamental concepts and keywords and techniques that this research will refer to.

This research is the amalgamation of several subject areas that are as follows: computerised assessment, multimedia, expert systems and chemistry. The main thrust of this research is that computerised assessment is in its infancy stages and is not fully developed. To date the multiple-choice question is at the core of computerised assessment and this has led to much debate within the academic community as to whether or not the multiple-choice question can assess higher cognitive skills other than recognition. It is not the aim of this research to investigate whether or not this is true but concentrate on extending computerised assessment beyond the multiple-choice format. This research will focus on whether other computerised assessment techniques can be developed in which the computer becomes more like an electronic notepad and behaves as an intelligent tutor being able to provide feedback to the student at each stage in solving a problem. The main focus is that of assessing process skills and not just assessing the end product.

The thesis will analyse a typical degree course in chemistry in order to determine what methods are currently used to assess students and those skills these assessments are assessing. A taxonomy will be produced as to whether or not these assessment methods could be computerised. This will provide a figure as to the extent computerised assessment could be implemented in such a course and the rationale for further development of this subject area.

The core of this thesis describes the work that was undertaken to determine if the hypothesis that computerised assessment can be extended beyond the multiple-choice format is true. The results are not of the conventional form of traditional wet chemistry PhDs. Instead the data to prove the thesis one way or the other are in the form of prototyped software packages. These packages by virtue of their existence will prove that the multiple-choice question format can be extended to be able to assess student process skills.

Acknowledgements

I would like to extend my deepest thank you to Dr J.A. Wood without whom this research project would not have been made possible. I am indebted to him not only for the many hours of work that he has invested in this project, the many debates and intellectual arguments that we have had, but also for his foresight and forward thinking for realising the worthiness of such a research topic. It is this ambition and initiative that will continue to extend the scope of research with in chemistry.

I would like to thank Dr D. Brady for his invaluable contributions throughout the duration of this research project on the many educational issues encountered.

I would also like to thank the CTI for Chemistry at Liverpool University especially to Dr S.Walker, Dr R.Gladwin and Dr A.Drury, for their advice at the start of this research and continued support throughout.

Finally, I would like to thank those subject experts and students that took part in answering questionnaires and providing an input / views on the prototype applications developed.

Section One

Literature Survey

and

Background Theory

1. Section Overview

Within this section the key terms in assessment are introduced to provide the reader with a fundamental understanding of concepts within this subject area. A concise review of the issues and trends within assessment is also described to illustrate different expert opinions and arguments as to the purpose of assessment. A categorisation of current assessment methods as defined by experts is also listed and described. This is to provide, for completeness, a definition of all the types of categories in which assessment methods may be defined, of which computerised assessment is one such category.

Some of the reasons (political and academic) as to why computerised assessment has evolved according to the literature are reviewed. The term *computerised assessment* is defined for the purposes of this thesis. A history of computerised assessment methods is described that have led to the definition of the four generations of computerised assessment. A review that has been compiled from the literature is then described as to the justification of the adoption of computerised assessment as a valid assessment technique with respect to the perceived problems of paper-based tests. The advantages and disadvantages as described in the literature are also listed and described. A compilation of computerised assessment methods is listed.

A critique of the multiple-choice question (MCQ) is detailed in the final sub-section. This sub-section reviews the advantages and disadvantages of the MCQ as an assessment technique and subsequent hybrids and improvements. The section concludes why this assessment method can not be a major assessment technique.

In the next sub-section Multimedia the techniques that have extended computerised assessment are described. The actual terms *multimedia* and *interactive* are reviewed. Hardware requirements for multimedia are described. Each of the multimedia elements is analysed.

The final sub-section Expert Systems describes the background theory of this subject area. The elements of an expert system are defined. The techniques used by experts systems are described.

1.1 Assessment Introduction

It is the aim of this sub-section to introduce the key terms and concepts that are fundamental to the understanding of this subject area. A concise review of the issues and trends within assessment is also described to illustrate different expert opinions and arguments as to the purpose of assessment. A list of the different types of assessment methods is included. This is included to illustrate some of the types of categories, into which assessment methods may be categorised, of which computerised assessment is one such category.

1.1.1 The Aim of Assessment

Assessments are used in a variety of situations in education and in the workplace to measure a number of constructs for example: mastery of a subject, to predict future performance, to measure personality, attitudes, and skills. It is also used to provide feedback to an individual to indicate how well the individual is doing within the specific area.

Assessment can be used in two primary modes [1]:

> *Criterion-Referenced*: This method is used to determine if a certain criterion has been achieved i.e. a simple pass / fail grading system;
>
> *Norm-Referenced*: This method is used to compare individuals with others to produce a rank order in which each individual's score is relative to one another.

If the assessment is given at the end of a course, for example to measure the total learning that had taken place within a course, then it is being used *summatively* [1,3]. If on the other hand feedback is given during the course the assessment is being used *formatively*[1,3]. If the assessment is designed to observe and measure those attributes that are and are not being used then it is being used *diagnostically* [1,3].

1.1.2 Assessing Assessment

An assessment itself is measured in terms of its *reliability* and *validity* [1]. The *reliability* of an assessment is a measure of the consistency of the test to produce the same result on any occasion that it is used. The *validity* is a measure of whether the assessment does assess what it purports that it is assessing.

1.1.3 Issues and Trends in Assessment

Before the 1970s the end-test or end-of-year examination was the primary mode of paper-based assessment within education. Some of the perceived advantages of this kind of assessment are [1]:

- each student does the same examination within a specific time period;
- each student does not know what is on the examination paper beforehand;
- each student works independently of each other and of any other source of information.

However, opponents argue that [1]:

- students are usually given a choice of questions from which they are required to do three from five for example, and hence each student does not do the same examination;
- students in some cases are given an indication of what will be on the examination and of the format of the paper;
- students may not perform to their ability "on the day" because of the pressure of doing examinations and, as a result, their score is not a true reflection of their actual ability;
- this form of assessment does not measure all the skills that a student should possess for example, oral skills;
- some students are bad at examinations due to stress and it is not a fair means of assessment for them.

As a result other types of assessment, for example coursework, have been introduced to measure the wide range of knowledge and skills in order to give several opportunities for assessment. It is hoped that implementing a range of assessment methods will provide a better picture of each student's ability.

Another shift of emphasis that has changed assessment is that courses have moved away from **course assessment** to **modular assessment** with the advent of Credit Accumulation Transfer Schemes (CATS). However, the major concern with **modular assessment** is that there is "assessment overload" and that students have not had the time to learn the material before they are getting assessed on it.

There has been a shift away from **learning objectives** to that of **learning outcomes**[1]. Assessments should be designed to assess these intended learning outcomes.

It is now recognised that students should not only develop and be assessed on their **subject expertise** but also develop and be assessed on their **transferable skills** such as: working as part of a team, time management, project planning, negotiation, information handling etc.

1.1.4 Types of Assessment Methods

There is a number of different types of assessment currently used to assess learning outcomes. Gibbs *et al* [2] came up with 53 different assessment methods. The list below represents their various headings that they used to categorise assessments:

- Essays;

  Students are assessed on their ability to communicate their answer in written form. The assessment is mainly subjective in nature i.e. is dependent on the marker - students may receive different marks depending on the marker.

- Objective tests;

[1] *"Learning outcomes include learning objectives but also include exploring what other things have been learnt or achieved, the unintended as well as the intended."* [1]

These tests can be marked absolutely i.e. there is an absolute value which the students mark is either agrees or disagrees with.

- Alternative exams;

  One of the perceived problems with the conventional unseen examination is that of regurgitation i.e. students rote memorise answers and regurgitate them during the examination. This will have a short-term consequence for learning in that a few days after the examination the student will have forgotten the subject material.

- Computerised assessments;

  These assessments are objective in nature i.e. they have an absolute value. The computer can then compare the correct answer to the student's answer to determine if they are correct or incorrect. The computer can not mark subjective assessments.

- Practical and project work;

  These assessments are based on a variety of skills, for example writing reports, orals and actually carrying out the practical. These assessments will be, in general, subjective in nature.

- Criterion-based assessments;

  In criterion-based assessments students are not compared against one another, as is the case in norm-referenced assessments but are compared against a set of criteria. The student will either pass or fail an assessment.

- Feedback methods;

  These are formative assessments in which the student receives feedback as to their performance.

1.1.5 Conclusion

This sub-section briefly introduced assessment and its key concepts. An attempt has made to highlight changing expert opinions as to how assessment should be implemented and what skills should be assessed. Some of the categories in which assessment methods can be listed were also described. This was to indicate how computerised assessment fitted in to assessment as a whole.

1.2 Computerised Assessment Introduction

This sub-section follows on from 1.1 *Assessment*, which reviewed assessment in general terms, to now focus specifically on computerised assessment. This sub-section reviews some of the reasons (both political and academic) as to why computerised assessment has evolved. The term *computerised assessment* is defined for the purposes of this thesis. A rationale for computerised assessment is described with respect to the perceived disadvantages of paper-based tests; the advantages and disadvantages of computerised assessment.

Known computerised assessment methods and question types have also been compiled and described.

The final sub-section focuses on the multiple-choice question that is at the heart of the computerised assessment debate. This sub-section reviews the advantages and disadvantages of this question type; subsequent hybrids and improvements in an attempt to silence its critics; and finally concludes why this question type can not be a major assessment technique.

1.2.1 The *Need* for Computerised Assessment

There has been a number of contributing trends in recent years that have led to the increased usage of computers both as a source for learning and as a means for assessment. These are as follows:

- There is an increase in student numbers entering Higher Education that is increasing the student / staff ratios. As a result there is more pressure on staff time with increased administration, assessments etc. Computerised testing is seen to be a part of the solution in helping to reduce the staff work load by administering and marking the tests for the lecturer;

- There is a shift away from lecturer-centred teaching to student-centred learning in which students are encouraged to learn for themselves as opposed being told what to learn. Computer-assisted learning is seen to be one such student-centred learning activity. Assessment is the tool that is used to measure learning. Computer-assisted learning will incorporate assessment to provide the student with a means of measuring their learning;

- There is a shift away from the traditional "end-examination" as being the only means of assessing all the learning outcomes of the course, towards the inclusion of other assessment methods such as course work assignments, open-book examinations, oral presentations etc. Computerised assessment is another form of assessment;
- There is an increase in the availability of computers within Higher Education;
- There has been a dramatic increase in the power of hardware and software. The result of this has been the phenomenon of the "multimedia" computer. This has enabled, for the first time, the ability to display photographs, video, sound, animation and graphics on the computer. A "multimedia" environment will provide an enriched learning environment for the student by enabling, where appropriate, concepts to be displayed visually or by audio or both, instead of in the traditional textual manner.

It has been said that computerised assessment *"cannot replace all other forms"* of assessment *"but it can relieve the pressure when used as a substitute for some of them"* [4]. The *"pressure"* being referred to here is with regards to increasing student numbers and providing them with adequate feedback at an early stage in the assessment process. However, the important omission in this statement is "at the present time". The scope of computerised assessment is only limited to the technology of the day - future developments in technology will broaden the range of those skills that are not presently conceivable as being assessed by the use of computers. The article goes on to say [4]:

> *"All forms of assessment ... have strengths and weaknesses. Where the strengths of computerised assessment outweigh the strengths of conventional assessment it should be used. Where they do not they should be rejected."*

It is one of the aims of this research to identify the strengths and weaknesses of computerised assessment in order to determine where computerised assessment should be used in chemistry.

The next sub-section, 1.2.2 *Definition of Computerised Assessment*, defines what the term *computerised assessment* will mean in the context of this thesis.

1.2.2 Definition of the Term *Computerised Assessment*

One of the most noticeable problems when reviewing the literature on computerised assessment is the broad range of terms that have been used to describe assessments that in some way use computers. For example, Computer Adaptive Testing; Computer-Aided Assessment; Computer-Assisted Assessment; Computer-Assisted Marking; Computer-Based Assessment; Computer-Based Education; Computer-Based Examination; Computer- Based Examination System; Computer-Based Multiple-choice Format; Computer-Based Testing; Computer Delivered Testing; Computerised Examination; Computerised Testing. This list is in no way exhaustive.

These terms refer to the many ways in which computers can be used in the assessment process. For example:

- to act as a delivery system to present material from which a paper-based assessment can be undertaken;
- to present assessment material that the student can answer using the computer - the responses can be stored in a database from which reports can be produced to provide feedback to the student and lecturer;
- to assist in the administration, marking and analysis of results etc.
- to describe the different types of computerised assessment.

Bull [3] has used the term *"technology"* when referring to machines that are used in the assessment process. She has identified two principal uses of technology in the assessment process:

- *"as a tool of assessment"*. This has been described as using technology to produce material on which assessment can be based. The example Bull [3] uses is the production of a video that students would be required to make an oral presentation based on the video. This would involve the students using a wide range of skills which they may not have used or been assessed on if doing a paper-based test.
- *"to assess learning"*. The application here is to use technology "to assign marks to an essay, practical or project." Technology is replacing the role of the lecturer to grade student's work. It is worth noting that Bull [3] does go on to say *"Essay marking by computers is not yet possible"*.

It is the intention of this thesis to focus on the second principal use and specifically the application of the computer as a means of presenting the assessment, assessing the student's work, recording their individual responses and providing constructive feedback.

The next section will provide a review of computerised assessment methods.

1.2.3 Types of Computerised Assessment Methods

The use of computerised assessment has been around for the last few decades. Gibbs *et al* [2] identified seven types of computerised assessment:

1. **Computer marking**;

   This method involves the computer being used to read and record responses made by the student on a card.

2. **Computer generated test papers**;

   The computer can be used to produce paper-based tests that have different responses to the questions when, for example, the student has to resit the test.

3. **Computer generated problems**;

   In the case of mathematical problem solving questions the computer can produce randomised numbers for each variable in the problem. This would then enable individualised tests to be presented to each student to avoid copying or so that the student could repeat the test but be provided with "new" problems. Gibbs *et al* [2] have provided a typical structure for such a program:

- a file containing the names of all the students;
- a question in a specified format;
- an algorithm to calculate the correct answer;
- a file containing permissible answers to the question;
- a lecturer's file in which the variables used for each student and the correct answer are stored;
- at the end a print-out would be produced containing the student's name, the problem / question and the student's mark.

They suggest that automatic marking could be achieved using a "computer marking scheme" (see above for details). However, they fail to take this idea one step further. The program could mark the student's answer. This would be much more effective than marking it by hand or by the "computer marking scheme" mentioned above. The student's answer could then be written to the lecturer's file.

4. **Computer feedback to students**;

   It is suggested that feedback other than a total mark or percentage would be beneficial for students learning i.e. tell the student which questions they got correct and incorrect. For the ones they got incorrect tell them why.

5. **Computer-based Keller Plan (Personalised System of Instruction (P.S.I.))**;

   6. Within the Keller Plan the student works on their own from a booklet. At the start of the course objectives are stated detailing what the student should be able to achieve on completion of each section. At the end of each section there are tests which the student must pass before they can continue with the course. Gibbs *et al* [2] have suggested that the computer could be used to deliver and mark these tests. When the student has achieved mastery of the subject by

achieving a certain percentage, the computer would print out the next set of objectives for the next unit.

7. **Assessed computer simulations**;

   It is suggested that by the way the student sets up the variables in the simulation it is possible to mark the assessment. The computer will be able to determine the answer from the variables that the student has chosen.

8. **Computer marked practicals**;

   9. Gibbs *et al* [2] explain that in practical sessions it is very difficult to determine how accurate a student's results are. It is very time consuming to check every student's calculations and that by the time students receive their reports back it is too late to rectify the mistakes and inaccuracies. However, by using the computer to mark student's accuracy and calculations in the laboratory immediate feedback can be given to the student during the laboratory session. A large time saving has been achieved because the student does not have to write up a laboratory report and the lecturer does not have to mark it. The computer can record each student's result and, if required, determine the mean for the class.

The question types mentioned by Gibbs *et al* [2] are dependent on objective questions in which there are clearly correct and incorrect answers.

In an attempt to bring together and classify computerised assessment there have been four generations of computerised assessment proposed by Bunderson *et al* [5]. These generations are as follows:

1. **Computerised testing (CT)**:

   These are "traditional tests" for example multiple-choice are delivered using the computer;

2. **Computerised adaptive testing (CAT)**:

   Within these types of test, the questions are tailored to the ability of the student i.e. the test continually adapts the level of difficulty of the next question based on whether the student answered the last question correctly or not;

3. **Continuous measurement (CM)**:

   Assessment is carried out continuously during a course. This is similar to the Keller plan which is mentioned previously in this section. Dynamic changes in the student's learning profile can be obtained;

4. **Intelligent measurement (IM)**:

   Knowledge bases and inferencing procedures are used to provide intelligent scoring, interpretations of student profiles and advice to lecturers and students.

Each of these generations supersedes the previous one. Of these four generations the first generation is the one that we are most familiar with and the one in which the majority of computerised assessment material has been developed.

In order to appreciate the current enthusiasm for computerised assessment we first need to examine what is wrong with the traditional paper-based tests and how computerised assessment is currently seen to be able to alleviate these problems.

1.2.4 Disadvantages of Paper-Based Tests

There are several problems that exist with traditional paper-based tests (PBTs). These include:

- Due to the increase of student /staff ratios there is ever increasing pressure on staff time to mark the increased PBTs and assignments. As a direct result of the increasing time to mark PBTs, increased administration etc. there is decreasing staff / student contact time.

- PBT items are fixed; each student does the same test as the other student. There is an argument that for the better students the test is too easy and for those of lower ability the test is too difficult. Therefore, the test is not beneficial to either group because the test does not really find out what each student really knows.

- One of the main criticisms of PBTs is that the time difference between doing the PBT and receiving feedback is large. This time difference could increase if the student numbers are to continue to rise. However, when the student receives their feedback for:

- formative assessment it is of little or no use due to:

  the pressure of marking on staff, they cannot give detailed feedback to the student;

  the time difference between the student doing the test and the student receiving feedback, the student has probably forgotten what the test was about and hence an opportunity for learning is missed.

- summative assessment it is usually in the form of a mark. This is of little use to the student's knowledge since they do not know the questions they got correct and those that they got incorrect. Students often think they have answered a question correctly when in fact they answered it incorrectly.

  Assessment can be an important part of the learning process. Assessment is specifically designed to determine what students know and diagnosing what they do not. However, due to the design of PBTs i.e. there is a large time difference between doing the test and receiving little or no constructive feedback, the assessment and learning process are not fully integrated. Assessment with good feedback would encourage the student to develop a reflective learning style.

- PBTs are marked by lecturers who do not have a common standardised marking scheme. Lecturers may and do weight marks differently for certain skills. As a result a student may be penalised under one lecturer's marking scheme and be rewarded under another's. PBTs are subjective tests even in science subjects - lecturers will disagree on what the important skills are for solving a problem.

- Hand marking in PBTs is also prone to errors - human errors. A lecturer may make errors in marking and tallying the scores etc.

To maintain this pro-computerised assessment perspective the next section presents the arguments for computerised assessment.

1.2.5 Advantages of Computerised Assessment

This sub-section is a compilation of the advantages of computerised assessment abstracted from the literature as described / perceived by academics.

> *"Conversion of paper tests to a computer is straight forward"* (- I assume this is MCQs) *"question banking, formal and informal testing, self-assessment with immediate feedback, random generation of questions in a test, collecting, analysing and tabulating of test data, flexibility of media."*[6]

There have been many advantages stated in favour of using computerised assessment. The following have been compiled from the literature:

- mark the assessment for the lecturer[7,8]
- report generating facilities[7,8]
- availability to test at any time[7,9]
- provide additional information and hints[7]
- provide instant feedback to student[7,8]
- enable statistical analysis of results; test validity[7,8]
- allow staff to concentrate on assisting students with problems and relieves some aspects of onerous repetitive marking[7]
- opportunity for the integration of learning and assessment and effective feedback[9]
- effective management, collation and transfer of assessment information[9]
- instantly accessible registration, scheduling, administration[8]

- solve some of the *"perceived errors in contemporary assessment techniques and potential cost savings"*[10].

The following list of advantages reflects a survey that was carried out by the National Council for Educational Technology of academic institutions. The percentage quoted beside the statement indicates the number of those who returned the survey who thought that this was an advantage [11]:

"saving of staff time (55%);

enabling easier analysis of results (55%);

providing a unified assessment framework (55%);

supporting national assessment requirements (55%);

providing objective information on student performance (47%);

allowing assessment to be made available at the time and place of student needs (47%);

encouraging students to take an active role in assessment (27%)".

The following quote is taken from an early study into assessment using computers to adapt tests to the ability of the student [12]:

"advantages of microcomputers and item response theory

they nearly eliminate error in deriving scores;

they reduce scoring time by up to 70 to 80%;

they provide a simple mechanism for storing and retrieving valuable information;

they have intrinsic motivation for the testee;

they have the speed to handle the evaluation of tests and their items (reliability, item difficulty, biserial correlations etc);

they have the ease to store data and to retrieve the data when the data has to be recalled;

they have the capabilities to detect aberrant response patterns;

they have the capabilities to provide ongoing group analysis of the test and item bias;

they have the capability to evaluate translations of measurement scales to different languages;

they have the capability to tailor the test to individual needs."

"advantages of computerised tests and computer adaptive tests over paper-based testing:

they have enhanced control in presenting item displays;

greater standardisation of test administration;

they offer improved test security;

they can enrich display information;

they can provide equivalent scores with reduced testing time;

they can improve the obtaining and coding of responses;

they can reduce measurement error;

they have the ability to measure response latencies for items and components;

they can provide improved scoring and reporting;

they can be automated for individually administered tests;

they can obtain records at a central site;

they have the ability to construct tests and create items by computer;

they have immediate test scoring and feedback;

they can provide an increased variety of testing formats, different languages."[5]

It can be easy to think after this volume of advantages of computerised assessment that computerised assessment is the way forward. However in order to form a balanced opinion the next section describes the disadvantages of computerised assessment and the advantages of paper-based tests.

1.2.6 Advantages of Paper-Based Tests

The perceived advantages of paper-based tests are:

- they can be done "anywhere" i.e. they do not need technology dedicated rooms to be administered;
- they are cheap in terms of materials and staff that do not need to be technology trained;
- students can do paper-based tests more easily than interacting with software with which they are not familiar;

1.2.7 Disadvantages of Computerised Assessment

The disadvantages of computerised assessment as reported in the literature are as follows:

"may not be immediately apparent"[9]

This disadvantage refers to the added time involved in the design and build of a computerised assessment - technical expertise is required to build computerised assessments in addition to the time required to write the assessment.

The following disadvantages are a result of the survey that was carried out by the National Council for Educational Technology that sent questionnaires to academic institutions:

"time required to input the data (55%)

lack of staff training (55%)

unsatisfactory software (31%)"

Other disadvantages were perceived to be:

"time needed to train staff

reluctance of staff to use the technology

difficulties they had with it

time taken to write objective questions

lack of computers for staff and students"[11]

The advantages purported in 1.2.5 essentially refer to multiple-choice questions. Computerised assessment only assesses for recognition and recall and does not assess higher order skills such as comprehension, integration and application of knowledge.

There has been an attempt over the sub-sections 1.2.4, 1.2.5, 1.2.6 and 1.2.7 to provide a balanced argument for and against paper-based tests and computerised assessments compiled from the literature.

In the next section a review is made as to what are the question types that can be actually asked in computerised assessment. This will determine exactly what computerised assessment can do in terms of the type of question that can be asked and hence the type of knowledge that is being assessed.

1.2.8 Computerised Assessment Question Types

There is a number of question types [9, 13-16] that have been developed for use in computerised assessment. These include:

- Randomised questions. There are two principal types of randomised question. The first is where a large database of questions (typically in the hundreds or even thousands) has been compiled (referred to as an item pool) and the computer then selects a number of questions from the item pool at random to produce an individualised test for each student. The second type refers essentially to numerical problems in which the computer can select randomised variables in the same question so that each student has a different question.
- True / false
- Multiple true / false
- Yes / No
- Multiple-Choice Questions (MCQs)

  In the standard format for MCQs the student is presented with a list of items, typically 3, 4 or 5, and the student has to select the correct answer from the list.

  Another form of MCQ is where a number of items in the list are correct and the student has to select all the items that answer the question.

  In the assertion / reason MCQ format the student is presented with a series of "assertions" and "reasons". The student has to identify the correct reason for the particular assertion.
- Numeric

- Text entry
- Pair matching
- Prioritising / re-ordering a list of items.
- Fill-in-the-gap

  These can be single and multiple gaps within a passage in which students type their answer directly in the gap.

  Another suggestion for this type of question is to provide help in the form of a list box that contains a multiple-choice list. The student then selects their answer from the list. If the student chooses this help facility the total score possible for that question is then halved.

These question types are objective in nature and in most cases involve selecting an option that is either correct or incorrect. Most of computerised assessment is concerned with the multiple-choice question. In the next section a critique of this question type is described.

1.2.9 Multiple-choice Questions

Over the last couple of decades the focus of computerised assessment has been around the standard MCQ [17-20]. This is primarily due to the "technology of the day". The computer is not (yet) capable of thinking for itself, producing tests and marking essay questions. However, MCQs are perfectly suitable for computerised assessment because:

- MCQs can be easily marked by the computer therefore eliminating human errors in marking and the subjectivity of the lecturer;
- item pools can be produced for MCQs. An item pool is the term used to describe a number of MCQs that are stored in a database. A program can generate randomised tests for each student based on these item pools. These pools can be used to generate repeatable tests[21] for students if the student does badly in the first instance;

- many of the advantages listed previously for computerised assessments including immediate feedback report generating etc. are also incorporated into MCQ computerised assessment.

However, critics of MCQs [18, 22, 23] argue:

- guessing is possible in MCQs and students should not be rewarded for this;
- students that have partial knowledge of the subject are penalised in MCQs since you are either correct or incorrect [24-26].
- MCQs do not let students express themselves. Instead the student has a set of predefined answers written down which uses the vocabulary of the lecturer.
- MCQs do not test for higher cognition skills such as problem solving, understanding etc. Critics argue that MCQs test only for recognition skills. There have been reports that MCQs do in fact test for higher order skills. I would agree with this in part but the level of student cognition is dependent on the student's attitude to approaching MCQs - the rogue guesser will always abuse the system.
- These fixed item tests (FIT) are too easy for higher ability students and too difficult for lower ability students.

Much research and time has been invested in:

a) investigating the effects of language, question sequence and adding more "distracters" in MCQs[27-34];
b) new scoring methods for MCQs in order :
    - to eliminate the guessing element [35-37];
    - to allow partial credit for partial knowledge to differentiate between the guessing student and those students who genuinely have a partial understanding of the subject content [24, 25];
c) generating new hybrids of MCQs:

- One suggested method to try to remove the guessing element in MCQs is as follows:

"Options shown individually on screen for five seconds but not at the same time, the student has to select a letter at the end. This reduces the probability that student will guess the correct answer by comparing the possible answers when listed together". [13]

- Generating problem solving MCQs.[38]
- Another method that has been developed to try to remove the guessing element is that of using a database or grid of items. The purpose is to increase the number options that the student has to choose from and hence decrease the chance of guessing the correct option.
- Computerised adaptive tests (CAT) [37, 39, 40-44] have been a recent development in computerised tests. They still largely employ the MCQ as the principal question type. In the CAT an item bank contains items (item is the technical term for MCQ) of varying degrees of difficulty. The student is asked a question and depending on whether the student gets the answer correct or incorrect the computer then selects the next question to be more difficult or less difficult respectively. In other words the computer "adapts" the test according to the ability of the student.
- An extension to the CAT is the self adapted test (SAT) [41, 45]. SATs differ from CATS in that the student selects the difficulty of the next question.

However, despite attempts to improve the format of the MCQ, it still is open to guessing and, more importantly, MCQs do not provide the student with an opportunity to express their knowledge and apply their interpretations of concepts etc. in written form - which is an important skill. MCQs in effect spoon-feed the student and do not allow the student to defend or show why they have reached their answer. It is alarming to hear in discussions that the MCQ computerised assessment should be used as the major examination in Higher Education.

1.2.10 Conclusion

Within this sub-section 1.2 *Computerised Assessment* a critique of this subject has been attempted. A review why this subject area has evolved in terms of political and academic reasons has been made and an analysis of what can be done in computerised assessment in terms of the types of questions that can be asked has also been achieved.

The conclusion that can be drawn from this section is that computerised assessment is still in its early development stages, with the major question type the multiple-choice question, which is still under much debate within the academic community. There has been an error within the academic community to use the terms computerised assessment and multiple-choice question interchangeably. It has to be stressed at this point computerised assessment is the overall set of which multiple-choice questions are a sub-set.

However, the potential advantages of computerisation, for example, the assessment is marked and administered for the lecturer - saving a large amount of the lecturer's time - justify further research into this area to expand the scope of computerised assessment beyond the multiple-choice question.

1.3 Introduction

It is the aim of this sub-section 1.3 *Multimedia* to introduce this subject area in terms of what multimedia actually is, what a computer must have to display multimedia capabilities and the techniques involved in producing multimedia. At the start of this research this was a relatively new subject area especially within computerised assessment. This section in addition to providing some background knowledge also includes details of some work that was carried out during this research. These details had not become common knowledge at the start of this research project. However, towards the end of this research publications were being released, as a result this information is detailed here [72-73].

1.3.1 Multimedia Question Types

Over the last decade the graphical user interface (GUI) has been developed and in recent years (1992-96) multimedia has become a buzzword in computer assisted learning (CAL) [45-68]. It has been said that computerised assessment *"is not as impressive as multimedia"* [4]. However multimedia is not just reserved for the development of CAL packages. It has very strong applications within computerised assessment. These technologies (GUI and multimedia) have extended the question types possible in computerised assessment to produce what some call "*Interactive Multimedia Computerised Assessments*". These question types include:

- Hotspots

  This is where an area of an image (photograph, diagram etc.) is selected (clicked on using the mouse) in response to a question.

- Drag and drop

  This feature enables objects on the screen to be picked up and moved (dragged) to a new position on the screen. For example this could be used to build molecules in chemistry or to label diagrams.

- Push button

The push button provides a familiar object (i.e. a button) to the user which they already understand the principles to use it i.e. push it in to turn it on or press it in to turn it off. Hence the function is to provide the student with a quicker method of selecting responses on screen instead of having to type the answer in (which may lead to typographical errors).

- Multimedia Pair matching

  The student is presented with a series of images and a list of items on screen. The student has to match the items to the images.[13]

- Selection from labelled grids[14]

  This is an extension of the grid of items for MCQs. Instead of being restricted to text, multimedia enables the information to include photographs, video clips, animations etc. A description of this question type follows:

  "An array of information is presented in the form of a numbered grid of 3x4, 4x5 or 5x5 cells. Contents can be pictures, words, ideas, equations, formulae, structures, definitions, numbers, operators or animated movies. Students consider the contents of each cell in response to a question and choose a cell or combination of cells. The number of cells chosen and the order of choice can be important. This method can assess deep learning of high-level skills, well beyond what is possible with the use of multiple-choice questioning."

Within academic groups the terms *multimedia* and *interactive* have been the source of much debate. The next two sections describe some of the arguments that have been central to this debate and conclude how the terms in this thesis will be used.

1.3.2 Definition of Multimedia

The term *multimedia* is itself the source of much confusion. Critics have pointed out that the multimedia computer is not truly "multimedia" because the senses of smell and touch are not possible. There have been a number of definitions that include the following:

- Nigel Woodhead in his book "Hypertext and Hypermedia" says it is a:

  "Superset of hypermedia and other interactive technology approaches".

- AimTech Corporation say:

  "Multimedia is not a product, nor even a technology. It is better seen as a platform ... a combination of hardware and software elements that together support a multisensory information environment".

- Dr. Jane Williams writing in the CTI Centre for Medicine's "Guide to Software and Resources" (July 1992) states that:

  "Multimedia combines several elements of communication: still and moving images, sound, text and graphics. Each of these in turn may be delivered by a variety of sources. Computers serve to combine these elements and control the delivery devices to provide powerful multimedia applications".

For the purposes of this research *multimedia* will be defined as:

> *"The combination of text with one or more or all of the following elements: graphics, animation, photographs, video and sound."*

The combination of text could be with ALL of these media elements but it does not necessarily have to be ALL; it could be any combination of one or more of these elements. The particular subject will determine the choice of combination. For example, music lends itself to the combination of text and sound whereas molecular chemistry lends itself to text and animation. (These examples are only intended to be generalised illustrations and not absolute rules.)

The definition of multimedia above is very similar to that of *hypermedia* which, is defined as:

> *"Hypermedia is a term widely used to mean text with additional material. For instance, some images or animation will be used in the creation of the courseware (computer-assisted learning materials) but the method of operation will still allow the student to browse or branch."*

1.3.3 Definition of Interactive

The term interactive has also been the focus of controversy. Heath [70] has concluded that the only genuine interactive exchange is between *two people*. This conclusion has been deduced in part from the definition of *interactive* based on Collins English Dictionary (1989):

> *""allowing or relating to continuous two-way transfer of information between a user and the central point of a communication system e.g. computer;*
>
> *two or more persons or forces acting upon or in close relation to each other"".*

Heath quotes Ann Malamah-Thomas:

> *""Interaction is ...more than action followed by reaction. Interaction means acting reciprocally, acting upon each other.""*

In comparison the Oxford English (1995) defines *interactive* as:

> *"reciprocally active; acting upon or influencing each other;*

(of a computer or other electronic device) allowing a two-way flow of information between it and a user, responding to the user's input".

As can be seen the Oxford English has produced a definition taking into account both opinions.

For the purposes of this thesis the term *interactive* will mean:

"A two-way flow of information between a computer and a user where the computer responds to a user's input to provide feedback".

The next section explores why multimedia is used.

1.3.4 Why use Multimedia?

Multimedia brings together the visual and sound communication media. Communication between individuals initially began with hand signals, grunts, drawings which eventually developed into language and written communication. Today, the MPC is just another communication tool in the communication evolution except that it enables us to provide true colour imagery and high quality sound that was not previously possible on any one machine.

One of the main disadvantages of textually based CAL packages is that the user is visually overloaded with information. With multimedia you can avoid visual overload by using narration and effective graphics, animation, photographs and video.

With such potential the MPC should enable us to assess and effectively communicate on an individual basis to students giving them detailed instant feedback which is not otherwise possible through a heavily overburdened lecturer.

In the next section hardware requirements are described in order for a computer to have multimedia facilities.

1.3.5 Hardware Requirements - MPC Level 1 and 2

Today, (1997), multimedia is taken for granted. Most computers come with multimedia capabilities already installed. At the start of this research project (1993) this was not the case. In order to have multimedia capabilities we needed to determine exactly what was required for a computer to be a *multimedia personal computer*. From the literature it was found that a standard had been produced in order for companies to sell a computer with a multimedia personal computer (MPC) logo. As such there are two standards specified by the MPC Marketing Council which are as follows:

MPC Specification 1.0 mid (1991) states that the:

> *base level MPC computer was a 12-MHz 80386 computer with 4 Mb of RAM and a 30 Mb hard disk; VGA Windows 3.0 with multimedia extensions or 3.1; Single speed CD-ROM drive, capable of transferring 150 kHz per second; 8-bit sound card, supports up to 11kHz.*

MPC Specification 2.0 Dec. (1992) stated that the:

> *base level MPC computer was a 80486SX/25, 4Mb of RAM, 160 Mb hard disk; VGA display with 65,536 colours; Windows 3.1; Double speed CD-ROM capable of transferring 300 kHz per second, XA compatible; 16-bit sound card, supports stereo sampling up 44 kHz, microphone, MIDI and joystick ports.*

However, these standards were not set very high. It was recommended by the CTI Centre for Chemistry, Liverpool for this project to purchase a 80486DX/50, 8Mb RAM, 540Mb hard disk; SVGA display, Double speed CD-ROM, 16-bit sound card and a video card (supports 256 colours).

The next section describes background theory of how the eye perceives colour and what requirements the multimedia computer needs to reproduce life-like colour images.

1.3.6 Reproducing Colour Images on the Visual Display Unit

The minimum number of colours that are required for the human eye to perceive a true-to-life photograph when using a visual display unit (VDU) display is 256 colours. The image must also consist of the same number of colours as that used by the VDU display.

Most VDU displays are set to standard VGA mode. This mode has a maximum of 16 colours available. If a photograph consisting of 256 colours is displayed on a screen set to 16-colour VGA mode, the resulting image is a very poor.

This can be easily corrected by loading the correct display driver, for example, 256 colours 640 x 480 resolution from the Windows Set-up program.

If you require a true colour photograph (photographic quality) then the 16.7M colour driver should be loaded. (For this option to be available it is dependent on the particular PC's available video memory.)

The next sections describe the various elements of multimedia - photographs, video and sound. Within each sub-section the techniques that are used to digitise each of these elements are described. These techniques were investigated using the equipment purchased for this project. The advantages and disadvantages found during this evaluation period are also described.

1.3.7 Photographs

1.3.7.a) Capturing Photographic Images

There were three methods investigated for acquiring these images. They were:

1.3.7.b) Kodak Photo CD

This is the best method for the acquisition of photographic images.[71]

In this method photographs are taken using a "normal" camera. When the film is finished the film is taken into a Kodak Photo CD registered retailer such as Boots.

The cost for 24 photographs is £12.99. The CD costs £4.99 but holds 100 photographs; once filled you have to buy a new CD. (1996 prices)

1.3.7.c) Hardware / Software Requirements of Kodak Photo CD

Kodak Photo CD compatible CD-ROM drive and Kodak Photo CD Access Software. A 8 x 10 inch Photo CD image requires 1.5 Mb of storage space.

1.3.7.d) Advantages of Kodak Photo CD

The main advantage of using Kodak Photo CD is that the images are high quality.

1.3.7.e) Disadvantages of Kodak Photo CD

The main disadvantages of using Kodak Photo CD are:

- Takes a "long" time to process, which is about one week.
- Can not see what the photograph looks like before it is taken; the quality of the photograph may be poor due to lighting, taking unsuitable photograph etc. Therefore, the photograph may have to be taken again. This will result in a further week's wait before the CD comes back, even then the problem may still exist.
- Currently there is a minimum order of 10 images which if only one image is required this is expensive.

1.3.7.f) Scanning

Scanners can be hand-held or flatbed, grey-scale (64 or 256 shades of grey or pure black and white) or colour (8-bit to 24-bit colour) and have a range of different resolutions.

1.3.7.g) Scanning using GrayArtist GS-800 Hand-Held Scanner

This scanner has three scan modes:

1. True 256 grey-scale supports a maximum of 256 true grey-scales and a resolution of 100 - 400 dpi (400 dpi being the best resolution). (Colour images can be scanned but will be grey scale!) ;
2. Halftone black and white, resolution 100 - 800 dpi ;
3. Line-art/Text black and white, resolution 100 - 800 dpi.

For the purposes of obtaining high quality continuous tone images, for example, scanning photographs the only suitable mode is in true 256 grey scale.

The scanner has a Brightness Control with which it is possible to adjust the overall whiteness of the scanned image.

1.3.7.h) Advantages of Hand-Held Scanning

The advantages of scanning are as follows:

- Scanning an image is quick it can be done immediately in real time;
- The image can be previewed on the screen before it is scanned in therefore ensuring the digitised image is correct before it is scanned.

1.3.7.i) Disadvantages of Hand-Held Scanning

The disadvantages of scanning are as follows:

- In order to scan an image, a photograph or photocopy of the image must exist in the first instance. Note, if the image is a photograph it has to be taken and developed before the image can be scanned therefore increasing development time and expense;

- Scanning requires large amount of free memory (e.g. 3.9 x 4.0 inch, 400 dpi, 256 grey-scale scanned image needs 2.5 Mb);
- The image to be scanned needs to be the same size as the scanner's width in this case 4.1 inches. Using a merge facility can scan images with a width greater than the scanner's width i.e. 4.1 inches. However, the seam where the two photos are joined does not always produce satisfactory results. However, by taking a photocopy (if the image is a text book page) and then reducing the photocopy in size, one can produce an original image that is within the scanner's maximum width;
- The scanner may move due to a "shaky hand" and as a result the image may be distorted. It is recommended where possible to use a guide with hand-held scanning;
- The scanner needs to be moved in a slow continuous manner. Judging the correct speed to move the scanner smoothly is difficult. If the scanner is moved suddenly the scanning is stopped.
- When scanning a page of a book the scanned 256 grey image needs to be edited to remove all the excess grey tones. It was found that the white page background becomes a shade of grey, editing was necessary to change the grey to white.

1.3.7.j) Video Captured Images

In terms of hardware a video capture board to capture a video image is required. The camcorder plugs directly into this. The image to be captured can be viewed in real time on the VDU display without having to save to the hard disk. When the image content is satisfactory the image can then be saved to file.

1.3.7.k) Advantages of Video Capture

The advantage of video captured images is that the image is captured in real time therefore there is no waiting time for the image to be developed.

1.3.7.l) Disadvantages of Video Capture

The disadvantages of video-captured images are as follows:

- If the subject is not near the computer, it is not known what the resulting image will look like on the VDU display because it is out of vision. The image may have to be retaken again.
- This method produces the poorest image quality of the three methods described.

1.3.7.m) Editing Photographic Images

Once the image is saved to disk it can be edited in terms of being, for example, resized, cropped, altered at pixel level, or the hue, saturation brightness, contrast etc may be changed. The extent to which the image is edited will depend on how much the image needs altering and if the time to do it is justified. Editing an image at pixel level can take a large amount of time and should really be done by a graphics designer especially in cases where the actual content of the image is to be altered. However, for simple edits such as resizing, cropping, adjusting colour, brightness, contrast parameters this is both quick and easy to do using the right package such as PhotoStyler.

1.3.8 Video

1.3.8.a) Video Capture

In order to capture images through a video recorder a video capture board such as Video Blaster, Video Spigot which are in the £250 region (1995 prices) is required.

Note! Video Blaster only supports 2 M colours but it was found that only the 256 colour driver was compatible within Windows. This meant that the 16.7 M colour driver could not be used to show true colour photographs. Video Spigot supports 16.7 M colours.

The quality of the video camera affects the resolution of the image. Ideally one should use a Hi8 or S-VHS camcorder with a CCD (charge-coupled device) that produces a minimum resolution of 400 horizontal lines [71] in order to capture video images.

Still-image video cameras are:

> *"usually better than that of a scanned Polaroid photo but not up to the standard of Kodak's Photo CD"* [4].

To capture full-motion video running at 30 frames per second (American frame rate; UK 25 frames per second) over 100 MB per second of data would be saving to the hard disk drive! The result of this means there has to be a compromise in terms of reducing the size of the video from full screen to a window, the number of colours displayed (256) and the smooth running between frames. The resulting video quality is good but it tends to be jerky where the PC can not calculate the data fast enough and has to skip frames to catch up.

1.3.8.b) Advantages of Video Capture

The advantages of digitised video are as follows:

- The video can be edited once it is saved to hard disk.
- Screen capture is possible using video software. This would be useful for demonstration purposes where different areas of the screen could be highlighted.
- Auditory capture is also available at the same time as video capture.
- It is possible to capture single bitmaps.

1.3.8.c) Disadvantages of Video Capture

The disadvantages of digitised video are as follows:

- Digitised video consumes large amounts of hard disk - Video Blaster records in uncompressed format. Compression is available but it has to be done after the video has been saved.
- There is a high development time to write the video script and edit it.
- Good facilities are needed to make high quality video.
- The length of video captured is limited to the space available of hard disk.
- The video image is jerky - at this point in time the computer for this project is not fast enough to do all the calculations that are required to produce full motion video.
- The video size is restricted to small window image.
- The video colour is restricted to 256 colours. This is due to the lack of computing power.

1.3.8.d) Editing the Video

Video Blaster allows the audio to be separated from the video and saved separately if required. Editing includes adding special effects, cropping, rearranging, extracting, and inserting to mention a few.

1.3.9 Sound

Hardware requirements that are necessary to produce sound are a sound board, such as Sound Blaster, (this has to be 16 bit capable of playing CD-Audio quality; 8-bit sound is very poor; it is sounds broken up which is comparable to old radio transmissions), speakers or headphones and a good quality microphone. Sound can take up a lot of hard disk space depending on the quality of the sound. An investigation was undertaken to test the capabilities of the sound card. This investigation involved determination of the amount of disk space consumed versus the quality of sound. Table 1.3.1 below represents the findings of this investigation.

Table 1.3.1 Illustrates the effect of digital sound quality on the time available for recording under those conditions if the resulting file was 12.5 Mb in size.

Output	Quality	Sample Rate / kHz	Size of file	Time / second	Comment
Mono	8-bit	11	12.5 Mb	1141	Very poor not acceptable
ditto	ditto	22	ditto	570	ditto
ditto	ditto	44	ditto	285	ditto
ditto	16-bit	11	ditto	570	Very good acceptable
ditto	ditto	22	ditto	285	ditto
ditto	ditto	44	ditto	142	ditto
Stereo	8-bit	11	ditto	570	Very poor not acceptable
ditto	ditto	22	ditto	285	ditto
ditto	ditto	44	ditto	142	ditto
ditto	16-bit	11	ditto	285	Very good usable
ditto	ditto	22	ditto	142	ditto
ditto	ditto	44	ditto	71	CD-Audio Quality

It can be seen that as the quality increases there is a dramatic reduction in the time available for recording under those conditions. It can also be seen that recording in stereo reduces the time by half.

From this simple test it was concluded that 16 bit 11 kHz mono sound was more than adequate for reproducing narration.

1.3.9.a) Advantages of Using Sound

The advantage of using sound is that there is very high quality produced.

1.3.9.b) Disadvantages of Using Sound

The disadvantages of using sound are as follows:

- CD quality audio consumes lots of hard disk space.
- Can not have speakers in a room of PCs since this would result in a lot of noise that would be distracting. If headphones were used for group work then student-student interaction would be lost.

1.3.10 Voice Recognition

One of the areas identified and researched for the incorporation of sound within assessment is voice recognition. One of the key criticisms that many users have made of the programs that have been produced during this research was moving the mouse cursor from one side of the screen to the other when interacting with the program. One solution to this problem was to implement voice recognition to replace or act as an alternative to moving and clicking the mouse.

The software used to produce voice recognition was a commercial package - Voice Assist which is available from Creative Labs. This software functions by recording in a database the waveform of words spoken by the user. Each word (command) then has an associated action recorded i.e. what actually happens when the user speaks the word. When the user then says a word Voice Assist takes the waveform of the word and searches in the database for a match. If a match is found the software then invokes the action that is associated with the command word. Within the assessment packages command words that were used with Voice Assist were the Next command i.e. page turning command and OK command to confirm responses to questions. This reduced the amount of interaction that the user would have to do by the mouse.

1.3.10.a) Disadvantages of Voice Assist

The main disadvantage of Voice Assist is that the user has to say the word exactly as it is recorded in the database. If it is spoken differently no match occurs and the action is not invoked. In order to minimise this Voice Assist lets the user record the word five times so that there is a aggregate of recordings from which to choose.

1.3.11 Conclusion

Within this sub-section key terms within multimedia were defined and the techniques that are involved in producing multimedia were described. The advantages and disadvantages of these techniques were also described.

1.4 Introduction

It is the aim of this sub-section to provide the background theory of expert systems and to mention some examples of where these systems are already employed in industry. The main focus is on the components of these systems and techniques that are used to make these systems work. It is not the intention of this thesis to derive a new philosophy of this subject area or to extend knowledge to advance these systems. It is the intention however, to apply that what is already known about these systems to the area of computerised assessment in a tertiary degree course. The main objective is to extend the question types available in computerised assessment beyond the multiple-choice question.

This subject area will be the major focus of this thesis. Therefore, the inclusion of this sub-section is vital to the understanding and thrust of the latter part of this work described in this thesis.

1.4.1 Applications of Expert Systems

Expert systems are widely used in industry. The knowledge that the system uses is both static and dynamic. For example, a dynamic system must be able to react immediately to changes in temperature that may ultimately lead to a disaster if not acted on instantly. Despite the obvious advantages of using an expert system it does depend on having an expert available to produce the system in the first instance. However, another problem may be that it is not possible to code the expert knowledge into the system. In these instances expert systems can not be implemented.

When considering designing an expert system the first step is to determine if the knowledge can be coded in the first instance. In science there is enormous potential for expert systems. However the system must have the following human traits. It must

> *"be able to cope with error-loaded or incomplete data, do everything a human expert can do in a limited domain and be able to explain its reasoning".* [77]

"Shells" are commercially available which provide an empty knowledge base, inference engine and a user interface. These terms will be described further on in this section. This makes the development of an expert system much easier. The user only has to concentrate on the inputting of the assertions into the knowledge base. Expert systems are currently used in science in the following areas: molecular modelling, chemical kinetics, analytical methods development, organic synthesis, imaging, diagnostic systems etc. Within chemistry they are widely used in the control of equipment and robotic samplers and acting as advisors in analytical procedures.

An example of the functions of an analytical system include:

> *"advise the user on methods of analysis for different samples;*
>
> *interpret spectra, chromatograms, electrochemical measurements and other data;*
>
> *instruct novice users of the system;*
>
> *monitor equipment performance;*
>
> *report defects and offer help when errors occur;*
>
> *schedule work and instruments if samples must be processed with different priorities;*
>
> *control instruments, robotic samplers or process lines."* [77]

The next section defines what the aim of an expert system is and the components that make it up.

1.4.2 Definition of an Expert System

An expert system (also known as a knowledge-based system) is a branch of artificial intelligence.

"Artificial Intelligence is an attempt to reproduce intelligent reasoning using machines." [77]

It is a program that attempts to solve problems just as a human expert would do. An expert system consists of three elements: a knowledge base; inference engine and a user interface [74, 77]. In the following sub-sections each of these elements will be described.

1.4.3 Knowledge Base

This is where all of the facts, rules, relationships and statements of probability are stored based on a specific domain of knowledge. The facts can be retrieved from any source such as textbooks or acquired from human experts who use "rules of thumb" [74].

Information can be stored in the knowledge base in a list form where data is unordered. If the database is large this can be a very hard structure to query; the time for the search could be long if the appropriate data is stored near the end of the database. It is therefore necessary to store the information in a logical manner to speed up searches.

A tree diagram is one such method for storing information that is inter-related.

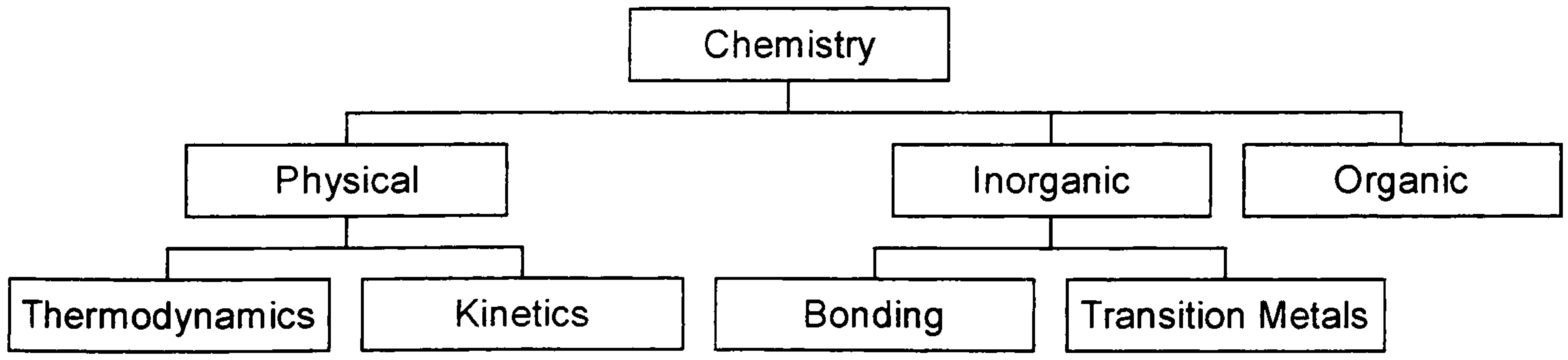


Figure 1.4.1 A Tree Diagram

This is a very simple example. The elements of such a structure are as follows: the top of the tree is called the root; the connections are branches and all extremities other than the root are called leaves. Each sub-category inherits the attributes of the categories above it. There are other types of structures for organising data.

Individual pieces of information are known as assertions. It is these assertions that the expert system uses to reason. The knowledge base also contains prescriptions or rules that instruct the inference engine on how to combine the assertions.
The next sub-section describes the second element of an expert system, which is the inference engine.

1.4.4 Inference Engine

This is the part of the system that uses rules to manipulate or draw conclusions based on the knowledge from the knowledge base. It looks in the knowledge base for the assertions and rules that are relevant in answering the questions of the user.

Rules for the combination of assertions include predicate logic, structured objects and production rules [74, 77]. Production rules are the most widely used within chemistry. They take the following format [77]:

IF assertion1 is true AND assertion2 is true AND .....

.....THEN draw conclusion1 AND draw conclusion2 AND .....

In order to help the inference engine locate the appropriate data, a decision tree or an AND / OR tree is used. The root defines a goal that will be met depending on the combination of sub-goals. Branches join the root and sub-goals.

The decision tree is used to delineate a problem and represent the production rules in a logical manner.

If B and C true, or D true
then conclude A

If E or F true,
then conclude B
If G true
then conclude C
If H and I and J true
then conclude D

If K and L true
then conclude G

If M or N true
then conclude K

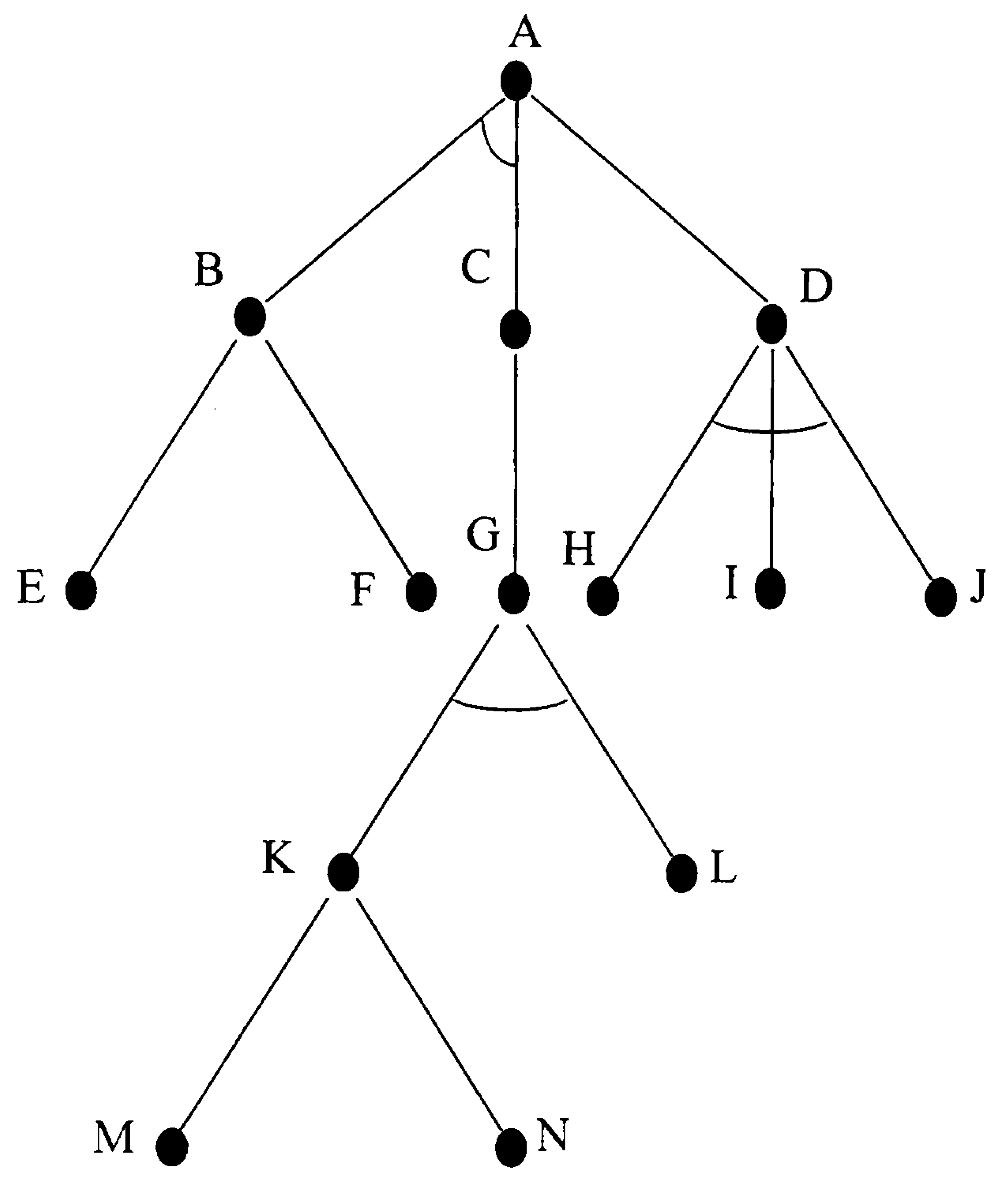


Figure 1.4.2 Production Rules

Figure 1.4.3 AND / OR Branches

The inference engine uses this representation within the knowledge base to prove its goal. Note B, C and D are not leaves.

Two methods of analysis that can be used for the inference engine to prove its goal are breadth-first searching and depth-first searching [74, 77].

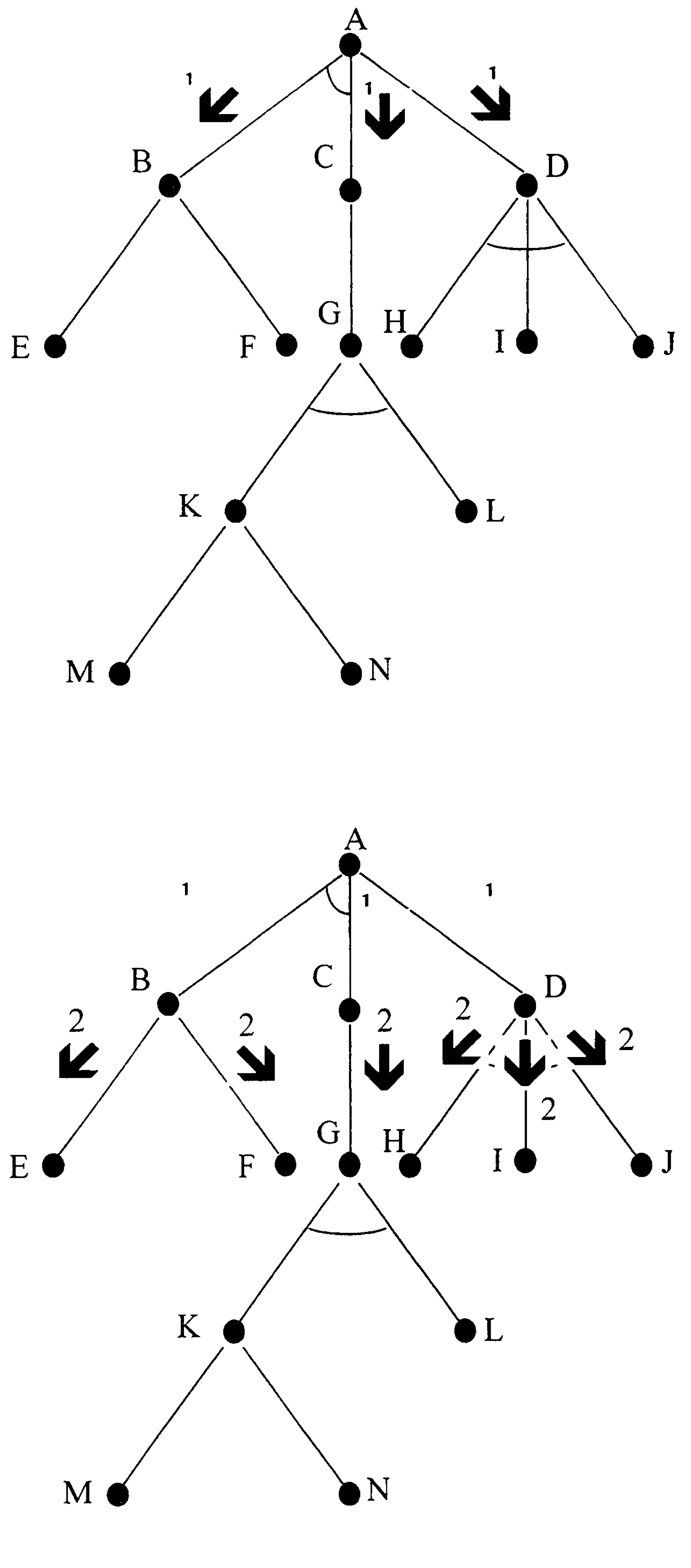


Figure 1.4.4 Breadth-First Search

1.4.4.a) Breadth-First Searching

In breadth-first searching all of the sub-goals or leaves in any one given level are searched at the same time until the goal is proved to be true i.e. once it finds a leaf or until all the branches have been searched until a leaf is found. This method finds the shortest path for the solution. However, it is not necessarily any faster than other search methods because if the solution is at the end of a branch of a very large tree and the search is involves searching every branch this will be very time-consuming task.

1.4.4.b) Depth-First Searching

In depth-first searching a branch a explicitly searched until a leaf is reached. If the leaf does not provide a solution to the problem the inference engine searches through the neighbouring branch examining the data in its leaves. This process is repeated until the goal is proven or until the entire tree has been searched. Throughout the knowledge base tips can be contained which the inference engine can use to decide which branches appear most promising than others.

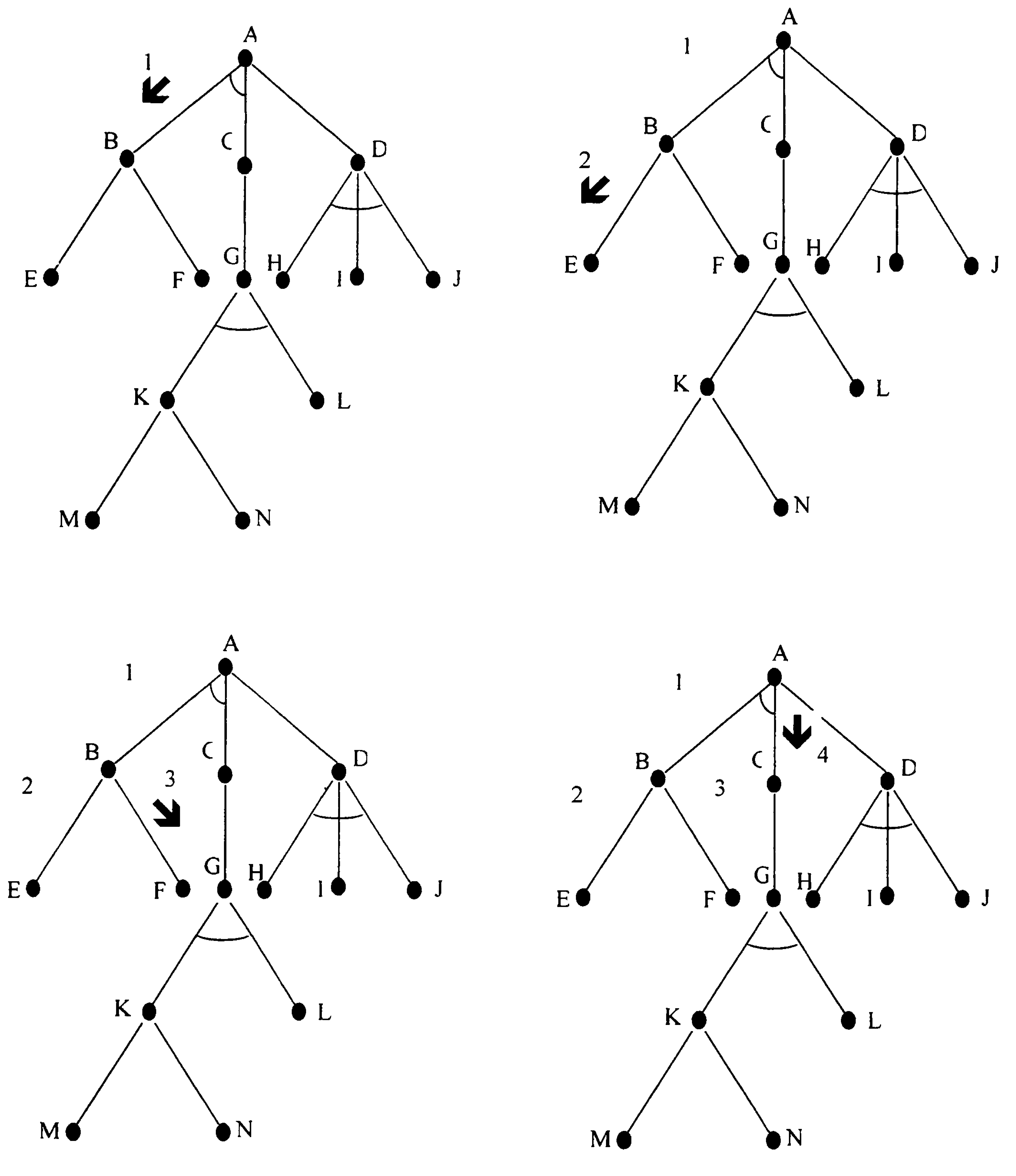


Figure 1.4.5 Depth-First Searching

For both search methods if all the possible solutions need to be known then the system must be instructed to do so otherwise the search will terminate once the first solution has been found using either method.

1.4.4.c) Forward-Chaining

Forward-chaining is used when we need the system to generate new assertions. This is achieved by asking the user for new information. The rules that are contained within the inference engine uses this new information and information within the knowledge base to produce the new assertion or conclusion.

1.4.4.d) Rule-Valued Searching

A more efficient method for searching than the previous two mentioned is rule-valued searching. This method tries to hone in on the best option / path to take. This then reduces the number of possible searches.

The next section describes the final element of an expert system, which is the user interface.

1.4.5 User Interface

This is the interface between the computer and the user. The interface is used to provide feedback to the user. The feedback is more than a suggested answer to the user's question. It is also necessary for the expert system to explain why it has chosen the response that it has in order for the user to gain confidence in the system's responses. The system must also be able to communicate to the user at a level they will understand. The expert system may contain information that is beyond the understanding of the user.

The next section describes a method called heuristic that is coded into the knowledge base to provide the system with expert knowledge in general terms when an answer is not absolute. Heuristics provide the system with a probability that an answer is correct.

1.4.6 Heuristics

Information is not always 100% true. It may be a hunch. A person may think something is true but is not certain. Within an expert system these "rules of thumb" are very important. However, they will only be useful if they can be coded into the knowledge base. Heuristics is the method by which these rules of thumb are coded into the knowledge base.

> *"A heuristic is a rule of thumb which gives guidance about how to solve a problem."* [77]

An example of a rule of thumb would be

> *"Most drugs are of low molecular weight."* [77]

This information could be used if given a compound of high molecular weight and a compound of low molecular weight in determining which would probably be a drug. Heuristics can be coded into the knowledge base as assertions or production rules and usually contain a probability weighting. For example [77],

DRUG=[MW<500,0.8]

IF product is black AND the product is gooey AND the product was synthesised by a student THEN the student will not make a good organic chemist (probability 0.95)

When the inference engine combines production rules and heuristics it does so by drawing a conclusion with a certainty value. They take the following format:

IF condition1 is true (with probability1) AND condition2 is true (with probability2) AND.....THEN draw conclusion1 (with certainty1) AND draw conclusion2 (with certainty2) AND....

The degree of confidence of each conclusion depends on:

- each conditions probability;

- each conclusions certainty if all conditions are true.

It is common practice to set the overall probability if all conditions are true equal to the probability of the condition that is least likely to be true. Despite this procedure not being statistically accurate many expert systems use the *weakest link rule* because there is insufficient information to indicate that the two conditions are statistically linked and the use of statistical methods has little effect on the conclusions made by the system. Since this addition calculation slows the system down but has little benefit it is normally left out.

> *"The power and the utility of an expert system is determined primarily by the knowledge that it contains, rather than by the sophisticated statistical methods that it uses to combine heuristics."* [77]

The outcome of the above statement is that expert systems often rank hypotheses. Because a question may have several answers to a question it is important that the system be able to deal with this variable and select the best answer. It is not important what the certainty of the conclusion is but rather which is the most likely conclusion based on the ranking order.

1.4.7 Conclusion

This section is included to provide some background theory of expert systems. It is intended to utilise expert systems to extend computerised assessment beyond the multiple-choice format to enable computerised assessment to assess process skills. Therefore, it is considered that background knowledge within this subject area is beneficial in understanding the core areas of this thesis.

1.5 Conclusion

It is the intention of this section to have:

- described key terms commonly used in assessment and the various types of categories in which assessment methods;
- reviewed current issues and trends within assessment;
- described the reasons for the development of computerised assessment;
- described computerised assessment methods that have been developed up till the start of this research;
- reviewed current issues and trends within computerised assessment;
- described why computerised assessment is not yet a widely accepted assessment technique within the academic community and therefore why there is a need for this research;
- described what multimedia is and some of the techniques involved to produce multimedia;
- briefly described the background theory of expert systems.

The following conclusions can be made after this literature survey:

- the main computerised assessment technique is the multiple-choice format;
- computerised assessment assesses the product, i.e. the final result, but not the process skills i.e. the individual steps that are taken to produce the final answer to a problem;
- expert systems offer the potential of extending computerised assessment beyond the multiple-choice format by acting as an intelligent tutor to assess student process skills as well as their product skills.

<u>Section Two</u>

Focal Theory

2. Section Overview

The previous section focused on the work that has been already undertaken in the areas of assessment, computerised assessment, multimedia and expert systems. It is the intention of this section to describe the aim of the research project and the methodology that would advance knowledge beyond that what is already known. This section will describe how these four areas will be integrated to achieve this goal.

2.1 Aim of Research

The main conclusion from the literature survey was that computerised assessment had not advanced much beyond the multiple-choice question. It was the overall aim of this research project to investigate and evaluate how computerised assessment could be extended with the then advances in technology such as the phenomenon of multimedia, the increased memory and storage capacity, and the 486 chip. This overall aim, as the literature survey progressed, was soon revised to *how computerised assessment could be extended beyond the multiple-choice question.*

The objectives of this research were to:

- undertake an extensive literature survey to determine the work that has been undertaken previously;
- analyse a current tertiary chemistry degree with the objectives of determining current learning strategies and assessment methods, and analyse the assessment methods with respect to whether or not these methods could be computerised and the advantages and disadvantages of doing so. This would provide the total number of assessment methods that could be potentially computerised within this course;
- investigate, evaluate and describe new methods of computerised assessment. This evaluation will describe the advantages and disadvantages of these new methods;
- produce a set of guidelines on how to determine *where, when and how* to develop computerised assessment.

2.2 Research Methodology

The research methodology would involve:

- an extensive and on-going literature survey;
- elucidation of expert knowledge with experts in a number of specialist fields - computer authoring languages, multimedia and education through individual meetings or conferences;
- collecting data for analysis through self-designed questionnaires;
- collecting data to provide evidence to support this thesis through the design and development of new and novel computerised assessment methods. The very fact that it is possible to produce functioning prototype computerised assessment methods that extend beyond the multiple-choice format will conclusively prove the hypothesis that computerised assessment can be extended beyond the multiple-choice question.

2.3 Introduction

It is the aim of this sub-section to compile the learning strategies and the means by which the learning outcomes are assessed within a tertiary level chemistry course. The assessment methods will be analysed in terms of whether they the can be computerised or not. This will provide a figure as to the extent that assessment methods within chemistry could be computerised. This will provide some evidence as to the justification of the development of computerised assessments in helping to reduce staff workloads or whether the development is really only an academic exercise with no real practical value.

2.3.1 Learning Strategies

The following learning strategies have been compiled from a typical chemistry degree course from a tertiary UK education establishment.

Table 2.3.1 Typical Learning Strategies in a Chemistry Degree Course

Learning Strategy	Description
Lectures	These are used as the primary method for the dissemination of information. These are used to illustrate the key principles and concepts to the student. Video, photographs, demonstrations and exhibits can used to enhance the learning experience.
Tutorials	These are designed to reinforce the lecture material by presenting the student with problems.
Seminars	These are group orientated and student led. They are intended to develop the student's oral skills e.g. ability to defend proposals. In some cases the oral may be videoed to help the student to analysis their performance.
Directed Self Study	These are in the form of case studies, guided reading etc. that are student-centred activities.
Library-Based Assignments	These are student-centred activities in which the student is required to go to the resource centre to gain hands-on experience retrieving information from various sources.
Workshops	These provide the student with hands on experience.
Computer-Aided Learning	These are designed to support lecture material or they can be used as stand alone study packages.
Practical work	These are vital in developing the students' ability in laboratory techniques.

2.3.2 Current Assessment Methods

The following assessment methods have been compiled from a typical chemistry degree course from a tertiary UK educational establishment. These assessment methods are used to assess the student learning outcomes. These assessment methods are as follows (a description will be given of each in the next sections):

1. Progress tests
2. Assignments (student-centred)
3. End test / Examination
4. Case study (written and oral)
5. Short diagnostic tutorial questions (10 minutes)
6. Laboratory work (individual and group)
7. Oral presentations (individual and group) tutor, self and peer group assessment
8. Time-constrained essay
9. Project (individual and group) report and oral, continual assessment
10. Open book test
11. Time-constrained problem solving (individualised)
12. Personal Interviews
13. Individualised information retrieval exercises
14. Seminar performance

2.3.3 Analysis of Current Assessment Methods

Within this sub-section it is intended to analyse the assessment methods that were listed in the previous section. Each method will be analysed in terms of:

- what the format is i.e. the question type;
- what type of knowledge does the assessment test for. A project called UNICAL, during the 1970's conducted an investigation into computer aided learning. This project identified five types of knowledge which are as follows[78]:

    - **Type A interactions**: recognition (e.g. multiple-choice questions);
    - **Type B interactions**: recall (e.g. fill in the missing word);
    - **Type C interactions**: reconstruction, understanding or comprehension (e.g. rewriting things in your own words);

  - **Type D interactions**: global reconstructions or intuitive understanding (e.g. problem solving);
  - **Type E interactions**: constructive understanding (e.g. student draws on previous experience to create something new).

  This study concluded that the aim of teaching is to get students producing constructive understanding i.e. Type E interactions. The assessment methods will be analysed in terms of these five types of knowledge.

- in what mode of assessment is the method used on the course i.e. is it formative or summative assessment;
- whether or not the assessment method can be computerised and why this is the case.

Table 2.3.2 Analysis of Current Assessment Methods

Assessment Method	Format	Tests for	Type	Potential Comptn
Progress tests	MCQ, Problem Solving	Recognition, Reconstruction, Global Reconstructions	Formative	Yes
Assignments	Literature, Further Reading, Producing Reports	Integration of Knowledge to produce report and solve problems	Summative	Partially
End test / Examination	Short, Long Essay Questions	Knowledge (Recall), understanding and application of concepts and principles to new situations, written communication skills	Summative	Partially
Case study (written and oral)	Involves literature work, further reading, gathering information, production of coherent report with a verbal presentation and defence of report	Integration of knowledge to produce coherent report and solve problems, communication skills written and verbal	Summative	No
Short diagnostic tutorial questions	Multiple-choice, Short Problems	Recognition, Reconstruction, Global Reconstructions	Formative	Yes
Laboratory work	Laboratory based sessions	Writing reports, Observational skills, Organisational skills, Integration of knowledge, Hands-on experience, Taking notes	Summative Formative	Partially
Oral presentations	Oral presentation	Verbal communication skills	Summative	No
Time-constrained essay	Ability to integrate knowledge, and present it in logical manor	Reconstruction	Summative	No
Project (individual and group) report and oral, continual assessment	Laboratory based sessions	Reports, Observational skills, Organisational skills, Integration of knowledge, Hands on, Writing, Communication within groups, Trying for Constructive Understanding	Summative Formative	No
Open book test	Multiple-choice, Problem Solving, Essay Type	Recognition, Reconstruction, Global Reconstructions	Summative	Partially
Time-constrained problem solving	MCQ, Problem Solving Exercises	Global Reconstructions	Summative	Yes
Personal Interviews	Oral	Verbal Communication Skills	Summative	No
Individualised information retrieval exercises	Go to library, find text and fill in short answers	Ability to retrieve Chemical Information	Summative Formative	Partially
Seminar performance	Student Lectures	Oral Communication and ability to understand, comprehend, present	Summative Formative	No

From the table it can be seen there are three options on whether an assessment method can be computerised or not: i.e. Yes, No or Partially. The reasons for these three options can be summarised as follows.

1. Yes: For assessments methods currently used within the course, those that are objective in nature will be computerisable. This is because it is possible to provide an absolute answer that the computer can compare to the student's answer. Typically mathematical problems, multiple-choice questions etc. fall into this category.

2. No: For those methods that can not be computerised. These are generally subjective in nature. In these cases there is no absolute answer to which the computer can compare. The marks that a lecturer will award are based on certain criteria and being able to use other senses, for example, being able to listen to an oral, which is not possible for a computer to do yet!

3. Partially: This option exists for those methods that will vary from subject to subject and will depend on the assessment in hand. For those that are objective then they will be potentially computerisable. For those that are subjective then they will not be computerisable.

2.3.4 Conclusion

Of the assessment methods listed above 1, 2, 3, 5, 6, 10, 11, 13 could be computerised to some extent. This means that up to 57% of the course assessment methods could be computerised to some extent! Note this does not refer to the total number of assessments a student may have to do on a course. This is a percentage based on the type of assessment methods and not the number of assessments.

It can be seen that there is a general trend of the suggested types of areas in which computerisation is a possibility. For multiple-choice, filling in blanks, numerical problem solving etc. question types computerisation is no real problem. However, whenever it comes to oral or essay-based assessment the computer fairs very poorly having no ability to comprehend language at either the oral or written levels.

From this simple analysis there is justification for the computerisation of assessment methods on a large scale using techniques that are a straightforward conversion from paper-based assessment methods to computerised assessment methods. But there is also justification for further research into new objective computerised assessment methods such as problem solving of a mathematical nature. This is because each stage in the solution to a problem of this nature has an absolute value that should be quantifiable which the computer can measure.

In the next section 2.4 *Analysis of Multimedia in Chemistry* an analysis will be performed with respect to whether or not chemistry is suitable for multimedia in the first instance. This will highlight specific areas that will be more suited for multimedia purposes and hence development of computerised assessment methods with multimedia.

2.4 Analysis of Multimedia in Chemistry

It is the aim of this sub-section to determine if multimedia is suitable for chemistry in general terms and more specifically for assessment purposes. The aim of this analysis is to identify those areas that would be suitable for the further investigation for the development of new and novel computerised assessment techniques.

As previously described in 1.3.2 *Definition of Multimedia*, multimedia is the combination of text with one or more of the following elements: graphics, photographs, video, sound, animation. In order to see how multimedia could be applied to chemistry one first must analyse what a degree course in chemistry actually consists of. Appendix I *Learning Outcomes* describes the learning outcomes of a typical degree chemistry course.

From Appendix I it can be seen that the chemistry course falls into one of two distinct categories, theoretical or practical chemistry. These categories exist for the subject areas of physical, inorganic, organic and analytical. (For the purposes of this discussion we will only discuss chemistry in terms of these four subject areas.)

From these learning outcomes it can be derived that theoretical chemistry consists of concepts and theories of atomic structure, bonding, mathematical derivations, manipulation of chemical equations, reaction mechanisms etc. In terms of representing these concepts with the elements of multimedia it can be concluded that theoretical chemistry is very limited in its use through photographs or video or sound. Theoretical chemistry uses text, graphics and animations to communicate its concepts etc. Therefore, computerised assessment of such learning outcomes using multimedia will be limited.

As can be seen from the learning outcomes for practical chemistry, this subject area consists of doing hands-on experiments, recording observations, writing reports etc. Practical chemistry is very much a real-life visual subject. This means that there is potential for using photographs and video in this subject area for assessment purposes.

In summary from this it can be concluded that:

- theoretical chemistry is more suited to the multimedia elements text, graphics, animations and sound;
- practical chemistry is more suited to the multimedia elements text, photographs, video and sound.

These very broad generalisations were derived from the fact that:

- theoretical chemistry is concerned with electrons, atoms, molecules, mathematical derivations and manipulations of chemical equations etc., theories, reaction mechanisms, synthesis etc. which are very difficult to illustrate and interact with through photographs and video. These are better done through graphical, textual and animated means. The lecturer could include narration if thought to be appropriate.
- practical chemistry is concerned with assembling of apparatus, carrying out experiments, recording observations i.e. very real objects that we can actually see. These can be more easily presented through video / audio demonstrations and photographs rather than animations or graphics.

From this simple analysis it can be concluded that depending on the subject matter, the suitability and combination of multimedia elements that are used for the development of new assessments will vary and that not all of the multimedia elements will be necessarily used all of the time.

For the purposes of this research, the primary subject area that will be focused on for developing assessment methods with multimedia will be practical chemistry for the reasons given above.

2.5 Discussion

Computerised assessment has real implications for a chemistry degree course curriculum. Up to 57 % of the assessment methods used could be potentially computerised. However, one of the main problems why computerised assessment has not had widespread integration within courses is because of the limited question types i.e. the multiple-choice format. Critics of this question type perceive that it only tests for recognition and does not let the student explain how they arrived at their final answer. It is the main aim of this research to investigate and evaluate how computerised assessment can go beyond the multiple-choice question type, if at all!

From the analysis in the previous section the application of multimedia within chemistry is restricted to certain areas and can not, realistically, be applied to every assessment method. From this conclusion it was decided to investigate the development of multimedia assessments within practical chemistry and chemical information and retrieval.

Early in this research it was concluded that multimedia would have a "limited" role to play (in that multimedia can not be used in every assessment situation) in the development of computerised assessment beyond the multiple-choice question. However, there are situations in which it could be used and therefore these would be investigated.

As can be seen from a study of the Appendix I *Learning Outcomes* much of a typical chemistry course is theoretical. It has been concluded in, 2.4 *Analysis of Multimedia in Chemistry*, that for chemical and mathematical theoretical problems multimedia does not have much of a role to play i.e. it is not the primary method of communicating knowledge within this subject area. Instead these are represented by formulae and diagrams for example. Within problems of this nature there are definite absolute solutions to the problem. The answer to each step is either correct or incorrect i.e. it is objective in nature. The solution follows predefined rules that dictate how the problem should be solved (of course there may be many correct solutions to a problem). As stated in 2.3.3 *Analysis of Current Assessment Methods* computerised assessment is adept at assessing objective tests. It is proposed as part of this research to investigate how computerised assessment can be extended to include marking each step in a solution to a problem. This would mean the computer would in effect become an *electronic notepad* and an *on-line intelligent tutor* capable of marking a student's solution to a problem and providing instant feedback. This part of the research is in the realms of expert systems, which explains why the background knowledge of this subject area is included.

For the purposes of this research it is not intended to push back the frontiers of multimedia or expert systems. This research takes knowledge from these subject areas to be applied to the subject area of computerised assessment. This is the extension of knowledge on which this research will focus.

It should be stressed there is a differentiation between the purpose of multimedia and expert systems to the application of computerised assessment. An expert system is an engine that can include multimedia. However, multimedia applications can not run without an engine. Multimedia is a form of communication that can be included in an application whereas an expert system is an engine that can be used to develop an application. It can be concluded therefore multimedia will not extend the question type in computerised assessment but does extend the media by which the computer can communicate.

Section Three

Data Theory

3. Section Overview

This section describes the work that was under taken during this project. It is split into six sub-sections. A brief description of each sub-section follows:

- section 3.1 provides details of the how the choice of authoring software was made;
- section 3.2 provides an introduction to the software that was developed;
- section 3.3 provides the details as to the work that was involved during the development of the Chemical Information and Retrieval package;
- section 3.4 provides the details as to the work that was involved during the development of the Identification and Usage of Apparatus package;
- section 3.5 provides the details as to the work that was involved during the development of the Determination of the Enthalpy of Reaction package;
- section 3.6 provides the details of a package that could be developed for assessing Reaction Mechanisms.

3.1 Introduction

It is the aim of this sub-section to describe how and why the authoring tool was chosen for this research. The section describes the categories of authoring tools that are available for development; the advantages and disadvantages of each category; the authoring tool that was chosen and why this authoring tool was chosen.

3.1.1 Software Development Tools for Assessment

One of the first decisions that had to be made was "which authoring tool should be chosen for this research?" After a review of the literature [79] the following conclusion was made; an assessment application can be developed in one of in two ways:

1. using a programming language application to write the software;

2. using an application specifically tailored for the production of assessments.

These two categories will be described in the following sub-sections.

3.1.2 Programming Language Applications

Programming language applications can be further sub-divided into three further sub-classifications.

3.1.2.a) Low Level Programming Applications

These programs include those that are pure low level programming languages such as, C, in which all the elements of the application are programmed from scratch.

Advantages

The main advantage of this type of application is that it is very flexible and powerful: any type of application can be developed from these types of development tools.

Disadvantages

There are disadvantages to low level languages. These are as follows:

- Initially there is a high learning curve to learn the language.
- The time required to develop the application is large since everything has to be programmed from scratch.

3.1.2.b) High Level Programming Applications

These applications do not require any previous knowledge of a programming language. The user is provided with a set of tools with which they "tell" the computer what they want and the computer fills in the values. Applications of this nature include Authorware Professional, Icon Author etc.

Advantages

The advantages of this type of application over low level applications are that:

- There is no need to learn a programming language. This will reduce the initial high learning curve.
- The development time of producing a program is relatively quick because many of the core functionality tools have been previously developed. The developer has only to concentrate on using these tools without having to develop them.

Disadvantages

However, there are still some disadvantages of using these applications for development. These are as follows:

- The developer still has to learn the fundamentals of the application and all the tools that it uses in order to use the package. Therefore, time is still required to learn the package.
- These types of packages reduce flexibility; the developer is restricted to the tools provided by the application. There may be something that the developer wants to do but can not because it is not possible to develop it in the application. This is because there is no functionality in the application to develop tools that the application has not provided.

3.1.2.c) Medium Level Programming Applications

This type of application falls in between that of high and low level applications. The developer still has to learn the application's language but there are tools for developing the interface very quickly. Therefore time is saved programming commonly used screen elements such as buttons, list boxes etc. Examples of this type of application include Visual Basic, ToolBook etc.

Advantages

The advantages of this type of package are as follows:

- There is a reduction in programming time compared to the low-level applications: this is because tools for development of interface etc. are built into the application therefore they do not need to be developed.
- These applications are both flexible and powerful: the programming language provides the developer to program to the required specification and hence is not limited to the restrictions of the application.

Disadvantages

Despite these advantages these packages do have the disadvantage in that a programming language still has to be learnt: this will mean there is a high learning curve initially.

3.1.3 Assessment Applications

These packages have been specifically designed for the production of assessment programs. These applications are basically 3.1.2b *High Level Programming Languages* that do not require the user to have any knowledge of low level programming languages. The advantages and disadvantages of this type of package are therefore very similar to 3.1.2b. Applications of this nature include Question Mark etc.

Advantages

The advantages of this type of application are as follows:

- Since these types of applications are specifically designed for assessment purposes there is a reduction in the development time for an assessment application. This is because many of the specifications for an assessment package are automatically built-in and therefore do not have to be developed by the developer.

Disadvantages

Despite, this type of application being specifically designed for assessment purposes one would imagine that this type of application would be the best method for developing assessment applications, there are still disadvantages associated with them. These are as follows:

- As with all packages there is still the initial investment of time required learning the application.
- These applications suffer from limited flexibility. There may be something that is required that the application does not have built-in.
- In some instances because the application has been developed by a third party a runtime fee may be required that is dependent on the number of machines that the application is to run on.

Despite this type of application being designed for assessment purposes it was not the one chosen for this research. This is because these types of assessment application offer assessment methods that are well known i.e. the multiple-choice format. It is the intention to go beyond this type of question. As a result these assessment methods do not offer the opportunity to develop new and novel assessment techniques. Therefore a lower level application would have to be chosen that would enable the development of new and novel assessment methods.

The next sub-section describes which authoring software was chosen and why.

3.1.4 Selection of Software Development Tool

It was decided to use Visual Basic Professional 3.0. The reason for this choice was that Visual Basic:

- offered the advantages of both the high and low level languages. It is easy to use [1]. The application interface can be quickly developed without the need for code through VB's built in tools. Simple applications can be produced without having to write any code.
- was very flexible and extendible[1]; the assessment application could be designed to meet the specific needs of the project. It empowers you to program what you want. There is nothing more frustrating than using an authoring tool that does not let you do what you want to do. With VB the author is in control. You can define your own customised tools using a C compiler.
- created programs that are stand alone. VB creates .EXE files. This means that there is no extra licensing costs for runtime files when you want to distribute you program.
- skills can be transferred to other applications. Excel 5.0 now uses VB as a macro language. Learning VB could be time saving in the use of other Microsoft applications.
- was cheap. With a Microsoft Select Agreement the software only costs £35 (before 12/94).
- was based on the BASIC language. BBC Basic had been used extensively throughout the department. The learning curve for the VB language was therefore very shallow since there is a very close correlation between it and BBC Basic.

However, despite these advantages of VB there are some disadvantages that are:

- In order to utilise VB's extendibility the developer needs to know C / C++ to develop .VBX files. If the author is not familiar with C / C++ the learning curve for learning C / C++ is very steep.
- There is no built-in hypertext.

- There is no built-in subscript or superscript facility that is very important in a subject like chemistry.

3.1.5 Visual Basic Professional 3.0 for Windows

Visual Basic is an object-orientated programming language used within the graphical user interface (GUI). Essentially visual objects, that ideally represent objects which the user is already familiar with such as, buttons, icons etc. are clicked using the mouse to produce an event such as turning a page. The aim of this interface is to reduce the initial steep learning curve of the user by providing the user with familiar objects with which they already know how to use. This is opposed to DOS based commands that require the user to learn many commands which can lead to information overload in the early stages of learning the language.

There are two steps that are necessary to produce any Visual Basic application:

1. Draw the interface using the tools provided;
2. Program the visual elements to respond to events generated by clicking the mouse for example.

Visual Basic comes with supplementary applications such as:

Help Compiler:

This enables the production of a comprehensive context sensitive help system that uses the Windows Help application.

Application Set-up Wizard:

This compiles all the application files that are required to run the program.

Database Manager:

Within the database manager databases can be created or opened. Visual Basic can also create and manage databases from within program code. Database options include photographic databases.

3.1.6 Conclusion

In order to develop new and novel assessment techniques a lower level application would be chosen. This is because it offers access to a language in which it was possible to develop most types of application. However, to save in programming time and to reduce the initial learning curve a medium level programming language was chosen - Visual Basic. The initial learning curve was significantly reduced due to prior knowledge of BASIC.

3.2 Introduction to the Software Developed

There were three prototype applications developed during this research. The very existence of these applications would support the hypothesis that computerised assessment can be extended beyond the multiple-choice question. These prototypes are described in the following sections:

- 3.3 *An Assessment Package to test for Student Knowledge on Chemical Information and Retrieval*;
- 3.4 *An Assessment Package to Test for the Identification and Usage of Apparatus*;
- 3.5 *An Assessment Package for the Determination of the Enthalpy of Reaction*;

The details of another application that could be developed are described in 3.6 *An Assessment Package for Assessing Reaction Mechanisms*. This details the methodology that would be followed to develop this application.

All of these subject areas have been investigated due to their potential suitability for computerised assessment as a result of the analysis carried out in Section Two. The first two prototypes, 3.3 *An Assessment Package to test for Student Knowledge on Chemical Information and Retrieval* and 3.4 *An Assessment Package to Test for the Identification and Usage of Apparatus*, are an investigation into the application of multimedia in computerised assessment. Note that the latter application is also really a forerunner of an expert system.

The sections 3.5 *An Assessment Package for the Determination of the Enthalpy of Reaction* and 3.6 *An Assessment Package for Assessing Reaction Mechanisms* are a study into theoretical problem solving within chemistry. As previously mentioned in Section Two subject areas of this nature do not lend themselves to multimedia and therefore it will not be investigated.

These applications are not of a generic nature due to the complexity of the applications and the time available to carry out the research. However, the general principle is that if these assessment applications can be developed for one specific example in these subject areas then many different problem solving exercises in these areas could be developed.

The interfaces for each of these applications were developed using the guidelines as presented in Appendix II *Interface Design Guidelines*. These guidelines are a result of a literature survey on the subject.

All these packages were designed to run on Windows 3.1.

The remainder of Section Three describes the work that was carried during this research and the results that were found.

3.3 An Assessment Package to Assess Student Knowledge on Chemical Information and Retrieval

3.3.1 Aims

The aims of this sub-project were:

- to produce a software package to assess student knowledge of chemical information and retrieval;
- to take photographs / obtain digital images of the chemical literature;
- to become familiar with the types of methods available to transfer photographs to the computer;
- to investigate the types of questions that could be used for assessment purposes using photographs;
- to investigate how the MPC can be used to assess student knowledge on chemical information and retrieval;
- to implement the software within the department;
- to evaluate the software in terms of student appraisal through a questionnaire;
- to investigate the advantages and disadvantages of this type of assessment.

3.3.2 Program Philosophy

The program is intended to be used for summative purposes after the theory has be taught through the traditional lecture.

The disadvantage of the older summative assessments is that the students have to wait for a week or more before they find out their results. Also feedback to each student is poor because of the increasing volume of students taking each test and the many other pressures on staff time. This results, in many cases, that the only feedback to the student is a percentage mark.

The advantage of computerised assessment is that each question can be marked immediately after the student answers the question. Also feedback can be given to the student immediately indicating to them why the question is correct or why they were wrong.

At the end of the test each student score can be recorded, even each student response can be recorded which will enable, if necessary, the analysis of the assessment.

The pressure on staff time on marking is thus reduced and the lecturer receives a percentage for each student. More importantly the student receives immediate feedback which will help to reduce any misconceptions of their understanding of a topic and help to provide the student with a deeper understanding of the topic. Assessment and learning go hand in hand.

From the staff point of view this software should be robust, and require no additional administration other than invigilators. Ideally the staff member wants to access a file to view / printout results and / or analyse the data.

3.3.3 Method

It was intended to provide the student with a virtual reference library of typical chemical literature. The student would be able to select a book, journal etc. On selection a photograph of the cover of the book would appear. This would enable students to see what actual reference material looks like before going to the resource centre. The student could also see the typical contents of the text to appreciate the style in which it was written - a journal detailing highly technical experiments is much different than, say, a textbook. A description of each of the categories in which chemical literature can be categorised is also included in the virtual reference library. In effect this assessment is an open-book assessment. The student is provided with reference material that they can view during the assessment.

The subject matter of this assessment was collated and designed in association with an experienced practitioner in this subject area. The photographs that would be central to this application were to be taken "in-house" i.e. resources in the department would be utilised to take the photographs. As part of this assessment it was the aim to investigate the methods of how to capture digital images. At the start of this research there was very little literature on the subject. The next sub-section details briefly the methods that were investigated to digitise photographs.

3.3.4 Evaluation

The evaluation of the software and student attitudes would be through a questionnaire. 27 B.Sc. 1 students did the assessment at the end of their first term in two groups. 24 questionnaires were completed and returned. No instruction was given beforehand on the use of the program.

3.3.5 Results

It is immediately obvious that computerised assessment can not give students vital hands on experience. There is no substitute for going to a library and doing a literature search. (However, it must be remembered that in the coming years chemical literature will become more and more electronically-based which evokes the question: will students need to do any paper-based searching? And if the literature is on-line should it not be possible to do hands on assessment of the digital literature hence increasing the scope of computerised assessment?)

However, what computerised assessment does assess is the students knowledge of the literature and their integration and application of that knowledge.

The principal types of usage of photographs in the package were:

- To illustrate what the chemical literature looked like. This included photographs of the bound work and a typical page.
- For informative purposes. The assessment was in the open-book format. Photographs and textual information were given illustrating typical chemical literature and literature categories. The student could use this information for help in answering questions but did not have to if they did not want to use it.
- To ask questions about a photograph using multiple-choice for selection purposes.

- To click on the photograph to identify the answer. This type of question was used with Chemical Abstracts to see if the students could identify the various parts of the abstract.
- Photographs were then used for feedback purposes to identify and explain the correct answer.

The next sub-section describes the problems that were encountered during the implementation of this assessment with the students.

3.3.6 Implementation

This assessment was carried out in the department's dedicated computer laboratory. There were 18 terminals of which 15 were used. The software was installed on the server. Each client was then able to gain access to the software via the network.

3.3.6.a) Implementation Problems

There is a number of important issues to be addressed with summative assessment that provides students with feedback during the test. These are as follows:

- Students will know the correct answer after the test. Some may be tempted to repeat the assessment and so restriction of student access to the assessment is of vital importance.
- What if the program crashes when the student is three quarters of the way through the assessment? Will they have to start again?

It is extremely important that a hard copy of each student response is made immediately after they have made the response. If the program just stores variables in memory and the computer crashes all the student responses would be lost. A hard copy was obtained by setting up a database file which would monitor student answers, time taken, date and the name of student.

However, obtaining a hard copy invoked another problem. In some institutions students are NOT allowed to write to the network even if it is not them but the program that is writing to the network. So, a hard copy would have to be obtained using a floppy disk. This would involve the lecturer manually having to put "n" floppies into drive A to compile one overall student database file in which to analyse their results. This defeats one of the advantages of computerisation that is having the computer automatically doing the work for the lecturer. The lecturer would also have to give the floppies to the students at the start of the assessment and collect the floppies at the end of the assessment.

This problem does have a positive side to it - increased security. The program was set up so that if the program does not find this database file it would not work.

Since the program is on the network it is accessible to students 24 hours a day. Ensuring they do not have access to the program at any other time other than when doing the assessment is of utmost importance to the success of summative computerised assessment.

The main issues to summative computerised assessment are:

- ensuring that the person who is supposed to do the assessment is doing it
- each student has only one attempt to do it.

Details of the measures of security implemented within this application will be described in the next sub-section.

3.3.7 Measures of security

There were three levels of security that were implemented in this assessment application. These were as follows:

1. user name;
2. password;
3. database file.

It was decided to see if students would actually cheat when doing a computerised assessment by leaving the system partially open to the potential abuser. A description of this loophole follows.

Access to the program required the student to type in their name and enter the password. There was no restriction to what was typed in for the user name, so it could be anything therefore allowing the student to access the program as many times as they wanted to. However, at the end of the test the student would hand in a print out containing their name and percentage scored, so they would have to do the test using their own name.

There was a time limit to the test - 45 minutes.

It was found that one student did actually cheat. This student initially entered their name using lower case letters and completed the assessment in 10 minutes, obtaining a score of 4%. The student then re-entered their name but in this case using upper case letters. Their score had risen to 75% but notice their psychology - they made sure they got a few wrong so as not to rouse any suspicions! Meanwhile they did not know all their actions were being recorded by the database that is used as a means to monitor if the system is being abused as well as to gather student results.

One way of solving this problem would be to set up a predefined list containing the names of those students who should be doing the test. When the student "logs on" they would select their name from the list. This would ensure that once logged on the assessment could be only entered once by that user.

However, this method would involve extra administration which is unpopular with staff i.e. having to change the list of examinees each year, and what if a student accidentally choose the wrong name or even deliberately selected it!

After this assessment it was decided to use each student's network user-name - this number is unique to each student and does not have to be entered by the student. In order to gain access to the program on the network each student has to log on i.e. type in their user name that is assigned to each student when they register at the university. The program automatically retrieves this username from the network. Once logged only one user can be logged on under that user name at any one time.

However, this user name is a number so a means has to be found of identifying the student without having to look up tables etc.

It was decided to ask each to type in their name.

This method is still open to abuse. If a student knows another student's user-name they could still do the assessment under that name and then repeat the test under their own name. In principle students are not supposed to know each other's user-name.

If there is still abuse of the system by looking in the database file you can determine the abuser since a record will be kept of users. An unknown name preceding a known student name should incriminate that student because they could cheat before they do the test for real!

This database file is encrypted so that users can not look in it.

The environment for summative computerised assessment is very important and needs to replicate the environment for the traditional examination. Summative computerised assessment does not mean students can do the assessment in their own time. Students will ask others what did they get for a question, copy each other etc., if left to their own devices.

Invigilators would still be required during the test. These are used in the traditional examination so there is no added complications.

One suggestion to solve the problem of students cheating would be assign each student to a specific PC. This is done in traditional examinations i.e. students are required to sit at certain desks. There is no reason why this should not be applied to computerised assessment.

If the computer crashes or someone tries to re-enter the program, the program is set up to take the student to the question after their last response therefore:

- allowing the student whose computer has crashed does not have to answer all the questions again;
- disallowing the student who is trying to cheat to repeat the questions.

3.3.8 Observations

This was the first time that this type of assessment had been run in the department. As a result it was done informally. General observations were:

- students had problems logging on;
- the printers in the computer room are shared but are not "smart" i.e. there is not a printing cue so you can not send anything to print until the printer has finished its current job, this caused the biggest problem;
- students would help each other;
- VDU screens were not set up as 256 colour drivers; photographs were very poor. There is a problem in the computer room there is a variety of machines that require different drivers. These would have to be loaded for every session because the network does not remember the setting and loads a default every time which is 16 colours - no good for multimedia. This meant that students did not see the real potential of computerised assessment - the displaying of colour photographs.

3.3.9 Questionnaires

Two questionnaires were given to the students at the end of the test. One was based on the student's general computing knowledge and the other was specifically aimed at the Chemical Information and Retrieval program.

3.3.9.a) Aims and Objectives of Questionnaires

The General Questionnaire was intended to investigate;

- if the sample suffered from computer anxiety;
- students' general knowledge of common elements of a personal computer;
- do students own a computer, have they used one before, where have they used it;

- have students used Windows, do they find it easy to use;
- do they find typing easy.

The Chemical Information and Retrieval Assessment Questionnaire was intended to investigate in general the samples attitudes to the software and this kind of assessment. See Appendices III and IV respectively for details on each questionnaire.

3.3.10 Student Results

Once the assessment was completed each student's percentage was collated. The following table shows the percentage bands and the number of students that fell in the appropriate band:

Table 3.1 The Distribution of Student Marks

Percentage Band	<40%	>40% <50%	>50% <60%	>60% <70%	>70%
Percent of Students	15	11	48	11	15

3.3.11 Discussion

From a sample of 24 B.Sc. 1 students who completed and returned the questionnaires 67% said that they suffer from computer anxiety. This experiment was deliberately set up to put the student in at the deep end to get an indication of their level of IT skills. In order for the success of computerised assessment it is essential that students are alleviated from this anxiety. It is of utmost importance that the student does not have to think about the tools they use, in this case the computer. They will have to use the computer as instinctively as they would a pen and paper. Time needs to be invested in the dissemination of IT skills to allow students to become confident users of technology. Surely this is only a question of familiarisation with the proper instruction and in time students should become confident users.

However, 83% of the sample did say that they preferred this method of testing despite their anxiety.

The advantage of using the computer is the ability to give instant feedback to the student. An overwhelming 96% of the sample did like to know if they get an answer correct or incorrect instantly and to have detailed feedback to explain why their answer is incorrect.

The most disappointing aspect of the experiment was the fact that the students did not see the proper appearance of the literature covers. It was of no consequent surprise that 58% said that the photographs were unacceptable in terms of quality.

However, despite this 75% of the sample thought that the ability to see what the popular chemical literature looked like before going to the library was useful.

An interesting observation was that 96% of the sample did not use a sheet of paper for taking rough notes considering the open book nature of assessment.

It was surprising to find that 50% of the sample would not prefer a lecturer explain how to use the package (remembering that 63% suffer from computer anxiety) but 54% would prefer a separate session to familiarise themselves with the package before doing the test.

All things considered student opinion was favourable. However, the students themselves reinforcing the need for a tutorial before running such a test highlighted computer anxiety and the fairness of doing such a test; it can not be assumed nor is it fair that summative computerised assessment should be thrust upon students.

3.3.12 Conclusion

One of the main advantages found by producing an assessment package of this nature i.e. a multimedia assessment package based on chemical information and retrieval, is that the student should be more efficient at retrieving the information from the resource centre. This is because the multimedia computer has provided the opportunity to assess student's knowledge through the use of photographs based on "every day" chemical literature. This was not previously possible on computers that have not got multimedia capabilities. The student can actually see what chemical literature bounds and contents actually look like. This application of photographs within assessment of an open-book nature has the benefit of being able to show students a wide range of chemical literature before they actually go to the resource centre. Students can be assessed if they actually know what the various elements of a abstract actually are. For example students are required to retrieve and use information found within Chemical Abstracts. Using a photograph students can be asked to identify the different elements before actually going to use the literature for the first time - which is a difficult and daunting task. This multimedia approach provides a close link between "the practical" and "the theoretical". Computerised assessment that had no multimedia capabilities could only assess student knowledge of a more theoretical nature using text based multiple-choice questions. For example, a question could be "What is Chemical Abstracts?" "What elements are found within a typical abstract?" The student would be given a selection of answers to chose from. These questions require students to know the answer, which could be done by rote memorising. However, with multimedia capabilities students may be required to classify chemical literature that they may never have seen before based only a cover and contents of the text. The students would be required to review other similar texts found in the information section of the package by looking at the contents pages especially to help them answer the question. The student is required to integrate and apply their knowledge based on the information presented to them to a new situation.

This type of assessment will encourage students to develop a more reflective approach to learning by being closely integrated within a structured framework. This framework should consist of say a lecture, which is then followed by the assessment followed by a hands-on session in the resource centre. The assessment will provide the student with immediate feedback from which they can reflect upon their knowledge. From the survey carried out within this research from the student sample that did this assessment if was found that 75% of the students thought that this ability to be assessed and see what chemical literature looked like was important.

Photographs have been used in this kind of assessment in the following roles: as a source of information that the student can use to review, as the focus of a question and as the feedback to a question. When a photograph is used as the focus of a question the photograph can be either an inactive or an active part of the question. The terms inactive and active refer to whether the student does not interact with the photograph or the student interacts with the photograph in order to answer a question respectively.

3.4 An Assessment Package to Test for the Identification and Usage of Apparatus

3.4.1 Aim

It is the aim of this part of the research to investigate the potential of using a multimedia solution as part of a strategy to assessing students' practical skills through reflective learning.

This investigation would primarily concentrate on practical skills of a typical degree course.

In order to devise an assessment the learning outcomes must first be identified. These learning outcomes must be analysed in terms of the knowledge and /or skills that the student will possess on successful completion of that part of the course. This analysis will serve as the basis from which an assessment strategy can be devised.

The learning outcome that was chosen for of practical skills was:

-selecting and setting up appropriate apparatus

This learning outcome requires the student to have a detailed knowledge of equipment: what it looks like, its name and its intended use. This learning outcome is particularly suitable for multimedia assessment due to the visual nature of practical chemistry. Students could be tested on the names of apparatus and its use through the incorporation of photographs into an assessment application.

It was decided to investigate the feasibility of creating an environment where the student could be assessed on their ability to build apparatus assemblies in a virtual environment through the use of photographs.

The aim of the application would be to provide the student with a virtual apparatus kit. From this kit they would select the appropriate piece of equipment to build the required apparatus assembly. It was intended to give the student a "free pathway" i.e. they could select equipment in any order to build the assembly as this scenario would be the most realistic. Students would then receive feedback at each step as to their progress. This assessment would be formative in nature. It is intended as a pre-laboratory assessment or it could be used in those cases where safety does not permit the experiment to be carried out or where, due to cost, the experiment is not feasible.

3.4.2 Objectives

The objectives of the project are to:

- identify the learning outcomes;
- produce photographs of apparatus on which the assessment could be based;
- assess the student's ability to identify apparatus;
- enable students to assemble a typical fractional distillation set-up and receive feedback as they assembled the apparatus;
- investigate how detailed feedback should be administered to the student to integrate the assessment / learning environments;
- evaluate the software in terms of staff and student appraisal;
- identify strengths and weaknesses of this type of assessment;
- identify further applications for this type of assessment.

3.4.3 Ideal Model of Application

The ideal model of the application would be to have a database of photographs of each piece of apparatus. The student would then select a piece of apparatus that would appear on the screen. This piece of apparatus would be freely rotatable through the x, y and z axes and the student would be able to "drag" it across the screen. With this piece of functionality the student would be able to assemble the apparatus. When two pieces of apparatus are joined the composition of the combined photographs would be adjusted to reflect the combination of the apparatus as would be expected in real life.

This ideal scenario just described is a very complex animation application and would use 256 colours. Due to the complexity of the development of such an application it would not be feasible within the time scale of this research. Instead a compromise was made to develop a scaled-down application but would, nevertheless, show the principles of how an assessment application of this nature would function.

Due to the complexity of developing a generic application that could be used to assess any pathway that a student may take to assemble any apparatus set up, it was decided to investigate the problems associated with developing a specific prototype designed to cope with just one scenario. The scenario that was chosen was the assembly of a typical fractional distillation apparatus. There was no specific reason for this choice other than this was a reasonably complex problem i.e. it involved the combination of several pieces of equipment and that it was therefore possible to build this apparatus in a number of different ways.

3.4.4 Experimental

This sub-section describes the steps that were taken in order to develop this software.

3.4.4.a) Consultation with Subject Experts

The first step in the development of any application is to plan in detail what exactly is required of the application i.e. its functionality. This will involve consultation between a subject expert(s) and the application developer. This consultation should elucidate the specification of the system that a subject expert would expect to see. In the case of the development of a typical fractional distillation the appropriate subject experts were consulted to:

1. gauge their impressions of the value of such an application to them and their students;
2. elucidate typical mistakes that students make when setting up a fractional distillation due to conceptual difficulties;
3. elucidate the "proper" method of setting up a fractional distillation arrangement;
4. identify how they would expect the application to function.

Three subject experts were interviewed. These experts had pre-concepts of computerised learning and assessment in general. One was against computerised learning, one was undecided and the third was in favour. These opinions would probably be reflected in their attitudes to the development of computerised assessment. This was found to be the case. The subject expert against computerised learning said that this type of package was below B.Sc. standard and would not be of benefit to B.Sc. students. A further comment was that hands-on experience could not be replaced by computer-based assessments. The expert that was in favour of computerised learning endorsed the concept of a virtual laboratory and said that it would benefit students in pre-laboratory sessions. Suggestions were made as to how to set up a fractional distillation. The most interesting interview was with the undecided expert. A prototype of the system had been developed to illustrate the fundamentals of the package. At first this expert said that this form of pre-laboratory assessment would not be beneficial to either himself or the students. However, when asked to elucidate the types of problems that students typically experienced during assembly of this piece of equipment the expert listed several problems. This is contrary to the first expert that implied students do not have any conceptual difficulties. The argument of the benefit of such an application would be to test each student's knowledge of setting up the equipment before a laboratory session. With adequate feedback the student's misconceptions could be highlighted and alleviated before any accidents occurred in the laboratory. When it was pointed out to this expert that these typical problems that they had elucidated could be assessed using this application, this expert concluded that this application would be beneficial to himself and students.

The following points on fractional distillation were elucidated from the experts:

1. The fundamental concept in fractional distillation is the design of the fractionating column. Separation depends on the surface area within the column. The greater the surface area the greater the separation. The student should be presented with say four types of column and questioned as to which one should be chosen to give the best separation. Four columns were given Empty, Vegra, Dufden, and Packed. The Packed column is filled with Raschig rings and a piece of copper spiral is dropped into the column before rings are added to stop the rings falling through.

2. In distillation the heating source is important. Students normally distil too fast. A good separation is achieved when the temperature is constant. However if

the temperature is increasing this leads to a bad separation. An oil-based heating source or isomantle are good heating sources for reflux but not for distillation. This is because it is not easy to adjust the heat flow with these sources. Instead a Bunsen burner is the best heating source. There are three ways in which the heat can be adjusted which are as follows: at the gas tap; blue / yellow flame or the actual height of the Bunsen burner beneath the gauze.

3. Some students have the retort stand the wrong way round because they can not fit the tripod on the stand base. This would lead to an unstable state. If the load on the stand becomes too great the stand will topple over which could lead to a serious incident. If the tripod does not fit on the stand's baseplate a wooden base plate should be fitted over the base of the stand.
4. Students forget to add anti-bumping granules to ensure smooth boiling.
5. Students do not grease joints.
6. Students forget to put stoppers on open necks of flasks. This would be dangerous especially if there is a flammable gas in the flask - it could ignite from the Bunsen.
7. Students use broken rubber tubing.
8. Students should put rubber tubing on the angled vacuum adapter when separating a flammable liquid. This is to reduce the fire risk by taking the flammable gas away.
9. Students do not position the thermometer correctly in the three-way adapter. It should be positioned at the arm mouth. If it is incorrectly positioned the student will not be reading the correct temperature of the distillate.
10. Students do not notice gaps in the joints of the apparatus. This will mean the distillate could escape and lead to a hazardous situation.
11. Students do not clamp round-bottom flasks below the lip of the mouth. This ensures that the flask does not "slip" through the clamp's jaws. Clamping on

the mouth of the flask could mean that could drop through the clamp's jaws that could lead to a serious accident.

12. Students use clamps that do not have cork. Clamps should always have cork.

Most of these common mistakes that students typically make would be included in the assessment as concepts that they would be tested on.

3.4.4.b) Identification of the Apparatus Required for the Assembly

The next stage in the development of the application was to identify the apparatus and steps involved in building a typical fractional distillation.

The apparatus that is required to build a typical fractional distillation set-up is given in Table 3.4.1 below:

Table 3.4.1 Apparatus Required to Assemble a Typical Fractional Distillation

stand	boss	clamp
round bottom flask	anti-bumping granules	fractionating column
grease	still head adapter	thermometer adapter
thermometer	Liebig condenser	rubber tubing
angled vacuum adapter	Bunsen	tripod
gauze		

3.4.4.c) A Suggested Method for Assembling the Apparatus

The main problem to the actual assembly of the apparatus is that the student could start at any particular point. A suggested method could be:

1. select stand and identify correct orientation

2. select Bunsen and position on stand
3. select tripod and position over stand and Bunsen
4. select gauze and position on gauze
5. select boss and position in correct orientation on the stand
6. select clamp and identify most stable position of clamp in boss
7. select round bottom flask and place in clamp making sure to clamp at the neck
8. select anti-bumping granules and place in the female joint of round-bottom flask
9. select fractional distillation column, identify the most efficient for separation, grease male joint and place in the female joint of round bottom flask
10. select still head adapter (three way adapter), grease male joint(s), and place in the female joint of distillation column
11. select thermometer adapter, grease male joint and place in the female joint of still head adapter
12. select thermometer and position the bulb of the thermometer at the mouth arm of the still head adapter
13. select stand and identify correct orientation
14. select boss and position in correct orientation on the stand
15. select clamp and identify most stable position of clamp in boss
16. select Liebig condenser, grease the male joint, position in the clamp and insert male joint of still head adapter in the female joint of the condenser
17. select angled vacuum adapter, grease the male joint and insert female joint into the male joint of the Liebig condenser
18. select stand and identify correct orientation
19. select boss and position in correct orientation on the stand
20. select clamp and identify most stable position of clamp in boss

21. select round bottom flask, position in clamp and insert female joint into the male joint of the angled vacuum adapter

22. select two pieces of rubber tubing and connect one piece to the top of the condenser and put the other end into the sink: connect the other piece to the bottom of the condenser and connect the other end to the water tap

23. select a piece of rubber tubing and connect one end to the angled vacuum adapter and the other end should be put well away from the heating source in a well ventilated area

This is one of several pathways that would be a correct method for building the apparatus. However a student may chose to take another route and they would not necessarily be incorrect. The application must not penalise the student for suggesting other routes. However, it is recommended that it would be good practice to encourage the student to think in a methodical manner. This could be the difference between forgetting to put a stopper in an open neck or not. For this exercise the suggestion is to build the apparatus from left to right which would mean the distillation column would be on the left-hand side.

3.4.4.d) Development Methodology

This application developed did not follow the ideal model. This was due to the time that would be required to develop the ideal model. Instead a modified model would be used. This model involves taking photographs of each and every stage in the apparatus assembly. As the student selected and joined each piece of apparatus a photograph would appear showing the constructed apparatus. The major disadvantage of developing a practical application using this model is that for every potential path the student could make a photograph would have to be taken. This would be both time consuming and expensive. However, for the purposes of prototyping the application to investigate its potential this would be the quickest method in the time available. In the case of the ideal model the constructed apparatus would be compiled using the individual photographs of each piece of apparatus. This would mean the total number of photographs would be equal to the number of pieces of apparatus that were available to the student. The scheme of the potential pathway using the suggested solution to the problem above was delineated. In total there were 70 steps defined in the apparatus assembly. The path to building the apparatus would be essentially linear i.e. the student would have to follow the path that was programmed in to the application. Where there was potentially more than one correct piece of apparatus to select the student would be guided to the appropriate path to focus on.

3.4.4.e) Identification of all the Apparatus to be Photographed

As previously mentioned the student would be presented with an apparatus tool kit. This tool kit consisted of photographs of many different pieces of laboratory equipment. A list was compiled of all the photographs that would be required. The apparatus included in the student tool kit is listed in Table 3.4.2 below:

Table 3.4.2 Apparatus in the Student Tool kit

stand	boss	clamp
1 neck round bottom flask	2 neck round bottom flask	3 neck round bottom flask
anti-bumping granules	empty distillation column	Vegra distillation column
Dufden distillation column	packed distillation column	beaker
grease	still head adapter	thermometer adapter
thermometer	Liebig condenser	rubber tubing
angled vacuum adapter	Bunsen	tripod
Buchner funnel	Hirsch funnel	glass funnel
vacuum flask	vacuum tube	measuring cylinder
filter paper	steam bath	isomantle
magnetic stirrer / hot plate	conical flask	3 way adapter
gauze	dropping funnel	stopper

These photographs would constitute the apparatus tool kit. In addition to these photographs, photographs were also taken of every stage in the apparatus assembly. These photographs had to be taken with the camera at a set distance from the apparatus. This was because the scale of each photograph needed to be the same to give a professional feel to the application. The following Table 3.4.3 details the photographs that were taken of each step in the apparatus assembly:

Table 3.4.3 Photographs of the Different Stages of the Apparatus Assembly

1. stand facing forwards
2. stand facing to the left
3. stand facing to the right
4. stand facing backwards
5. boss right way up
6. boss upside down
7. boss + stand
8. clamp + boss + stand
9. round bottom flask + clamp + boss + stand
10.packed distillation column + 9
11.still head adapter + 10
12.thermometer adapter + 11
13.thermometer + 12
14.Liebig condenser + stand / boss / clamp + 13
15.angled vacuum adapter + 14
16.round bottom flask + stand / boss / clamp + 15
17.rubber tubing on condenser (water in) + 16
18.rubber tubing on condenser (water out) + 17
19.rubber tubing on angled vacuum adapter + 18

These photographs were digitised using the Kodak Photo CD method. A professional photographer was employed to take the photographs of the apparatus. This was because taking photographs of glassware caused problems due to the internal reflections.

3.4.4.f) Identification of the Application Functionality

The next step involved in the development of the application is to design the interface and identify the functionality.

The application must have the following features:

- provide instruction to the student indicating what is required;
- provide feedback to the student as to their success of completing a task;
- have an apparatus tool kit to enable the student to select apparatus;
- have an area to display photographs of the apparatus from the tool kit;
- enable a student to confirm that this is actually the piece of apparatus they want to select;
- have an area to display assembled apparatus;
- enable the student to interact with photographs to allow the student to build the assembly;
- enable a student to confirm their responses;
- monitor student responses to determine if their answer is correct or incorrect;
- provide adequate feedback to the student;
- have a tutorial on how to use the application.

3.4.4.g) Designing the Interface

The apparatus tool kit was placed in the standard Windows 3.1 tool bar area i.e. beneath the menu bar. A button with an icon picture of the apparatus represented each piece of apparatus. This would graphically represent to the student what each button represented. The icons were produced using Microsoft Image Editor Version 3.10 which comes with the Microsoft Windows Software Development Kit Version 3.10. Each icon was 24x24 pixels in size and was saved as a .BMP file. Despite the good detail that can be achieved within each icon some of the icons may be ambiguous. Therefore, tools tips were also produced for each button in the tool kit. A tool tip is a textual piece of information that appears when the mouse pointer in over a button for a period of a few seconds. When the mouse pointer moves off the button the tool tip disappears. The tool tips would appear beneath the apparatus tool kit in the same place every time to aid student learning of the interface.

During the application design it became apparent that the tool kit would have dual functionality. Not only would it be used, as a source for providing the apparatus to the student it would also have "action" buttons also. Buttons would be available for joining apparatus together and greasing the appropriate joints.

There were three other major areas that the screen would be used for:

- textual information;
- showing apparatus via the tool kit;
- showing assembled apparatus.

The screen resolution for this application was 640x480. Once the title bar, menu bar and apparatus tool kit were taken out, the space left was 640x400. The main body of the screen was reserved for the apparatus assembly area. This area was positioned to the right hand side of the screen. The width of this area was 390 pixels. The remaining width, 250 pixels, would be reserved for the textual information and for displaying photographs of the apparatus. From Appendix II *Interface Design*, it has been found that the most appropriate place for textual information to be displayed is at the top left-hand corner. This area would be 100 pixels high. The area below this was reserved for the photographs of the apparatus i.e. 250x300 pixels.

Ideally photographs should be as large as possible in any application. The space available in each photographic area was adequate for most cases. However, detail was lost for larger pieces of apparatus. For this reason a zoom facility was added to each area to enable the student to see in good detail the apparatus. It is important that the photographs should be of the highest quality possible. When the zoom button was clicked scroll bars horizontal, vertical or both would appear as appropriate with an enlarged photograph. If the zoom button was clicked again the photograph would be zoomed out.

3.4.4.h) Joining and Greasing Apparatus

In this application the apparatus is joined together by selecting the join button. This puts the student into join mode. The student is then asked to identify the joint of the apparatus they have currently selected. This is done by simply clicking on the appropriate joint in the photograph. Visual representation is given to the student to indicate where they have selected by a red filled circle. If the student wants to change their choice they just click on the photograph again. Students confirm their answer clicking on the confirmation button. This button has a green tick drawn on it. This has been chosen to be user friendly i.e. students will be familiar with the concept that a green tick means OK, agreement. When this button has been clicked feedback will be relayed back to the student as to whether they are correct or incorrect. The student is then required to identify the joint in the apparatus assembly area by clicking on the photograph, confirming their choice and they will then receive the appropriate feedback. A photograph will then be placed in the apparatus assembly area showing the assembled apparatus.

The student can grease male joints by clicking the grease button. They will then be required to click the joint that they wish to grease. A photograph appears with the appropriate joint greased. One of the advantages of digital photographs is that they can be easily manipulated by applications specifically designed for that purpose, for example PhotoStyler V 2.0. Instead of having to take a photograph of an ungreased joint and then a greased joint, an image of a greased joint could be easily produced using the photograph of the ungreased joint. This was done using the smudge tool within PhotoStyler. In this way duplication of photographs could be avoided. Another use of digital wizardry was utilised for the orientation of apparatus. For example a photograph was produced showing a stand with the baseplate facing forwards, then to the left, to the right and to the back. This was produced by using only one photograph which was the baseplate facing forwards. Through digital manipulation the other three orientations were produced. The zoom facility utilised the Visual Basic (VB) Image Control. Instead of producing two photographs i.e. zoomed in and zoomed out, just one photograph was produced - a zoomed in one. The VB Image Control can be set to automatically size the photograph to the size of the control or size the control to the size of the photograph. In this manner the user gets the impression that the photograph is being zoomed in / out.

The following five figures illustrate the application's interface.

Figure 3.4.1 illustrates the opening screen when the student enters the package.

Figure 3.4.2 shows how the student's knowledge of apparatus orientation is tested. The student is required to click on the photograph to select their answer and then click the green tick button to confirm that this their answer.

Figure 3.4.3 shows the instruction when the student selects a piece of apparatus. Notice that the apparatus name is given below the apparatus toolkit at the top of the screen in the yellow "tool tip" panel. When the student has selected a piece of apparatus they must click the green tick button to confirm their answer.

Figure 3.4.4 shows how the student is "led" through the apparatus assembly when there is more that one piece of apparatus that could be assembled. In this case the student could for example add rubber tubing to the condenser. The application indicates to the student the area on which to concentrate by highlighting the area using a red rectangle.

Figure 3.4.5 shows an example of the feedback that is provided to the student if they input an incorrect answer. This figure also shows what the completed assembly looks like.

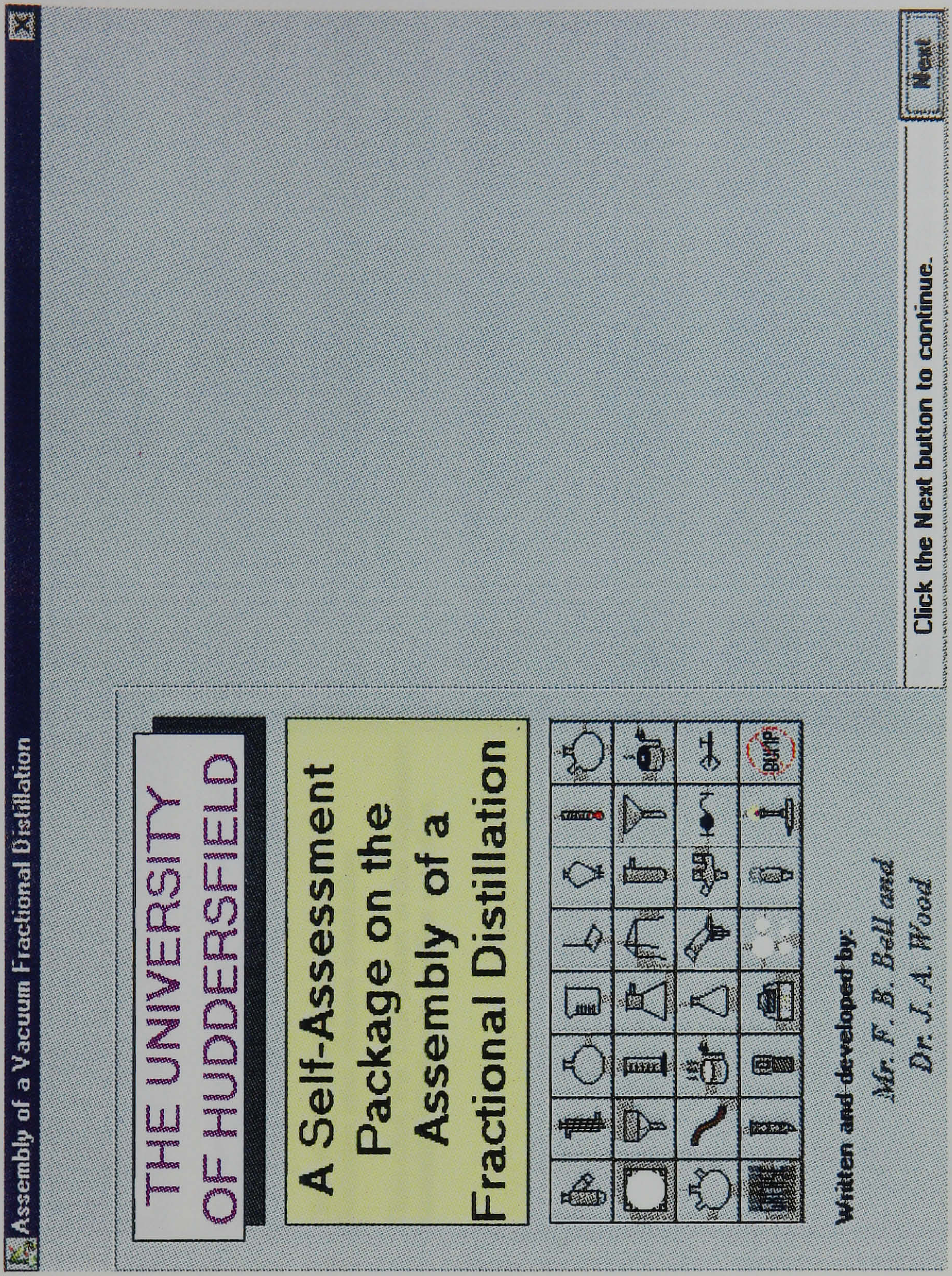


Figure 3.4.1 Opening Screen of the Fractional Distillation Package

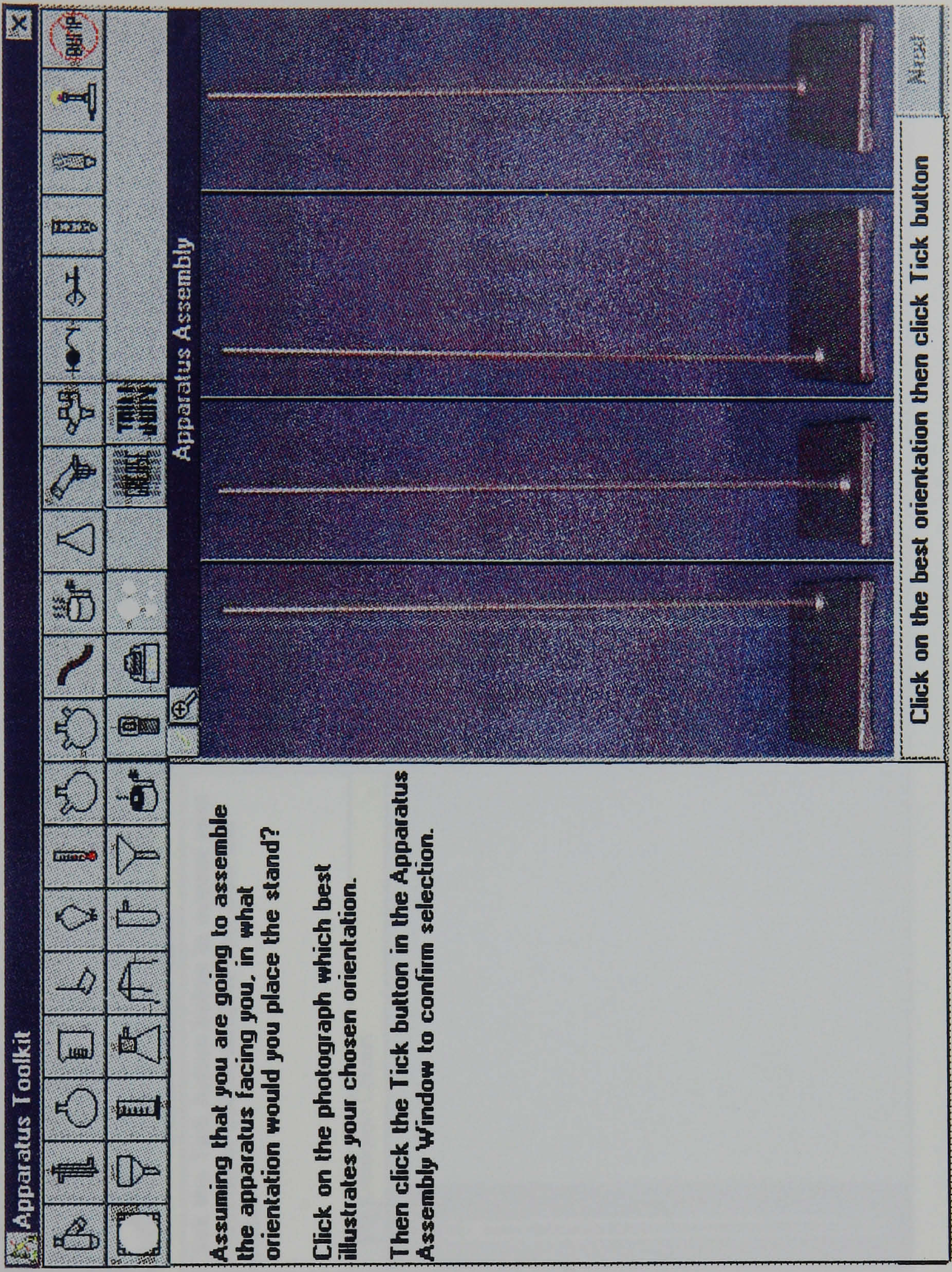


Figure 3.4.2 Question to test Student Knowledge on the Most Stable Orientation of a Stand

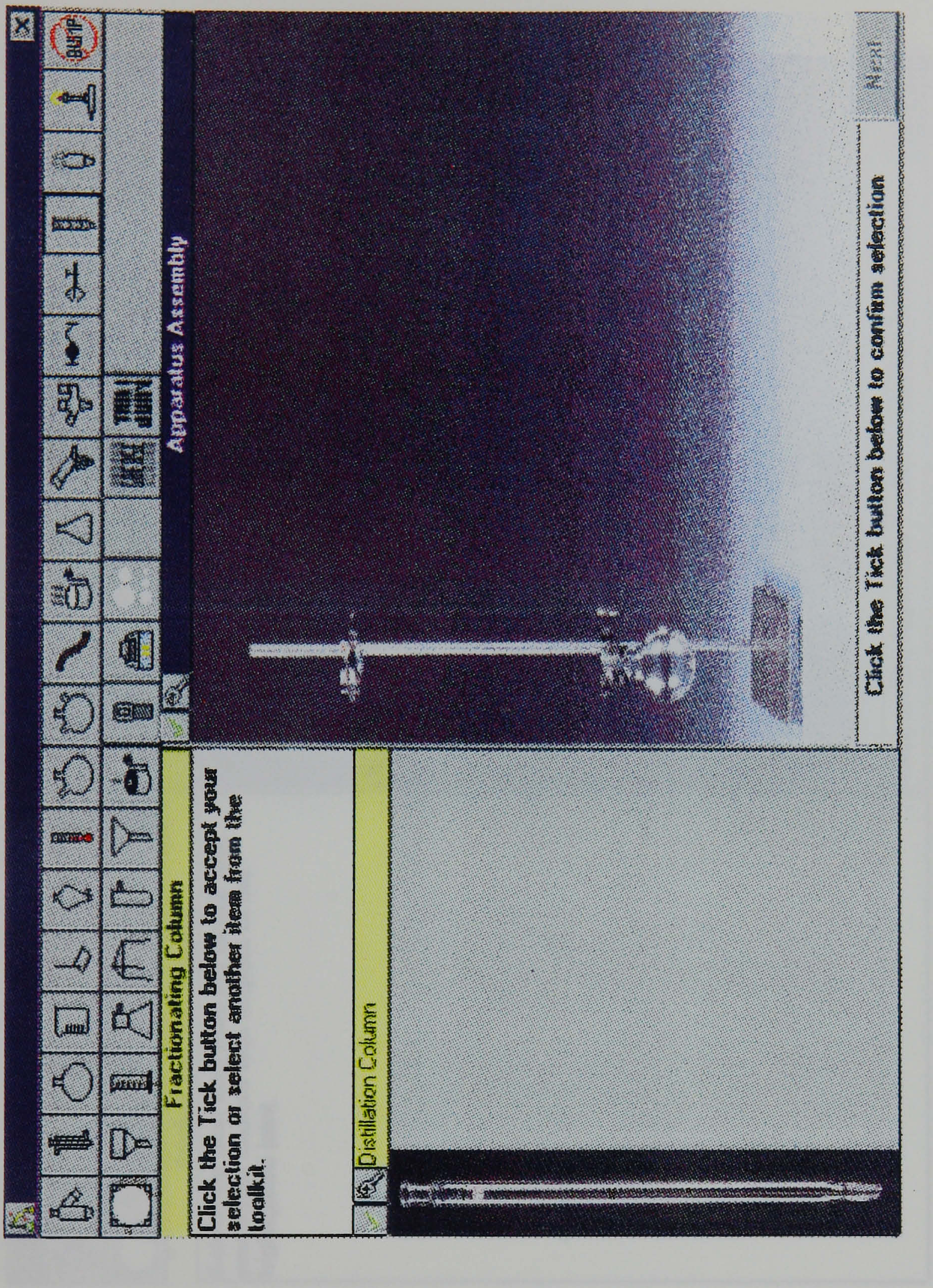


Figure 3.4.3 Illustration of Selecting a Piece of Apparatus

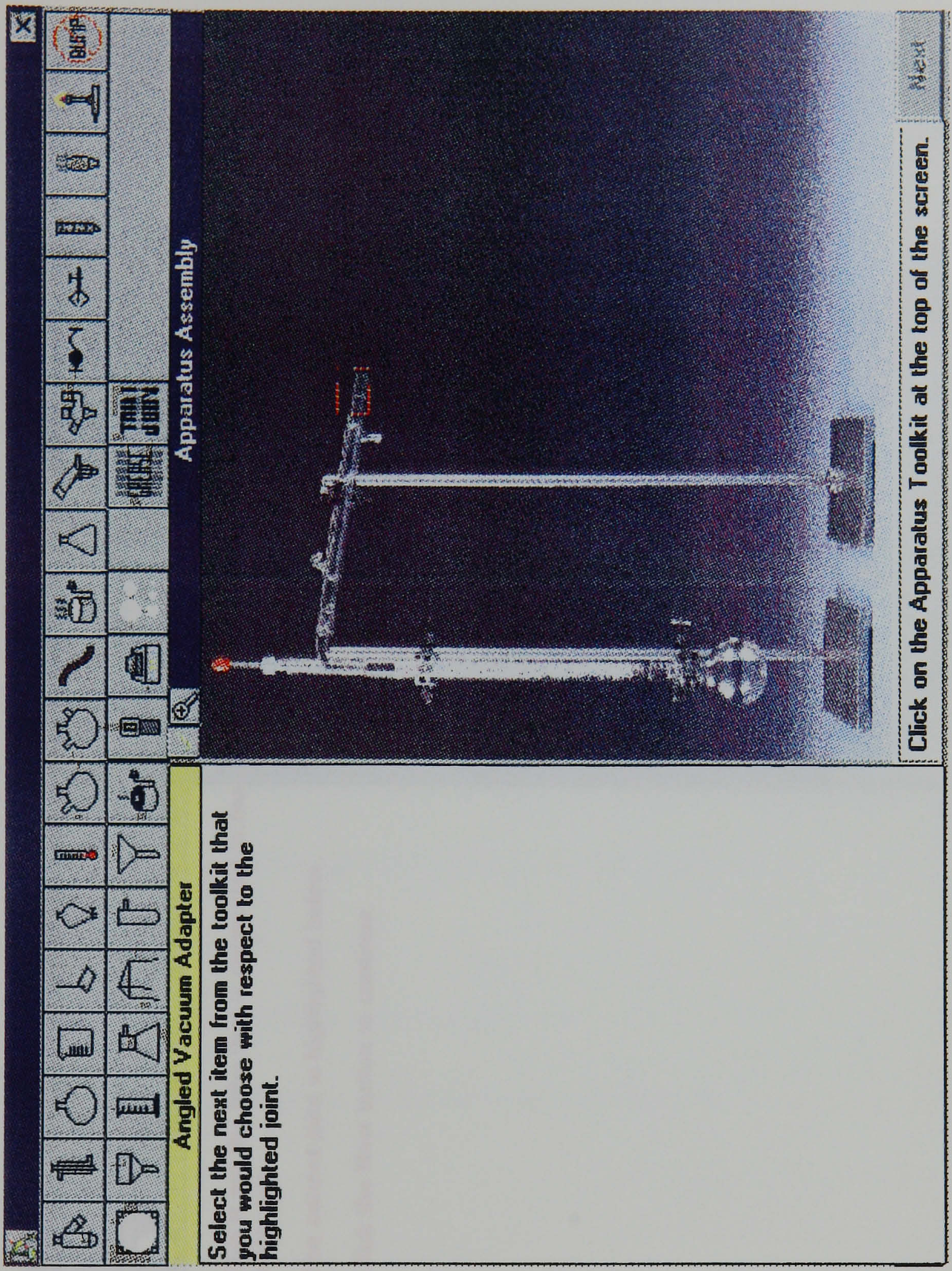


Figure 3.4.4 Illustration to Show How the Student is "Led" When Selecting a Piece of Apparatus

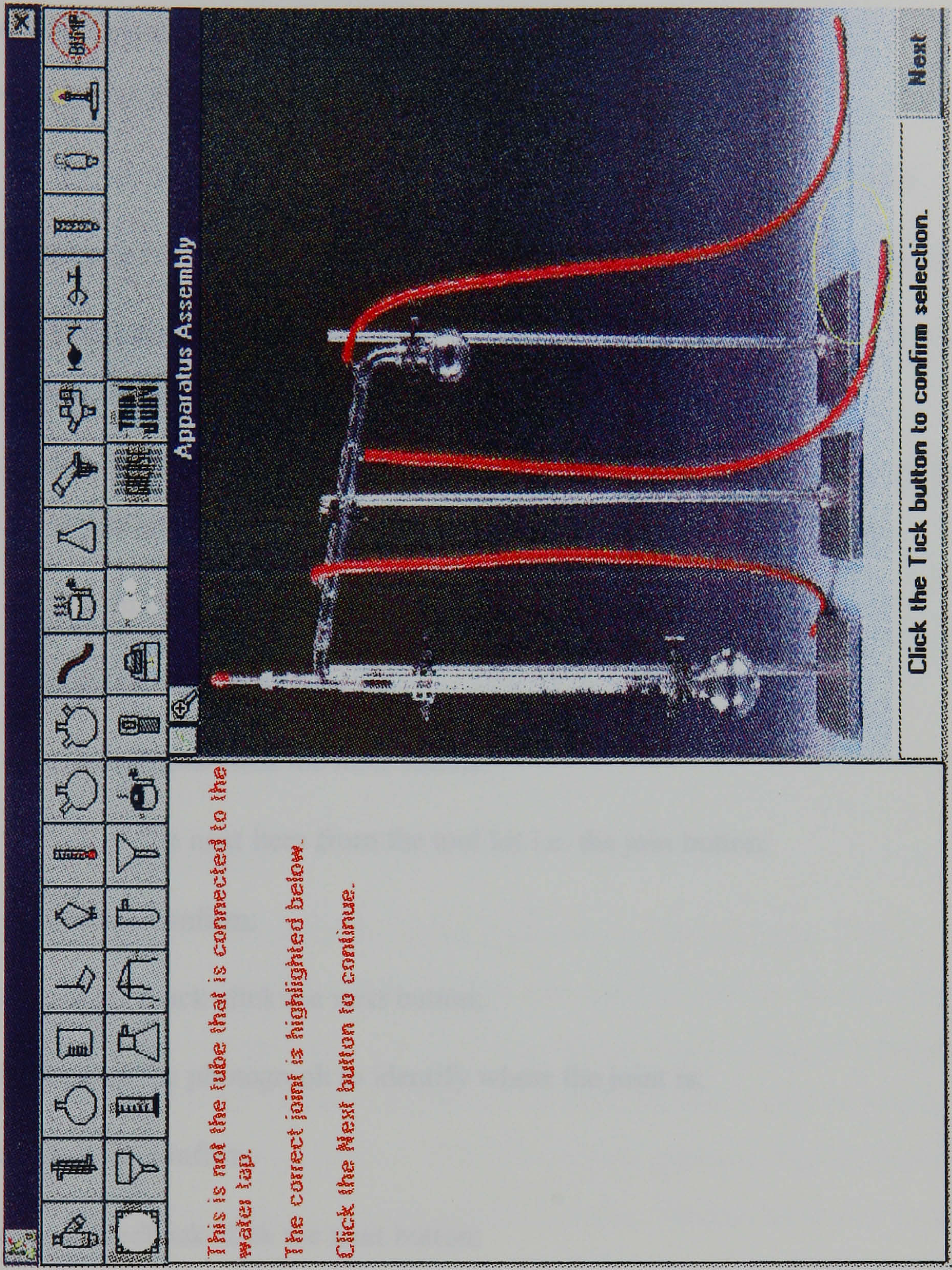


Figure 3.4.5 Illustration to Show Feedback When an Incorrect Answer is Given

3.4.4.i) Typical Steps to Building the Apparatus

Typically the steps involved for building the apparatus would be:

- to click on the tool kit to select the first piece of apparatus i.e. the stand;
- to click the confirm button in the photograph display area to confirm the choice;
- feedback would then be displayed to the student indicating if their choice was correct or incorrect;
- to click the next button once the feedback has been read and understood. The apparatus assembly area would then display a photograph of the stand;
- to click on the next piece of apparatus from the tool kit i.e. the boss;
- to click the confirm button;
- to read feedback click the next button;
- to click on the next item from the tool kit i.e. the join button;
- to click the confirm;
- to read feedback click the next button;
- to click on the photograph to identify where the joint is;
- to click the confirm;
- to read feedback click the next button;
- to click on the assembly photograph to identify where the other joint is;
- to click the confirm;
- to read feedback click the next button.

This procedure is repeated until the apparatus is assembled.

Each student response is stored in a database. The database is created and filled dynamically at runtime. The database records what the student response is, what the actual answer is and the mark awarded to the student. At the end of the assessment a score is relayed back to the student.

3.4.5 Evaluation

This application has been shown at a mini-conference within the University of Huddersfield. The immediate feedback was "what about if the student selected another path" or "this would not be the way that I would do it". As always with any piece of educational software the *"not invented here syndrome"* always arises. However, the main emphasis of this package is to demonstrate the potential of using computers as an assessment tool and to push back the boundaries of computerised assessment beyond the multiple-choice question. An objective and encouraging viewpoint was obtained from a representative of the Open University who highly appraised the worthiness of such an assessment method.

3.4.6 Conclusion

This sub-project investigated the further application of photographs within computerised assessment. It was found that photographs could be used to enable students to assemble apparatus. The ideal application could not be built due to the time restrictions in this research. In the ideal system students would be able to build the apparatus without any guidance. This removes the guessing element that was normally associated with the multiple-choice question. Interaction would be achieved by "drag and drop" techniques. The system would be given a set of information based on each piece of apparatus that would indicate where the apparatus joints were, if it was a male or female joint, if there are any joints at all. This information can then be used in a set of rules to determine if the pieces of apparatus could be joined where the student was currently indicating.

The system that was prototyped demonstrated the principles and the potential of this kind of application.

The main advantage of this type of system is that students can be exposed to more apparatus than they would through their "normal" laboratory experience and can see actually see what the apparatus looks like. This was not previously possible using a computer without multimedia capabilities. The students can build the assembly before actually doing the experiment. By doing a pre-laboratory session students can be assessed on their knowledge of the assembly before actually doing it for real. This is beneficial because often in the laboratories students frequently have problems assembling apparatus. This has been demonstrated in earlier sections of this sub-project where the experienced practitioners when asked to list student problems on one apparatus assembly listed 12 points. The assessment can be closely linked with these problem areas so that the students can be assessed on these specific points and provided with the appropriate but immediate feedback. Pre-laboratory assessments based on the practical itself as opposed to the theoretical material will help the students to assemble the apparatus more precisely due to their familiarity with the apparatus.

This kind of assessment can be used in other ways for example to view and / or assemble apparatus that the institution does not have available or due to safety reasons can not. The development of a package of this nature in which students can assemble any kind of assembly would be beneficial to secondary and even primary level. This will enable these students to see apparatus and build apparatus that they may not get to see until tertiary level. The other advantage to this group is it will increase the level of practical knowledge so that there will be less of a burden at the tertiary level where lecturers have to introduce basic types of apparatus to new students.

3.5 An Assessment Package for the Determination of the Enthalpy of Reaction

3.5.1 Aim

It is the aim of this part of the research to demonstrate, through the development of a prototype expert system, how many of the theoretical problem solving exercises within a chemistry degree course could be computerised. If paper-based problem solving exercises will be replaced by computer-based problem solving exercises the computer in effect will become an *electronic notebook.* However, unlike paper-based exercises, it will act as an intelligent tutor enabling the student to input individual steps of their solution monitor student's responses in real time and provide them with immediate feedback as to their methodology in solving the problem.

The hypothesis here is that if one problem solving exercise can be computerised then it must follow that other similar problem solving exercises can be also computerised.

3.5.2 Method

The method used to develop such a prototype was to take a typical numerical problem solving question of degree standard and analyse / elucidate the steps involved that a lecturer would employ in marking the student's solution. This would be done through close liaison with the appropriate subject specific lecturers. Next the problem would be delineated with respect to all the possible solutions using an appropriate methodology. This "knowledge base" would then be coded into the computer to be used to enable the computer to "think" like a human counterpart.

The problem that was selected was the determination of the enthalpy of reaction using simultaneous equations as the method used to solve to the problem. The reason for selecting this problem was that this question provided a good example of mathematical \ numerical problem solving within chemistry.

Within the specific problem of determining the enthalpy of reaction using simultaneous equations there are several rules that have to be strictly adhered to. These include:

- correct notation of chemical formulae;
- balancing equations;
- multiplying /dividing equations by a constant;
- adding / subtracting equations.

3.5.3 A Model for the Determination of the Enthalpy of Reaction using an Expert System

A model of how an expert system would operate with respect to the determination of the enthalpy of reaction would be:

- Define all the possible methods to solve the problem.
- Write modular programs that would contain the knowledge base necessary for the application to enable the student to solve by which ever method they wanted.

Due to time restrictions it would not be possible to produce an expert system that could comprehensively cope with all the methodologies possible to solve the problem in question. It was decided to investigate, propose and develop a prototype of how an expert system would solve the problem using simultaneous equations.

3.5.4 Preliminary Stages

Before starting to develop the system it is very important that the project is well planned with respect to:

- defining the aims and objectives of the assessment i.e. is it for summative or formative purposes; when should feedback be given to the student - after each step or at the end of the assessment?

- what is the level of interaction that the user will have to do? Is the authoring tool capable of this? For example, if they have to type in chemical equations is there provision for sub- or superscripts or, if they are required to draw chemical molecules, is there provision for this?
- how is the knowledge base going to be compiled i.e. from text books or experts. If it is from an expert, a questionnaire will need to be prepared to elucidate the necessary knowledge from the expert.

3.5.5 System Requirements

A typical specification of the system requirements with the specific problem would be to:

- generate a randomised question for each student;
- enable the user to interact with the environment as easily as possible (however, the key is to provide students with training in the use of the assessment package);
- enable the user to input individual steps for their solution as well as the final answer;
- calculate the correct answer with respect to the student's answer and calculate what the absolute answer should be and compare and mark the student's answer;
- provide the student with feedback with respect to their answer and the absolute answer;
- insert the correct answer for this stage and ask the student to repeat the calculation based on this new information and re-mark the student's answer;
- repeat the problem until the student has solved it.

The question generated by the computer, the student responses and feedback can all be stored in a database if required by the lecturer for analysis after the assessment - this would be strongly advised.

3.5.6 Screen Shots of the Interface

The following four figures illustrate some of the steps that would be required to solve the problem.

Figure 3.5.1 illustrates the opening screen of the package.

Figure 3.5.2 illustrates where the student inputs their answer and the automatic subscripting and letter-casing for the chemical formula that the system does for the student.

Figure 3.5.3 illustrates the feedback that is given to the student when they have completed a step of the solution.

Figure 3.5.4 illustrates the instruction that is given to the student to help them interact with the software.

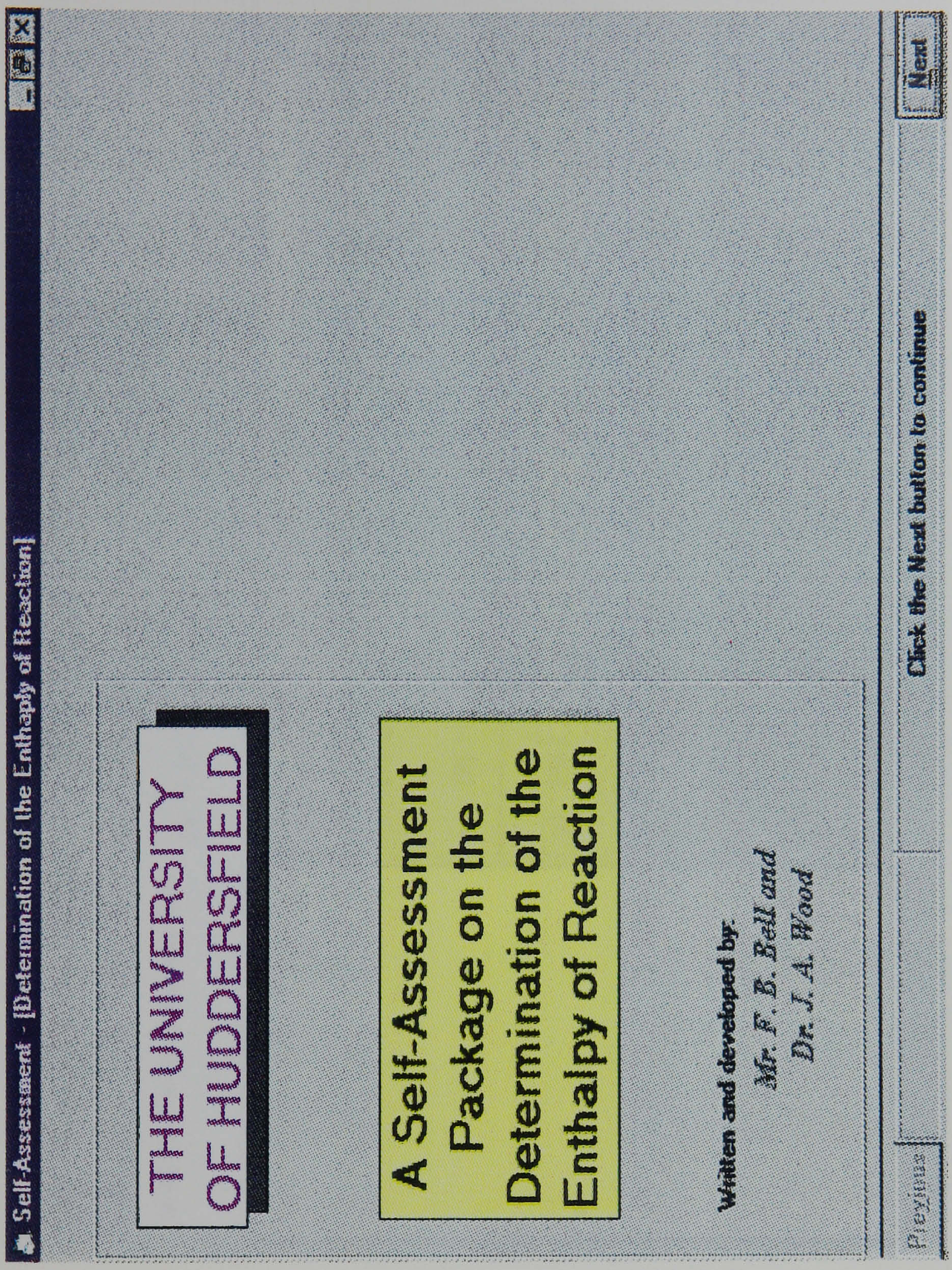


Figure 3.5.1 Opening Screen of the Determination of the Enthalpy of Reaction Distillation Package

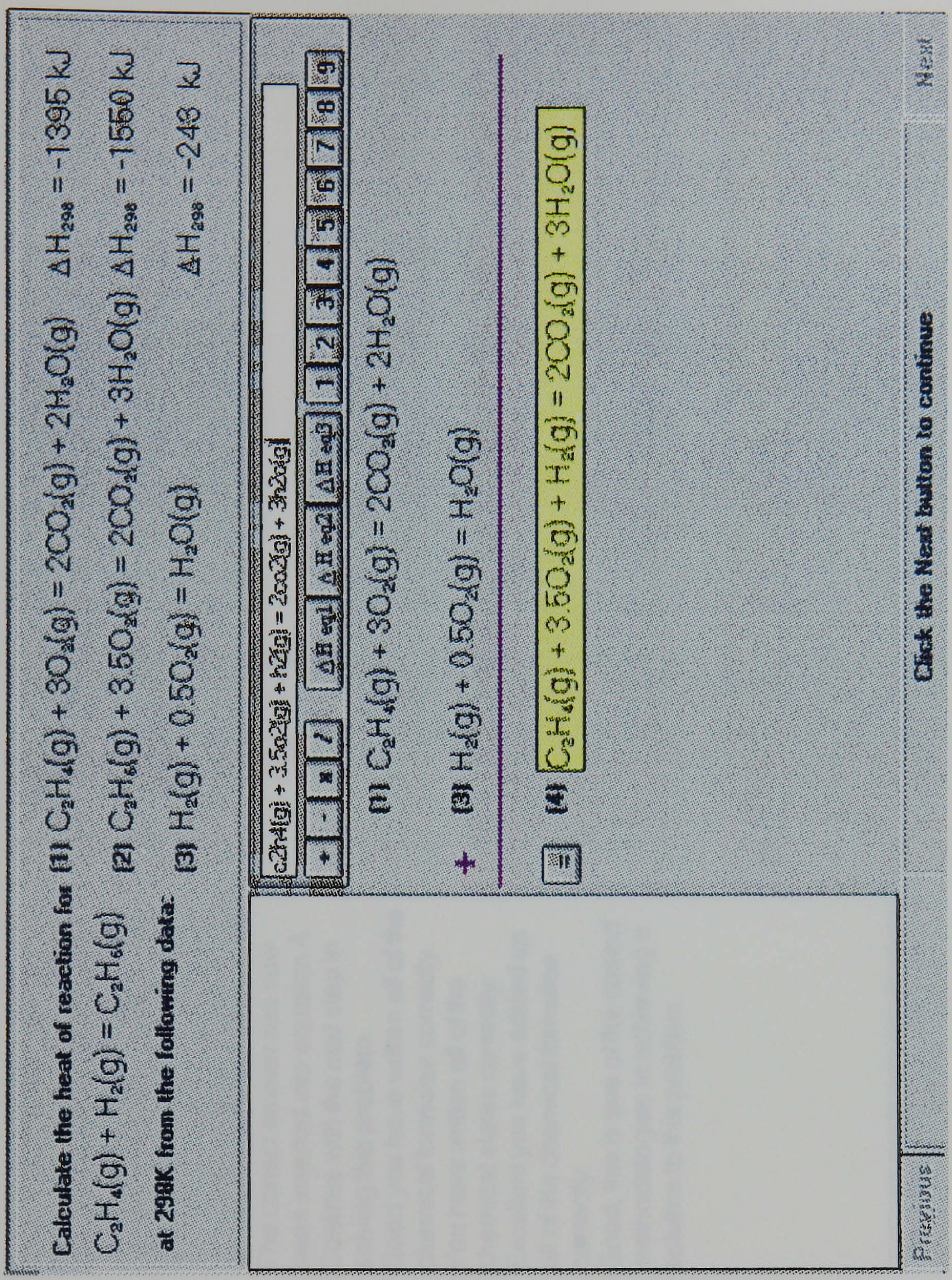


Figure 3.5.2 Illustration of How the Student Inputs their Answer

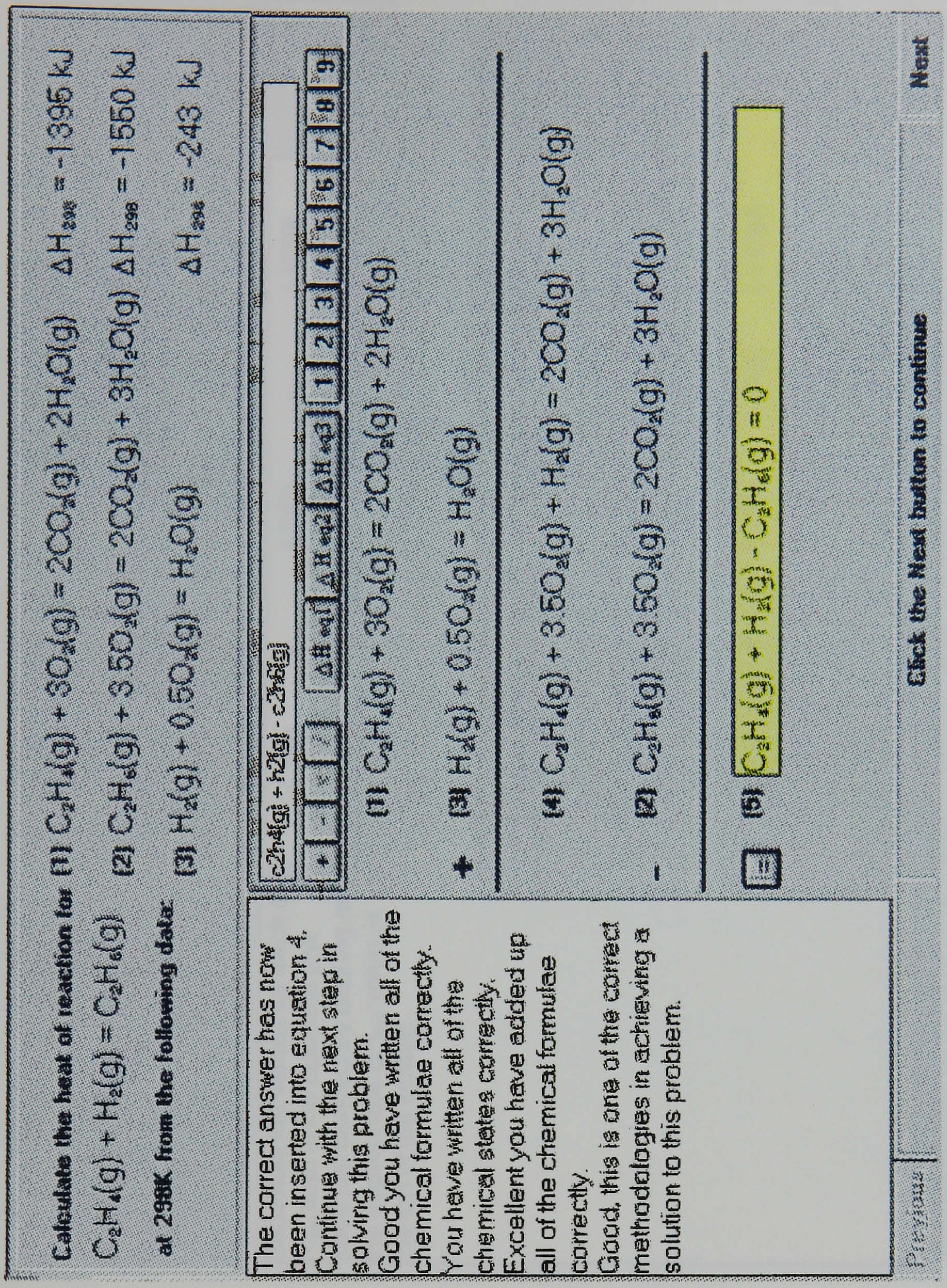


Figure 3.5.3 Illustration of the Feedback that is Provided to the Student

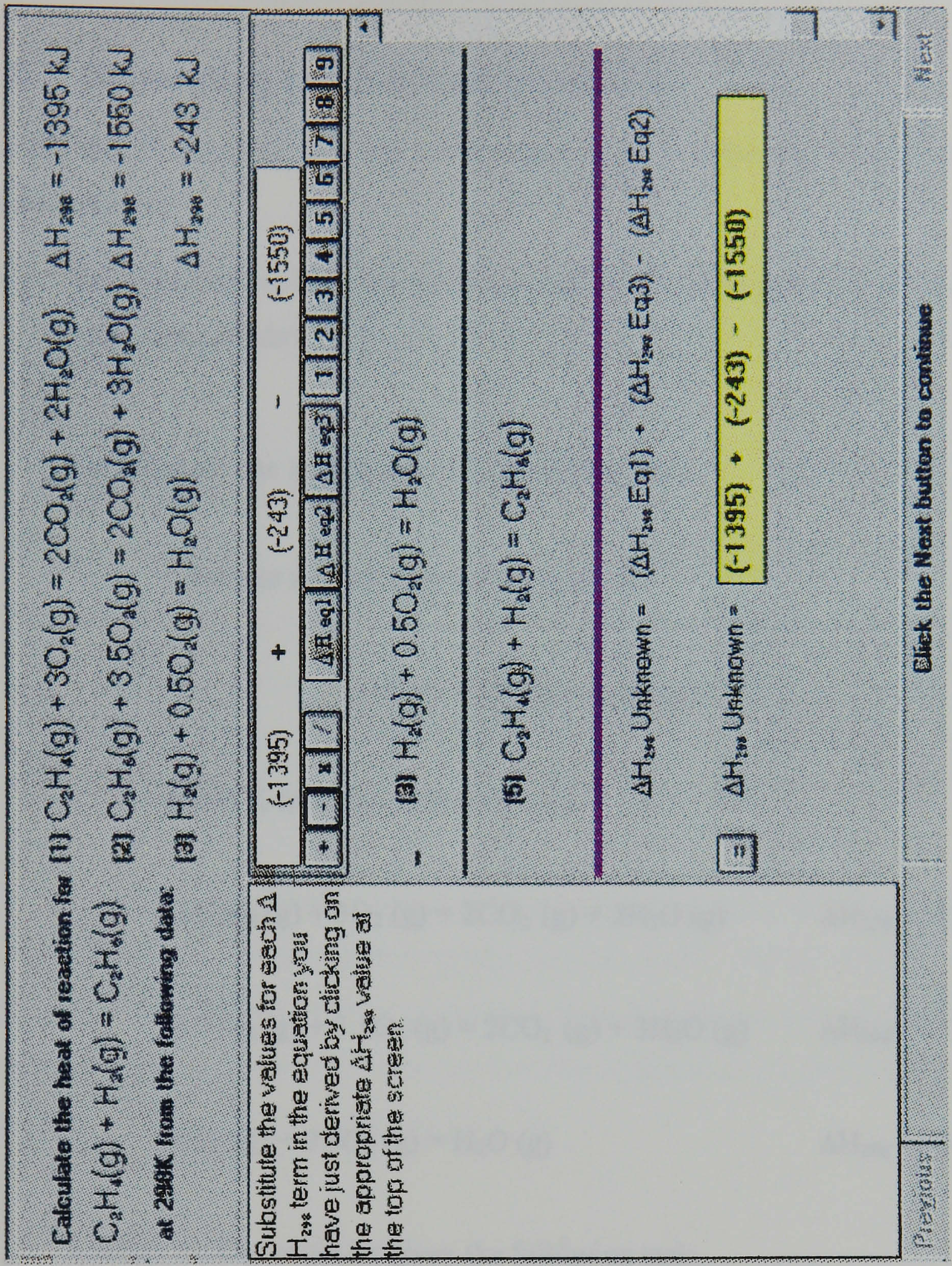


Figure 3.5.4 Illustration of the Instruction given to the Student to Help them Interact with the Application

3.5.7 Generating randomised questions

For the specific problem there are two possible methods in the way the question could be randomised;

1. the equations could be randomised based on the given alkane or alkene (or any other compounds);

For example one student could be given the task:

Calculate the enthalpy for the reaction:

$$C_2H_4\,(g) + H_2\,(g) = C_2H_6\,(g)$$

at 298 K from the following data:

1) $C_2H_4\,(g) + 3O_2\,(g) = 2CO_2\,(g) + 2H_2O\,(g)$ $\Delta H_{298} = -1395$ kJ

2) $C_2H_6\,(g) + 3.5O_2\,(g) = 2CO_2\,(g) + 3H_2O\,(g)$ $\Delta H_{298} = -1550$ kJ

3) $H_2\,(g) + 0.5O_2\,(g) = H_2O\,(g)$ $\Delta H_{298} = -243$ kJ

Another student could be given the following task:

Calculate the enthalpy for the reaction:

$$C_3H_6\,(g) + H_2\,(g) = C_3H_8\,(g)$$

at 298 K from the following data:

1) $C_3H_6\,(g) + 4.5O_2\,(g) = 3CO_2\,(g) + 3H_2O\,(g)$

2) C_3H_8 (g) + $5O_2$ (g) = $3CO_2$ (g) + $4H_2O$ (g)

3) H_2 (g) + $0.5O_2$ (g) = H_2O (g)

Note in both cases (in fact all cases for this particular example of alkenes and alkanes if the order of the equations is the same as above) the solution is the same i.e. Eq1 + Eq3 - Eq2 or any rearrangement of this solution.

2. the equations could also be randomised based on the order that they are presented to the students.

   For example, for student A the equations could be in the order 1, 2, 3 but for student B the same equations could be presented in the order 3, 1, 2 and for student C the equations could be presented in the order 2, 3, 1 etc. There is a total of six possible combinations of the way these three equations could be presented: (1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1).

It is possible to determine the number of combinations possible with three given equations by the two randomised methods described above. For the range of alkenes with 2 carbons to 5 carbons the total number of randomised questions would be equal to the number of questions times the number of combinations possible with three equations which in this example is 4 x 6 = 24.

In order to develop an expert system capable of producing its own randomised questions one must provide the rules for doing so. If we concentrate on the specific example of alkenes and alkanes this means the computer will have to "understand" the relationship between the formula of alkenes and alkanes (i.e. C_nH_{2n} is the general formula for alkenes and C_nH_{2n+2} is the general formula for alkanes,) and how to balance chemical equations in order to be able to produce balanced equations.

Essentially a balanced equation is given by the following rule (Conservation of Mass):

The number of atoms on the left-hand side must equal the number of atoms on the right hand side.

This can be illustrated through the following balanced equation example:

$$C_2H_4\,(g) + 3O_2\,(g) = 2CO_2\,(g) + 2H_2O\,(g)$$

Taking an atom balance using the above rule:

No. of atoms on left-hand side	=	**No. of atoms on right-hand side**
2 carbon atoms from C_2H_4	=	2 carbon atoms from $2CO_2$
4 hydrogen atoms from C_2H_4	=	4 hydrogen atoms from $2H_2O$
6 oxygen atoms from $3O_2$	=	6 oxygen atoms from $2CO_2$ and $2H_2O$

The computer can replicate this expertise by providing it with an explicit definition of these rules.

When any alkene reacts with hydrogen to produce an alkane we can produce the following generic equation:

$$a\ C_nH_{2n}\,(g) + b\ H_2\,(g) = c\ C_nH_{2n+2}\,(g)$$

By applying the conservation of mass we can derive a set of equations that define *a*, *b*, and *c*.

No. of atoms on left-hand side	=	**No. of atoms on right-hand side**
No. carbons = a x n	=	No. carbons = c x n
No. hydrogens = (a x 2n) + (b x 2)	=	No. hydrogens = c x (2n + 2)

From the carbon balance we can say:

$$a \times n = c \times n$$

therefore $a = c$

From the hydrogen balance we can say:

$$(a \times 2n) + 2b = c \times (2n + 2)$$

substituting a with c

$(c \times 2n) + 2b = 2nc + 2c$

$2nc + 2b = 2nc + 2c$

$2nc - 2nc + 2b = 2c$

$2b = 2c$

$b = 2c / 2$

therefore $b = c$ $(= a)$

From the definition of these simple rules and then compilation of the variables into an equation it is possible for the expert system to produce a balanced equation for the reaction between any alkene and hydrogen.

We know that if a hydrocarbon reacts with oxygen the products are carbon dioxide and water. For example, for the following generic equation for the reaction of alkenes and oxygen;

$$a\ C_nH_{2n}\ (g) + b\ O_2\ (g) = c\ CO_2\ (g) + d\ H_2O\ (g)$$

applying the above rule, we can produce a set of rules to describe c in terms of a and n; d in terms of a and n; and b in terms of c and d

No. of atoms on left-hand side	**=**	**No. of atoms on right-hand side**
No. carbons = $a \times n$	=	No. carbons = c
No. hydrogens = $a \times 2n$	=	No. hydrogens = $d \times 2$
No. oxygens = $b \times 2$	=	No. oxygens = $c \times 2 + d$

Note $a = 1$

IF n = 4 THEN

formula of alkene = C_4H_8

$c = a * \text{n}$

therefore no. of carbon dioxide molecules = 4 (CO_2)

$d = (a * 2\text{n}) / 2$

therefore no. of water molecules = 4 (H_2O)

$b = (c * 2 + d) / 2$

$b = ((a * \text{n} * 2) + (a * \text{n}) / 2)) / 2$

therefore no. of oxygen molecules = 6 (O_2)

From the definition of these simple rules and then compilation of the variables into an equation it possible for the expert system to produce a balanced equation for the reaction between any alkene and oxygen.

By following a similar mechanism it is possible to do the same for alkanes.

$$a\ C_nH_{2n+2}\ (g) + b\ O_2\ (g) = c\ CO_2\ (g) + d\ H_2O\ (g)$$

No. of atoms on left-hand side	=	**No. of atoms on right-hand side**
No. carbons = a x n	=	No. carbons = c
No. hydrogens = a x (2n + 2)	=	No. hydrogens = d x 2
No. oxygens = b x 2	=	No. oxygens = (c x 2) + d

Note $a = 1$

IF n = 4 THEN

formula of alkane = C_4H_{10}

$c = a * \text{n}$

therefore no. of carbon dioxide molecules = 4 (CO_2)

$d = (a * (2\text{n} + 2)) / 2$

therefore no. of water molecules = 5 (H_2O)

$b = (c * 2 + d) / 2$

$b = ((a * n * 2) + (a * (2n + 2)) / 2)) / 2$

$b = (2na + na + a) / 2$

$b = a (2n + n + 1) / 2$

therefore no. of oxygen molecules = 6.5 (O_2)

From the definition of these simple rules and the compilation of the variables into an equation it is possible for the expert system to produce a balanced equation for the reaction between any alkane and oxygen.

The equation for the enthalpy of reaction between hydrogen and oxygen is constant. Therefore this equation can be entered into the computer directly as a constant.

If the system is used to generate the randomised number, n, between certain limits the system can generate three randomised equations providing individualised questions for each student within the limits specified. Once the equations have been generated they can be put into any order to further individualise the question. The purpose is that students will not be able to copy directly from another colleague.

3.5.8 Student Interaction

One of the major problems with Visual Basic (VB) 3.0 Professional was that there was no direct provision for sub- and superscripted text. This problem was overcome by using a third-party tool designed to allow sub- and superscripted text. However, this tool can not receive input directly from the keyboard. A compromise had to be made. When students input text from the keyboard it would be received in VB text box control. This text would then be passed to the third-party control where the appropriate conversion would take place to format the text for the student i.e. it capitalised the chemical formulae and subscripted numbers as appropriate.

In order for the system to capitalise chemical formula and subscript numbers, it was necessary to define the rules for the system to understand chemical formulae. Consider the following equation again:

$$C_2H_4\,(g) + 3O_2\,(g) = 2CO_2\,(g) + 2H_2O\,(g)$$

The main problem for the system is that there are two different functions for the numbers that appear in the equation i.e. numbers are used for stoichiometric quantities and in the chemical formula. However, they can be easily distinguished from one another. From the equation above it can be seen that subscripted numbers always appear after a letter whereas integer values always appear after a space. By "memorising" the previous alphanumeric character and the current alphanumeric character it is possible to determine whether the number should be subscripted or not. Capitalisation is easily achieved by forcing capitals from standard VB commands.

Rule

IF (previous alphanumeric character is a space or the start of an equation) AND the current alphanumeric character is a number

THEN current alphanumeric character is integer

IF (previous alphanumeric character is a letter) and the current alphanumeric character is a number

THEN current alphanumeric character is subscript

3.5.9 Understanding What the Equation Means

Once the student has typed in their result there are a number of functions that the system must perform to act like a human counterpart. A human expert will be marking for methodology i.e. is this the correct way to solve the problem in the first instance, has the student made any silly mistakes for example transcription mistakes, forgotten to include the chemical state etc. The diagram below describes the criterion for assessment:

Table 3.5.1 Errors that students may make

	Typographical: silly mistakes	Nonsensical
		Formula: incorrect subscript, letter, or chemical state
	Mistakes made by	Incorrect chemical state
Delineation of	student due to	Performed operation incorrectly
the errors that	conceptual	Incorrect methodology
a student could	misunderstanding	forgot to multiply by -1
make		wrong operation on operand
		forgot to change sign as term was moved to the other side of the equation
		produced incorrect ΔH equation
		inserted incorrect ΔH values
		did not produce correct ΔH value
		incorrect units
		forgot units

A typical solution is given below for the following example and then the functions that a human expert would perform are identified:

Calculate the enthalpy for the reaction:

$$C_2H_4\ (g) + H_2\ (g) = C_2H_6\ (g)$$

at 298 K from the following data:

1) $C_2H_4\ (g) + 3O_2\ (g) = 2CO_2\ (g) + 2H_2O\ (g)$ $\Delta H_{298} = -1395$ kJ

2) $C_2H_6\ (g) + 3.5O_2\ (g) = 2CO_2\ (g) + 3H_2O\ (g)$ $\Delta H_{298} = -1550$ kJ

3) $H_2\ (g) + 0.5O_2\ (g) = H_2O\ (g)$ $\Delta H_{298} = -243$ kJ

A typical solution using simultaneous equations could be:

	1) $C_2H_4\ (g) + 3O_2\ (g) = 2CO_2\ (g) + 2H_2O\ (g)$
+	3) $H_2\ (g) + 0.5O_2\ (g) = H_2O\ (g)$
	4) $C_2H_4\ (g) + 3.5O_2\ (g) + H_2\ (g) = 2CO_2\ (g) + 3H_2O\ (g)$
-	2) $C_2H_6\ (g) + 3.5O_2\ (g) = 2CO_2\ (g) + 3H_2O\ (g)$
	5) $C_2H_4\ (g) + H_2\ (g) = C_2H_6\ (g)$
	$\Delta H_{unknown} = \Delta H_{Eq\ 1} + \Delta H_{Eq\ 3} - \Delta H_{Eq\ 2}$

A human expert would perform the following when marking this solution:

- Has the student identified the correct terms that must be eliminated from the given equations in order to derive the target equation;
- Calculate what the correct answer should be based on the operands and operations the student has selected;

- Check that the correct chemical state has been entered;
- Check that the correct chemical formulae have been entered;
- Check that the correct operations have been selected correctly;
- Check if the resultant operands are in fact proper equations i.e. do they have an equals sign;
- Check that the student has performed the operations correctly;
- Subtract marks for mistakes made;
- Has the student produced the correct equation for $\Delta H_{unknown}$?
- Has the student substituted the correct values for each ΔH?
- Has the student put the correct units?

There are other functions that a human expert knows about equations, which are:

- If one term is "moved" from one side to the other side of the equation the sign changes;
- If a resulting term is zero normally it is not included in the resulting equation;
- If no terms remain on either side of the equation then that side becomes equal to zero;
- If an equation is multiplied by a constant then all of the terms in the equation are multiplied by that constant;
- When adding or subtracting two equations only terms that are the same can be added or subtracted.

If an expert system is to perform like its human counterpart all of these rules have to be explicitly defined.

There is a number of approaches to give the student control over solving the problem. One would be to let them type in anything. This would mean the system would have to cope with many possible scenarios. Feedback would be presented at the end of the question. This type of approach would be similar to a paper-based summative examination. At the other extreme the student would have to follow a set method for solving the problem as presented by a lecturer in a lecture. Feedback could be at every stage in the solution of the problem. This means that the student would always have at the start of each step the correct answer to begin the next stage of the solution. This scenario would be typical of a drill and practice formative assessment. For the purposes of this research it was decided to concentrate on the latter scenario. The reasons being it is the intention to demonstrate that problem solving is possible with computerised assessment. The subscripts for chemical formulae would be generated automatically for the student. This would mean it is assumed that the student understands chemical formula and would not be assessed on their ability to write chemical formula. Instead the main emphasis for assessment is on the student's methodology to solving such a problem and the ability of the system to mark such a methodology.

An outline of the steps that the student would have to perform to solve the problem by simultaneous equations using the designed system prototype would be:

- Select two equations by clicking on the appropriate ones. The system would then automatically provide a copy of these in the work area for the student. This means that the student would not be assessed on their ability to "write down" the equations thus avoiding transcription errors and silly mistakes at this stage;
- Select the appropriate operation to perform on the selected equations;
- Type in the result of the operation on the operands;
- Click the equals sign. This is where the system will mark the student's answer in terms of transcription errors, silly mistakes and methodology. In general terms, the system will work out the correct answer based on the student's selection of operands and then compare this correct answer with that given by the student. If the student has entered a wrong answer for whatever reason the system will provide the correct answer and the student will continue from this point with the correct answer;
- This is repeated with the final operand;

- From the derivation the student then has to compile an equation for $\Delta H_{unknown}$ and provide a value for the enthalpy of reaction.

3.5.10 Rules to provide the expert system with knowledge on determining the correct answer based on student operands.

In order for the system to determine the correct answer based on the student's operands each term in each equation is split into three component (variable) parts - integer, chemical formula and chemical state. The integer variable contains the sign and the stoichiometric amount of the substance. The chemical formula contains the chemical formula of the substance only. The chemical state contains the chemical state for example a gas is represented by (g). A fourth variable is used to record which side of the equation the term is on i.e. whether it is the left-hand side or the right-hand side.

For example in the equation below:

$$C_2H_6\,(g) + 3.5O_2\,(g) = 2CO_2\,(g) + 3H_2O\,(g)$$

the following would be:

integer variables	1	+ 3.5	2	+ 3
chemical formula	C_2H_6	O_2	CO_2	H_2O
chemical state	(g)	(g)	(g)	(g)
equation side	left	left	right	right

The rules that have been defined in order for the system to interpret equations input by a student are as follows:

Rules for integers

IF first character is a letter THEN first integer variable is plus one.

IF first character is a minus sign AND next character is a letter THEN first integer variable is minus one.

IF first character is a minus sign or a number THEN find position of the first letter - first integer variable is (characters from the start of the equation to position of first letter) minus one.

The next integer variable is determined by searching the string for the position of the operation whether it is a plus sign or a minus. This indicates the start of this variable. The end of the variable is determined by finding the position of the start of the first letter thereafter. This is repeated for subsequent integer variables until the string has been searched. An exception is that for the first integer variable following the equals sign. A similar process is followed as previously described except instead of a plus or minus sign indicating the start of the variable an equals sign indicates the start.

Rules for chemical formula variables

When a letter is encountered this will indicate the start of a chemical formula variable. The length of the chemical formula is determined by finding the first bracket, (, character.

Rules for chemical state variables

When a bracket, "(", is found this indicates the start of the chemical state variable, the end of this variable is determined when the closing bracket, ")", is found.

Rules for equation side variables

This variable is determined by finding the position of the equals sign in the equation and comparing the position of each chemical formula to it.

IF position of chemical formula is less than position of equals sign THEN equation side variable is left-hand side

ELSEIF position of chemical formula is greater than position of the equals sign THEN equation side variable is right-hand side.

Performing operations on equations

Once the equations have been split into their component parts it is possible to define the rules that will enable the system to determine what the correct answer should be according to the student's current operands and operations.

One function that the student may wish to perform is to multiply an equation by a constant. Depending on how the specification of the system is defined this could be achieved by two methods. The student could perform the multiplication manually therefore the student is tested on their ability to multiply an equation by a constant: or it could be assumed that the student knows how to do this and have the system perform this operation for the student. The latter can be easily achieved if this is a requirement of the assessment. Since each equation is split up into its component part, by multiplying the integer variable by the constant and then recompiling the equation, the system can in fact be made to automate the multiplication of an equation by a constant if required to do so.

The next function the student will perform is the addition or subtraction of one equation from another. Once the student has typed in their answer and clicked the equals sign the system first determines what the correct answer should be based on the student input.

This can be achieved by finding like chemical formula terms in each equation. The system then has to perform the operation on the integer terms based on a set of rules that tell system how to perform such an operation.

IF sign of integer 1 is plus AND sign of integer 2 is plus AND operation is plus
THEN result equals integer 1 plus integer 2

ELSEIF sign integer 1 is plus AND sign of integer 2 is minus AND operation is plus THEN result equals integer 1 plus integer 2

ELSEIF sign integer 1 is minus AND sign of integer 2 is plus AND operation is plus THEN result equals integer 1 plus integer 2

ELSEIF sign integer 1 is plus AND sign of integer 2 is plus AND operation is minus THEN result equals integer 1 minus integer 2

ELSEIF sign integer 1 is plus AND sign of integer 2 is minus AND operation is minus THEN result equals integer 1 plus integer 2

ELSEIF sign integer 1 is minus AND sign of integer 2 is plus AND operation is minus THEN result equals integer 1 minus integer 2

Once the correct answer has been determined the system then takes the student answer and splits it up into its component parts. These component parts are then compared to the correct answer. The system needs to be aware that if a term is on the opposite side of the equation with its sign changed that this is in fact the correct answer; if a term with a zero integer does not appear in the student answer that this is a correct answer etc. Feedback is then provided to the student depending on if their answer is correct or incorrect. The feedback has been previously determined by the human expert and is stored in the system. If the student's answer is incorrect the system inserts the correct answer so the student will continue with the correct answer. One of the advantages of computerised formative assessment is that the system will monitor student progress and provide feedback at each stage and provide the student with a correct starting point for the next step in the solution. A lecturer with many students would find it difficult to see every student and assist them at every point in their solution to the problem.

Table 3.5.2 Delineation of some of the possible paths in the determination of the heat (enthalpy) of reaction by simultaneous equations.

Possible paths that student may take in trying to solve problem

Step 1	**Step 2**	**Feedback**
Eq 1 + Eq 2	+ Eq 3	Step 1 incorrect. Step 2 never reached
	- Eq 3	Step 1 incorrect. Step 2 never reached
Eq 1 + Eq 3	+ Eq 2	Step 1 correct Step 2 incorrect operation
	- Eq 2	Correct solution
Eq 1 - Eq 2	+ Eq 3	Correct solution
	- Eq 3	Step 1 correct Step 2 incorrect operation
Eq 1 - Eq 3	+ Eq 2	Step 1 incorrect. Step 2 never reached
	- Eq 2	Step 1 incorrect. Step 2 never reached
Eq 2 + Eq 1	+ Eq 3	Step 1 incorrect. Step 2 never reached
	- Eq 3	Step 1 incorrect. Step 2 never reached
Eq 2 + Eq 3	+ Eq 1	Step 1 incorrect. Step 2 never reached
	- Eq 1	Step 1 incorrect. Step 2 never reached
Eq 2 - Eq 1	+ Eq 3	Step 2 incorrect operation
	- Eq 3	If x -1 would be correct solution
Eq 2 - Eq 3	+ Eq 1	Step 2 incorrect operation
	- Eq 1	If x -1 would be correct solution
Eq 3 + Eq 1	+ Eq 2	Step 1 correct Step 2 incorrect operation
	- Eq 2	Correct solution
Eq 3 + Eq 2	+ Eq 1	Step 1 incorrect. Step 2 never reached
	- Eq 1	Step 1 incorrect. Step 2 never reached
Eq 3 - Eq 1	+ Eq 2	Step 1 incorrect. Step 2 never reached
	- Eq 2	Step 1 incorrect. Step 2 never reached
Eq 3 - Eq 2	+ Eq 1	Correct solution
	- Eq 1	Step 1 correct Step 2 incorrect operation

The methodology for solving the problem was determined by manual delineation of the possible paths. Some of the paths are given in Table 3.5.2. This delineation applies only to the equations being given in the pre-set order as shown in the example above. Of course if the equations were randomised these paths would be different. However, the aim of this research is to show that an expert system can in fact be used as an on-line tutor - this application is only a prototype to demonstrate the principle that it can be done.

In step 1 the system has been "given" what the correct solutions are. Solutions where an equation can be multiplied by minus one to provide the correct answer will be ignored to avoid unnecessary complications. Instead only the paths that involve the student selecting and performing an operation on the operands without multiplying an equation by minus one will be considered. From Table 3.5.2 there are twelve possible paths for the student to select from in Step 1. It can be seen from the feedback column in Table 3.5.2 that there are only four possible paths that are correct.

1. C_2H_4 (g) + $3O_2$ (g) = $2CO_2$ (g) + $2H_2O$ (g)
2. C_2H_6 (g) + $3.5O_2$ (g) = $2CO_2$ (g) + $3H_2O$ (g)
3. H_2 (g) + $0.5O_2$ (g) = H_2O (g)

Ignoring the randomisation of the order of the given equations the correct paths for step 1 are: Eq 1 plus Eq 3; Eq 1 minus Eq 2; Eq 3 plus Eq 1; Eq 3 minus Eq 2. Note that there are no correct solutions for positive Eq 2.

These paths can be provided using the following rules:

IF equation 1 plus equation 3 has been input by student THEN this is correct path for step 1

This rule is repeated for the other correct paths.

ELSE for any other combination THEN this path is incorrect

It can be seen from Table 3.5.2 that for the paths that are incorrect e.g. Eq 1 plus Eq 2 step 2 will never be reached. This is because the feedback will indicate to the student that the operation based on the student selection is incorrect (in this example it should be Eq 1 minus Eq 2). Once the student has finished with the feedback the system will automatically correct the operation and ask the student based on this new information to re-input the answer. The system will remark the student answer to determine if the student has correctly performed the operation on the operands.

Once the first step has been correctly done the remaining equation is selected. Again the appropriate rules have been input into the system to enable it to mark the students answer. The system checks step 1 and the student input for step 2. From this information the system can determine the appropriate feedback to relay to the student.

The system checks the student equations to determine if they have input the correct chemical formula. This is achieved by comparing each of the student chemical formulae variables to those calculated by the system.

Once the principle has been shown to be possible i.e. using the computer as an alternative to paper-based formative assessment based on a rigid solution to the problem the next stage would be to enable the student to input any number of steps to solve the problem. This, for example, could be inputting just the final answer or not completing every step in the prescribed solution. In the example solution given above in stage one we input directly the result of the addition of Eq 1 and Eq 3. However, a stage that would be perfectly acceptable before this stage would be the grouping together of the terms before the actual addition was carried out.

	1) C_2H_4 (g) + $3O_2$ (g) = $2CO_2$ (g) + $2H_2O$ (g)
+	3) H_2 (g) + $0.5O_2$ (g) = H_2O (g)
	4) C_2H_4 (g) + $3O_2$ (g) + H_2 (g) + $0.5O_2$ (g) = $2CO_2$ (g) + $2H_2O$ (g) + H_2O (g)
	5) C_2H_4 (g) + $3O_2$ (g) + $0.5O_2$ (g) + H_2 (g) = $2CO_2$ (g) + $2H_2O$ (g) + H_2O (g)
	6) C_2H_4 (g) + $3.5O_2$ (g) + H_2 (g) = $2CO_2$ (g) + $3H_2O$ (g)

The system must have comprehensive rules to deal with any situation that a student may do and they must not be penalised for providing a perfectly acceptable solution to a problem. The key is the elucidation of all the rules from the human expert in the development of a comprehensive knowledge base. It must be remembered that in the prototype developed not all of the rules have input into the knowledge base however, these can be added at a later date.

3.5.11 Conclusion

This sub-section has detailed the steps that were necessary to develop a computerised assessment that extended computerised assessment that were well beyond the multiple-choice question. The prototype that was developed enabled the input of each stage of a solution to the Determination of the Enthalpy of Reaction. There was no guidance from the system as to what the student would have to select - it was up to the student to input each step of the solution into the system. Once the student has input their step the system goes through a series of processes and depending on the choices / input that the student made would calculate what the correct answer would be. The system would then calculate what the answer would be based on the input made by the student. These two answers would then be compared. The system would use a series of rules to compare the two answers. The student would then receive feedback to whether their answer was correct or incorrect. If the student's answer was incorrect the system would provide the student with the correct answer to the step so that they would start the next stage with the correct answer.

The prototype was developed to be used for formative purposes. However, the system could be utilised for summative purposes. The system would mark the student's answer at the end when the student had finished the question. Feedback would be provided after the question had been marked.

3.6 An Assessment Package for Assessing Reaction Mechanisms

It is the aim of this sub-section to describe a further extension of computerised assessment beyond the multiple-choice question that would enable students to draw reaction mechanism schemes within organic chemistry. The student would be able to show the intermediate steps of the synthesis of molecules. Again an expert system would be basis of the application.

This assessment application has not been developed in this research due to time restrictions. However, an investigation into the feasibility of such an application has been undertaken. The steps involved in the development of the application and the application's functionality have also been identified. The rules that the expert system would use to perform the actions of a human counterpart have also been defined.

It is envisaged that this application could be used in either summative or formative assessment. However, the system described is based on the formative mode. The emphasis of the assessment is on student's process skills as well as their product.

The specification of the expert system would be to:

- enable the student to draw chemical molecules and structures easily and readily;
- enable the student to draw curly arrows (of any size) to represent reaction mechanism schemes quickly and easily;
- monitor student responses and react accordingly providing the student with instant feedback.

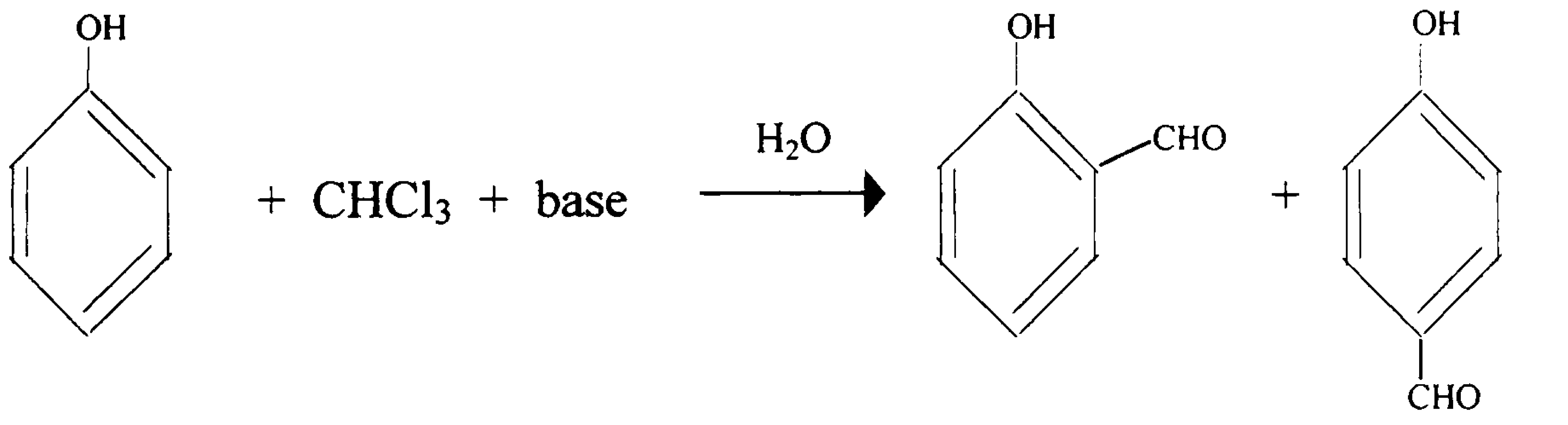
3.6.1 Rationale

The justification to support the claims that an actual expert system of this nature would be possible stems from the fact that applications such as ChemDraw exist today. These applications are routinely used in education and industry to enable the drawing of complex molecules and reaction mechanisms. The drawing of a benzene ring can be done with a click of a button. “Curly” arrows to show the movement of electrons in a reaction mechanism can also be done quickly and easily with a click of button and drag techniques using the mouse.

So we know that the graphical requirements of the expert system can be “easily” obtainable. The main problem of this package is giving the system intelligence to understand and therefore assess a student’s answer. This investigation will centre on a specific problem. Only by analysing a specific problem can the limitations and difficulties be examined. This will then provide evidence as to the success or failure of a system to do the task at hand.

3.6.2 The Reimer-Tieman Reaction

The example that is going to be the focus of this investigation is the Reimer-Tieman Reaction. This reaction is an example of one of the reactions of carbenes i.e. insertion reactions. An aldehyde group is added to an aromatic phenol.



salicyl aldehyde

p-hydroxy benzaldehyde

Figure 3.6.1 The Reimer-Tieman Reaction

This example has been chosen because it is a typical problem at degree level. Students must be able to show the mechanism for this reaction. The mechanism for the formation of salicylic aldehyde is as follows:

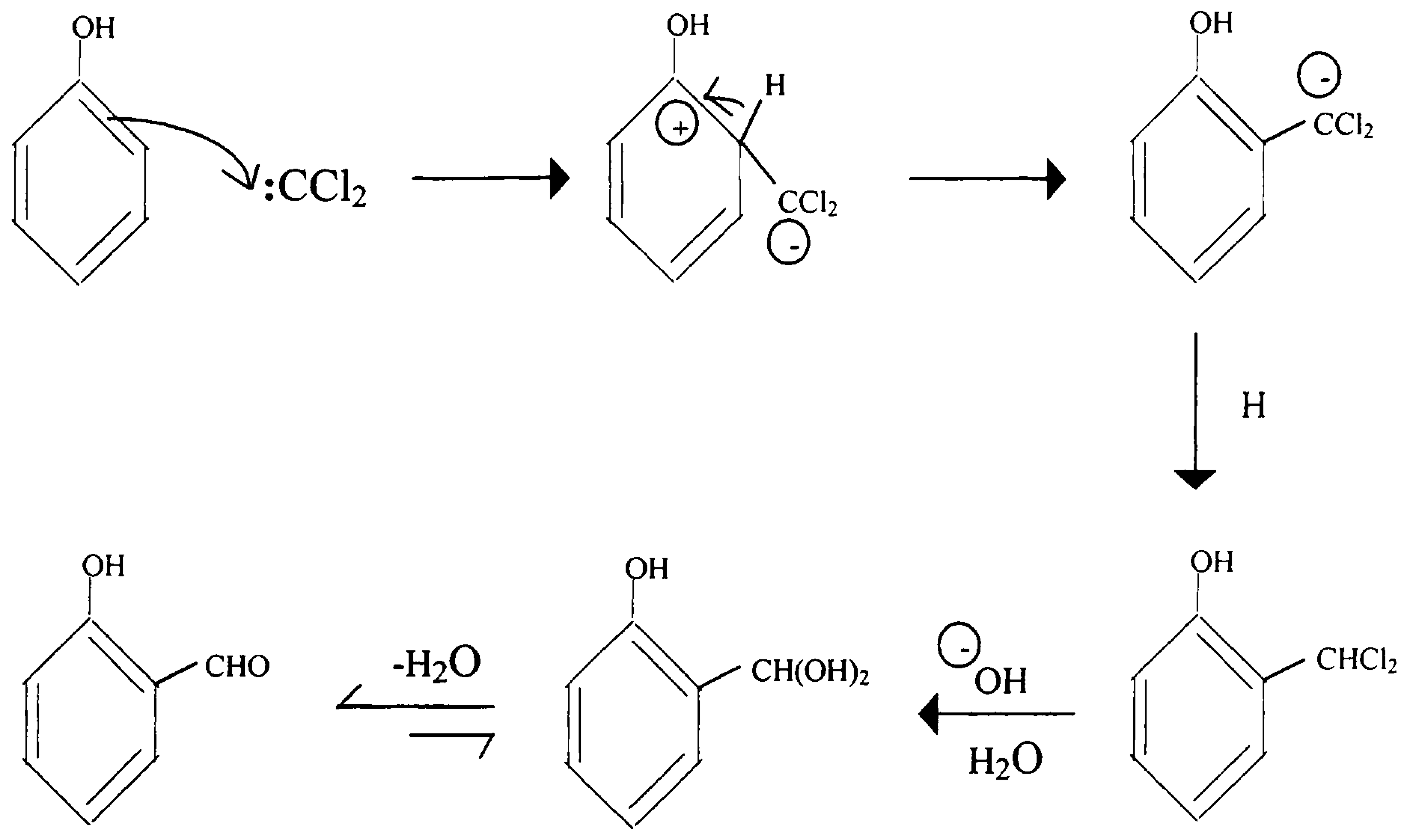


Figure 3.6.2 Reimer-Tieman Mechanism

Within this example there are numerous rules that must be elucidated from an expert to enable the system to "act" like its human counterpart. The system must understand what aromaticity is; how these reactions occur; chemical notation and rules etc. Again in this example it is proposed to use this assessment for formative purposes: students will receive immediate feedback after each step. In the first step of this reaction the mechanism relies on electrophilic substitution on the aromatic ring in the ortho position of the electrophile $:CCl_2$. Electrons are donated from the ring producing an intermediate step where there is a positive charge on the carbon at position one of the aromatic ring. As shown in the diagram the bond from the ortho hydrogen is broken as the electrons are donated from this bond to the positively charged carbon to provide the stability back to the aromatic ring.

3.6.3 Identification of a Solution Pathway

The steps involved for the student to replicate this first step in an application would be:

- The student must select / construct the reactants. This is the first problem. Does the application provide the student with the reactants already drawn on the screen and request the student to draw in the "curly" arrows and draw the intermediate molecule? Or is the student required to select the appropriate reactants from a list of compounds that could potentially be in millions? The answer to this problem depends on what is being tested by this problem and what exactly the question is. Is the student given the overall equation and asked to provide the mechanism by which it takes place? Or is the student asked "Please provide the mechanism by which the Reimer-Tieman Reaction proceeds" and this is all the information that is given to the student. The latter requires the student to have a thorough knowledge of the reactants and products of this reaction. This would mean that, since the student's knowledge of the reaction is being tested, a list of several compounds would have to be supplied in order to avoid giving the student any unnecessary hints. The former provides the student with a "hint" by providing them with the starting materials. This would mean the student would not have to select any starting compounds since they are already given. However, the student would be expected to provide the intermediate compound.

- Once the correct reactants have been chosen the student must understand electrophilic substitution. This requires knowledge of where the electron-rich centres are and where electron-deficient centres are. The student also requires knowledge of the notation that chemists use to show how this mechanism takes place. This is done using curly arrows. The head of the arrow represents the flow of the electrons. If the arrowhead is a single line this means that only one electron is being donated whereas if there are two lines for the arrowhead then two electrons are being donated.

Figure 3.6.3 Representation of the Number of Electrons Donated

- Once the initial mechanism has been shown then the intermediate compound must be given showing the necessary positive and negative centres.

3.6.4 Identification and Description of the Algorithms that *Describe* Benzene

Within chemistry the carbon atoms in a benzene ring are numbered from one to six. This, by convention, is done in a clockwise direction. When a benzene ring has a functional group bonded to it, it is convention to number the carbon atom to which it is bonded as one. The other carbon atoms are numbered from this one, again, in a clockwise manner. Using these rules the position where an electrophile will attack the ring can be established. However, the main technical problem to overcome with respect to a computer application is that the aromatic ring is a drawn image so how does the system interpret where the student has clicked on the ring? If the application enables the student to draw the molecule free hand the OH could be in any position through 360 degrees. This problem is overcome if the system draws the molecule for the student since the position can be predetermined by it.

It is proposed that no matter which route is taken the position of each carbon atom in the ring is determined by "hotspots". Each carbon atom will have an area whose top left co-ordinate and bottom right co-ordinate is represented by the co-ordinates (x_n, y_n) and (x_{n+1}, y_{n+1}) respectively. If the student puts the mouse pointer within these co-ordinates and then clicks on it the system will be able to determine which carbon atom the student has selected.

For example, the following algorithms can give the co-ordinates for a benzene ring:

if the student selects to draw a benzene ring this object is drawn on the screen and it is enclosed by an invisible rectangle:

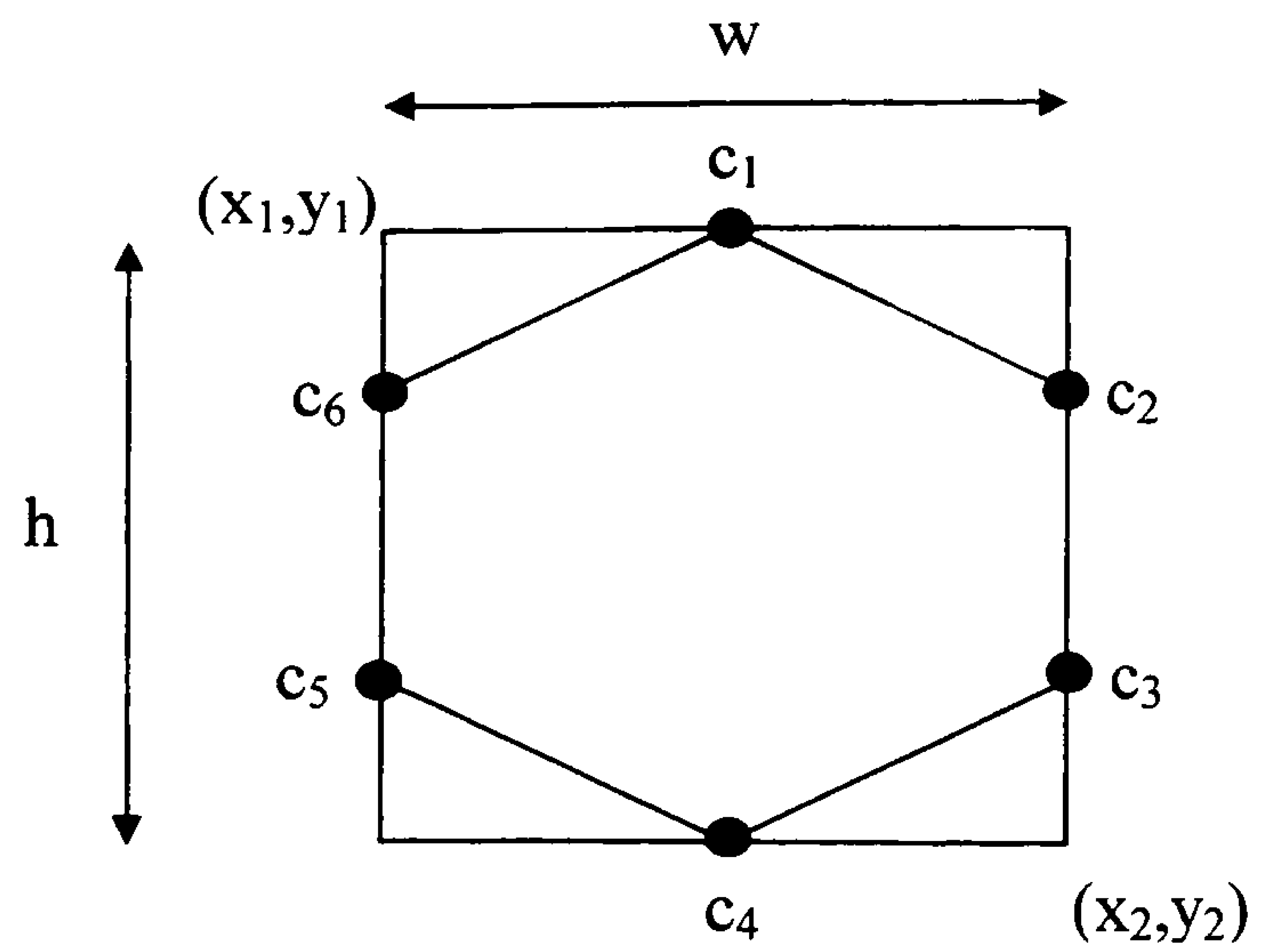


Figure 3.6.4 The Benzene Object

where

w is the width of the rectangle

h is the height of the rectangle

(x_1,y_1) is the co-ordinate of the top left of the rectangle

(x_2,y_2) is the co-ordinate of the bottom right of the rectangle

c_1, c_2, c_3, c_4, c_5, c_6 represents each of the carbon atoms in the ring

w, h, (x_1,y_1) and (x_2,y_2) are all known variables. The following algorithms can express the co-ordinates of each carbon atom:

Table 3.6.1 Algorithms to Describe the Carbon Atom Positions in Benzene

c1 co-ordinates equal $x_{c1} = x_1 + w / 2$ $y_{c1} = y_1$	c4 co-ordinates equal $x_{c4} = x_1 + w / 2$ $y_{c4} = y_1 + h$
c2 co-ordinates equal $x_{c2} = x_1 + w$ $y_{c2} = y_1 + h / 3$	c5 co-ordinates equal $x_{c5} = x_1$ $y_{c5} = y_1 + 2h / 3$
c3 co-ordinates equal $x_{c3} = x_1 + w$ $y_{c3} = y_1 + 2h / 3$	c6 co-ordinates equal $x_{c6} = x_1$ $y_{c6} = y_1 + h / 3$

Note! The co-ordinates (0,0) within VB start at the top left hand side of the screen and not the bottom right hand side as is the case with the Cartesian Co-ordinate System.

These can be easily implemented within the expert system to enable it to determine the co-ordinates of each carbon atom. Other rules can be defined to enable the system to determine which carbon atom the student has selected. This is achieved by creating hotspot areas around each carbon. The locations of these rectangular areas are determined from the co-ordinates of each carbon atom (determined in the previous step). The size of each hotspot depends on the width and height of the original benzene ring. Let's assume the ratio of the width / height is a proportionality constant, k. The top left co-ordinates and the bottom right co-ordinates can be determined with respect to each carbon atom's co-ordinates.

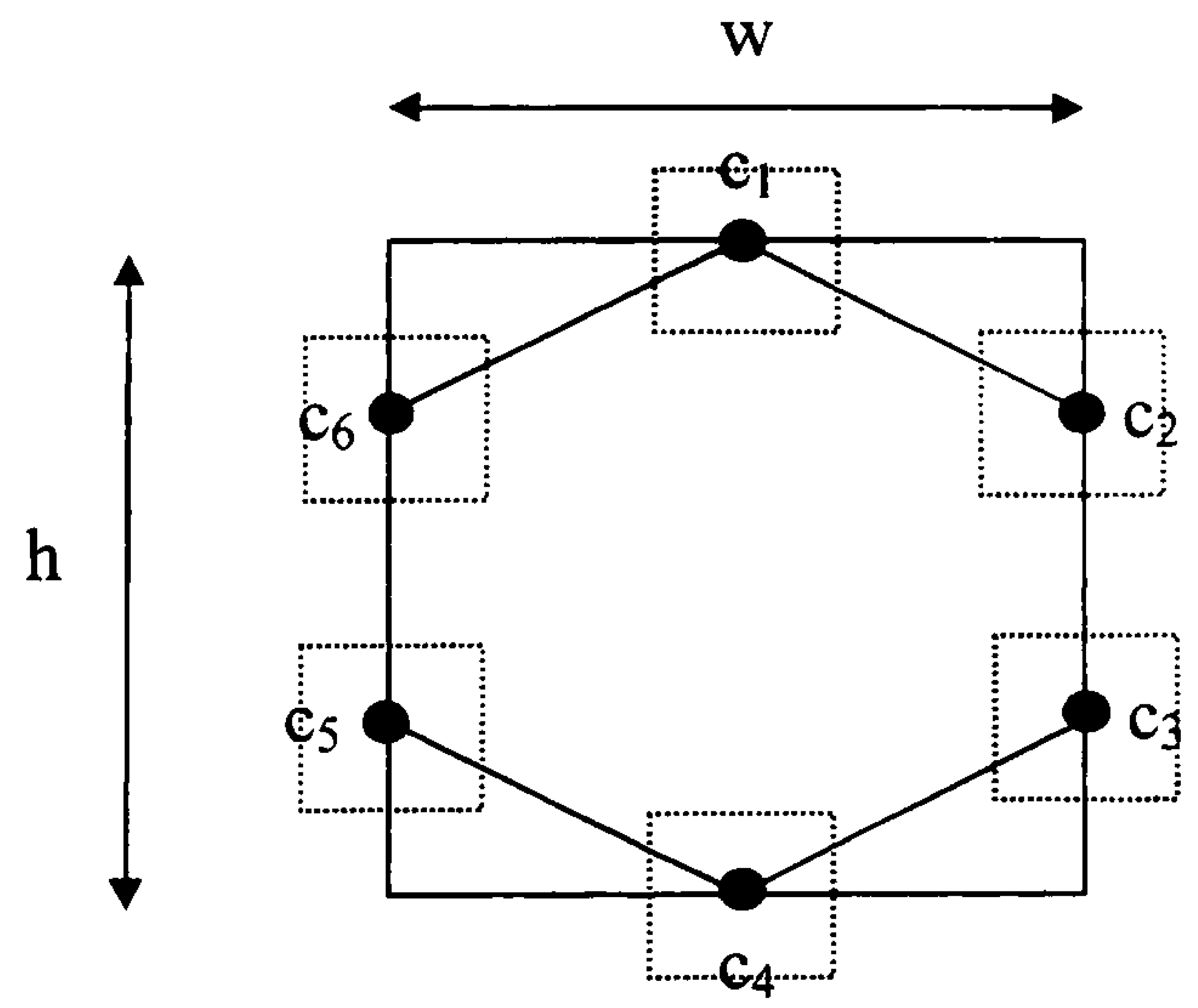


Figure 3.6.5 Hotspot Regions Around Each Carbon Atom

For c1 the dimensions are defined as follows:

c_1 width, (c_{1w}) = k w

c_1 height, (c_{1h}) = k h

c_1 hotspot top left x co-ordinate = $x_{c1} - c_{1w} / 2$

c_1 hotspot top left y co-ordinate = $y_{c1} - c_{1h} / 2$

c_1 hotspot bottom right x co-ordinate = $x_{c1} + c_{1w} / 2$

c_1 hotspot bottom right y co-ordinate = $y_{c1} + c_{1h} / 2$

These rules can be applied for any carbon atom, n as follows.

c_n width, (c_{nw}) = k w

c_n height, (c_{nh}) = k h

c_n hotspot top left x co-ordinate = $x_{cn} - c_{nw} / 2$

c_n hotspot top left y co-ordinate = $y_{cn} - c_{nh} / 2$

c_n hotspot bottom right x co-ordinate = $x_{cn} + c_{nw} / 2$

c_n hotspot bottom right y co-ordinate = $y_{cn} + c_{nh} / 2$

3.6.5 Rules to Determine Which Carbon the Mouse Pointer is Over

The location of the mouse pointer can be readily obtained at any time to determine if it is in the hotspot area of each carbon atom. The rules that would be used to determine if the mouse pointer is within a certain hotspot region would be as follows:

Let

x_{mp} be equal to the x co-ordinate of the mouse pointer and y_{mp} be equal to the y co-ordinate of the mouse pointer;

TL represents top left;

BR represents bottom right.

IF $x_{mp} > x_{c6TL}$ AND $x_{mp} < x_{c6BR}$ THEN

IF $y_{mp} > y_{c6TL}$ AND $y_{mp} < y_{c6BR}$ THEN

Conclude that mouse pointer is over carbon atom six

ELSEIF $y_{mp} > y_{c5TL}$ AND $y_{mp} < y_{c5BR}$ THEN

Conclude that mouse pointer is over carbon atom five

ELSE

Conclude that mouse pointer is not over any carbon atom

END IF

ELSEIF $x_{mp} > x_{c1TL}$ AND $x_{mp} < x_{c1BR}$ THEN

IF $y_{mp} > y_{c1TL}$ AND $y_{mp} < y_{c1BR}$ THEN

Conclude that mouse pointer is over carbon atom one

```
        ELSEIF y_mp > y_c4TL AND y_mp < y_c4BR THEN
                Conclude that mouse pointer is over carbon atom four
        ELSE
                Conclude that mouse pointer is not over any carbon atom
        END IF
    ELSEIF x_mp > x_c2TL AND x_mp < x_c2BR THEN
        IF y_mp > y_c2TL AND y_mp < y_c2BR THEN
                Conclude that mouse pointer is over carbon atom two
        ELSEIF y_mp > y_c3TL AND y_mp < y_c3BR THEN
                Conclude that mouse pointer is over carbon atom three
        ELSE
                Conclude that mouse pointer is not over any carbon atom
        END IF
    END IF
```

By definition of these rules it is possible for the system to determine which carbon atom the mouse pointer is over.

When a functional group is added to the ring, the student will have to identify where it is to be joined to the ring by clicking on the appropriate position. If there are no other functional groups then as mentioned the carbon at this position will be numbered one. If other functional groups exist on the ring then rules will have to be defined as to which functional groups will have precedence over others when it comes to numbering the carbons. The compound is named so as to produce the lowest numbers. The properties of each function will also have to be defined. These can be stored in for example an array, collections or tables. The primary property with respect to this problem is whether the functional group is electron-withdrawing or electron-donating i.e. its electronegativity as this will determine the reaction mechanism scheme.

In the Reimer-Tieman Reaction the first step involves dichlorocarbene bonding in the ortho position in a substitution reaction (carbenes do not undergo electrophilic addition). This involves a double-headed arrow to be drawn from the double bond from carbon one to carbon two to the dichlorocarbene. Again it must be remembered that this could be done anywhere on the screen and in any number of confirmations, therefore the system must be given rules to handle this infinite number of possibilities.

From the Kekule representation of benzene, it has two canonical forms:

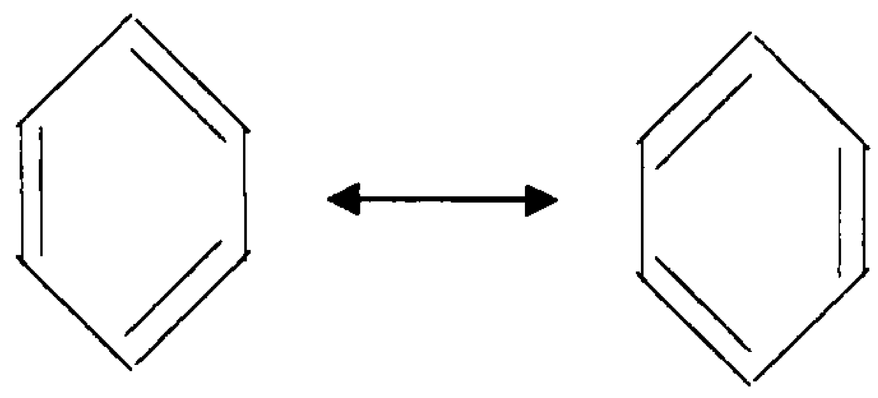


Figure 3.6.6 Kekule Representation of Benzene

Phenol could be represented in a number of ways. For example using the first Kekule structure from the diagram above all the following diagrams would represent phenol:

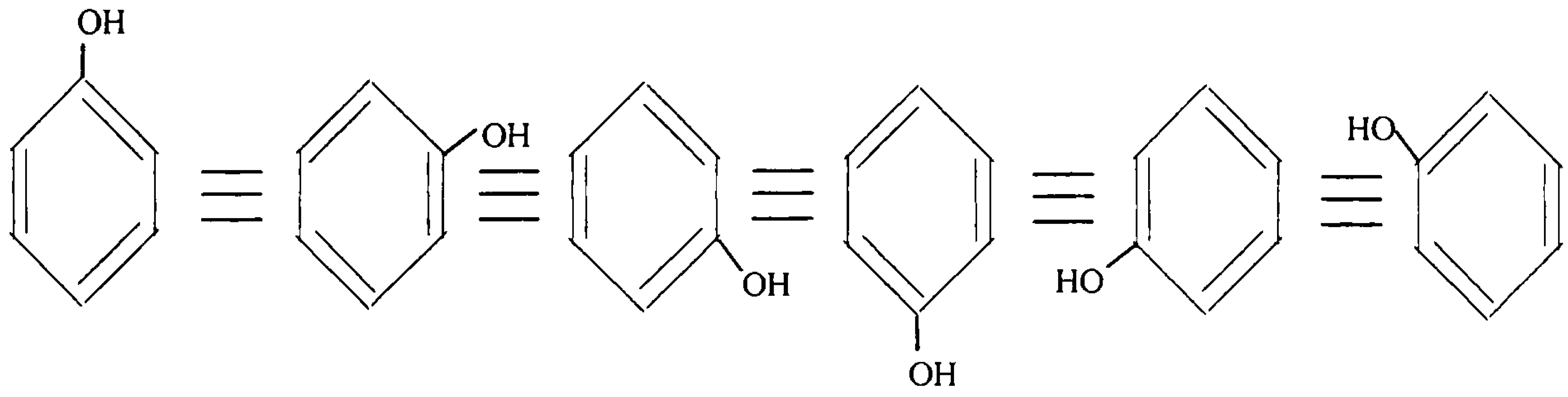


Figure 3.6.7 Some of the Orientations of Phenol

The solution to this problem is to create a phenol object that has algorithms to describe each atom of the molecule using the same method as described with benzene. The molecule will then have a number of hotspots that will enable the identification of where the student has clicked on the molecule.

3.6.6 Monitoring the Flow of Electrons

When the student draws the arrow that represents the flow of electrons this can be "monitored". This process begins when the student selects the arrow tool from a menu or tool bar for example. The arrow is then drawn on the screen using the mouse. When the left-hand mouse button is pressed down this will indicate the starting position of the arrow. When this button is released this will determine the finishing co-ordinates of the arrow. A function will draw the "shape" of the arrow for the student. From these co-ordinates the flow of electrons can be monitored by continual evaluation of where the start and end of the arrow is with respect to the reactant's defined hotspot regions. See Appendix V for further details.

3.6.7 Identification of the Subject Specific Rules

The previous sub-sections have dealt with the identification of how the system will be able to interpret student interactions through the interface. In this sub-section the subject specific rules will be identified and described. These rules will provide the system with the ability to "understand" the chemistry and make decisions on what is the correct answer and compare this to the student's answer. These subject specific rules can be elucidated from a subject expert and / or literature sources. For the purposes of this example literature sources were used because the reaction under investigation is well documented.

In organic chemistry a reaction is between an electrophile and a nucleophile. An electrophile is *electron-loving*. A nucleophile is electron-*donating*. An element's affinity for electrons is governed by the electronegativity that is the ability of an atom to attract electrons. As a result the electrons in a polar covalent bond are shared unequally and makes each atom slightly positively charged or slightly negatively charged. This produces what is called the Inductive Effect. This is the shift of electrons in a chemical bond due to the difference in the electronegativity of each atom. Some of the electronegativity values are shown in the table[88] below:

Table 3.6.2 Electronegativity Values for Some Common Elements

Period	IA		IIA		IIIA		IVA		VA		VIA		VIIA	
1	H	2.2												
2	Li	1.0	Be	1.6	B	2.0	C	2.5	N	3.0	O	3.4	F	4.0
3	Na	0.9	Mg	1.3	Al	1.6	Si	1.9	P	2.2	S	2.6	Cl	3.1
4													Br	3.0

In aromatic chemistry the reaction mechanism can be Electrophilic Aromatic Substitution or Nucleophilic Aromatic Substitution. Benzene is nucleophilic and does not undergo electrophilic addition reactions - it undergoes substitution reactions. However, if a benzene ring has groups bonded to it, these substituted substituents can influence the reactivity of the ring. They can either activate the aromatic ring toward further electrophilic substitution or deactivate the ring toward nucleophilic substitution.

These groups have been classified and the information is readily obtainable. The following table[88] describes the effects that some substituents have on an aromatic ring when undergoing electrophilic aromatic substitution:

Table 3.6.1 Effects of Substituent on an Aromatic Ring

Substituent	Reactivity	Orientation	Inductive Effect	Resonance Effect
-CH_3	Activating	Ortho, Para	Weak, Electron Donating	None
-OH, -NH_2 (+ derivatives)	Activating	Ortho, Para	Weak, Electron Withdrawing	Strong, Electron Donating
-F, -Cl, -Br, -I	Deactivating	Ortho, Para	Weak, Electron Withdrawing	Weak, Electron Donating
-$N^+(CH_3)_3$	Deactivating	Meta	Weak, Electron Withdrawing	None
-NO_2, -CN, -CHO, CO_2CH_3, -$COCH_3$	Deactivating	Meta	Weak, Electron Withdrawing	Strong, Electron Withdrawing

It is proposed to keep these data in a series of lookup tables. By creating a series of lookup tables of the properties of elements and compounds, and defining the rules to enable the system to use this information it will therefore be possible to determine if a reaction should proceed by a certain mechanism or not.

For example applying this information to the Reimer-Tiemann Reaction the following outline describes the pseudo code that would be involved to assessing the mechanism between phenol and dichlorocarbene:

```
IF phenol is a nucleophile AND dichlorocarbene is NOT the same as phenol
    THEN
    Determine where the electron flow is from and to
    IF electron flow is from a double bond THEN
            IF double is in the ortho position THEN
```

```
                IF electron flow is to the carbon in dichlorocarbene THEN
                    Message  "This is the correct answer"
                ELSE
                    Message "The electrons are donated to the nucleophilic
                    carbon in dichlorocarbene"
                END IF
            ELSE
                Message "This reaction proceeds via the ortho position"
            END IF
        ELSE
            Message "For this reaction to proceed electrons must be donated from
            a double bond"
        END IF
    ELSE
        Message "In order for the reaction to proceed the reactant must be different."
    END IF
```

These rules would be expanded but essentially the steps would be repeated for the rest of the mechanism and will not be further described. However, the main result is that a step of a mechanism has been proven to be possible.

3.6.8 Conclusion

It has been the aim of this sub-section to describe how computerised assessment can be extended to include the assessment of reaction mechanisms. Each step of the solution can be assessed using the appropriate rules and tables that have been defined in a knowledge base. Each compound becomes an object with associated properties. These properties can be compared to other compounds in order to determine how the compounds will react.

Despite this application not being developed into a prototype this section has described the functionality of how the system might work. In some instances the code has been developed to determine if certain key functionalities are possible. The conclusion that has been reached is that it is possible to assess reaction mechanism schemes.

Section Four

Contribution

4. Contribution

4.1 The Real Meaning of Computerised Assessment

At the start of this research the term computerised assessment was synonymous with the multiple-choice question. However, one of the main aims of this research was to investigate and demonstrate how computerised assessment was not just restricted to the multiple-choice question. It can be concluded from this research the term computerised assessment should be used as the name of the *set* of which computerised assessment techniques are *sub-sets.* Hence, computerised multiple-choice questions are an example of a sub-set of computerised assessment. Therefore, the term computerised assessment should be synonymous with a range of techniques that can be computerised. With paper-based assessments a range of assessment methods are associated with this term. The same is true with computerised assessment. The reason why this has not been the case is that computerised assessment is an evolving technology. Before this research was undertaken the multiple-choice question was the primary technique. This is no longer the case. This term now refers to a range of techniques. Figure 5.1 below illustrates the relationship between computerised assessment and the multiple-choice question. The boxes without any text have been deliberately left blank. These represent the techniques that will become computerised in the future.

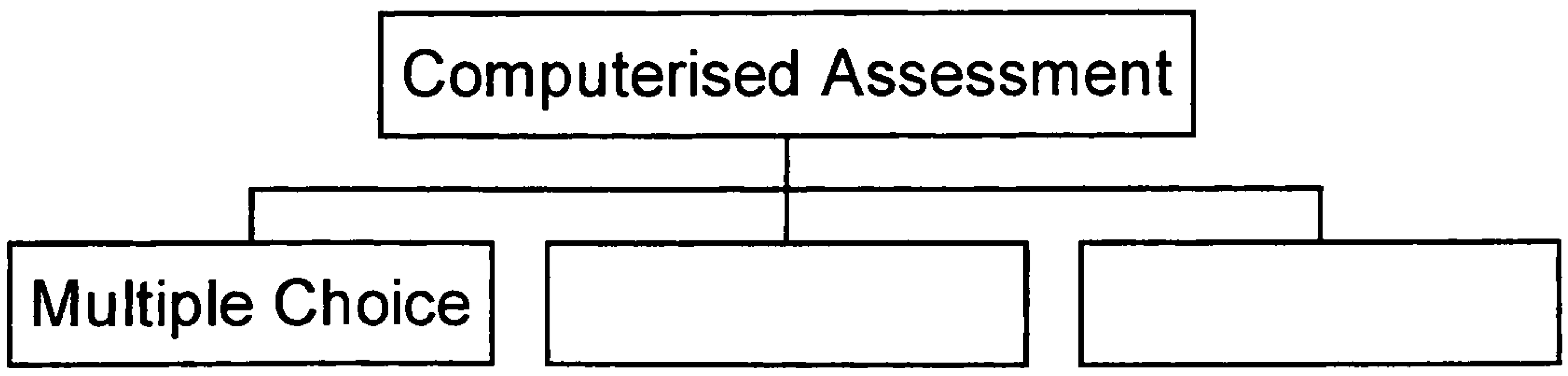


Figure 5.1 The Computerised Assessment Hierarchical Structure

4.2 Problem Solving

This research focused on the sub-set of problem solving questions that were objective in nature, for example, mathematical or numerical. It has been found that other sub-classes of computerised assessment techniques include problem solving that have definite pathways that can be marked objectively. The main point of the term *problem solving* is that the student has to provide a solution to a problem that has several steps. Before this research the multiple-choice would only enable the student to *select* the answer to a question that was of a problem solving nature. The term *select* is used because if the student determines the answer to the problem that does not match any of the pre-selected answers the student still has to select an option from the list. The student has no means to show how they solved the problem and hence arrived at their answer.

With the techniques developed during this research the student would input each of the steps of their solution directly into the computer. The computer would assess if their answer was correct or incorrect at each stage. The students' process skills are now being assessed as well as their final answer.

During this research several computerised assessment techniques were prototyped and described within chemistry. These prototypes were developed to prove that computerised assessment could be extended to include problem solving within chemistry.

4.3 Computerising Current Assessment Methods

There were many situations within chemistry that computerised assessment could be used as a result of this research. It was found that up to 57% of computerised assessment methods currently used on a typical degree course could be computerised to some extent. These assessment methods were essentially objective in nature. Of the 43% that could not be computerised these were skills that the computer can not realistically assess due to the limitations of the current state of technology. The assessment methods that were found to be unable to be assessed were: Essays, Oral Presentations, Personal Interviews and Seminar Performance. These assessment methods require skills that are not yet possible to assess using the computer. These skills involve the ability to listen, comprehend and understand the spoken word. However, it is not a conclusion of this research that the computer will never assess these skills.

4.4 The Interface and Student Interaction

Computerised assessment can be used in those assessments that require the student to solve problems that have definite solutions. This includes all areas of theoretical chemistry - physical, inorganic, organic etc. Within these subject areas there are definite rules which human experts apply when marking student scripts. These rules can be embedded in a computerised assessment as illustrated by the assessments developed during this research project. The primary concern within computerised assessment of a problem solving nature is the interface between the student and the computer. This interaction needs to be as easy for the student as doing a paper-based assessment. This interaction with the current technology (1996) is dependent on the students having good keyboard and mouse skills. However, this research investigated the potential use of voice recognition, which is a fast developing technology, as a third type of interaction that is possible through the computer. It had been noticed that during the development of the applications that interacting with the mouse and the keyboard can be irritating. If mouse interaction was used this can become tiresome by having to move the mouse pointer from the top of the screen to the bottom to do repetitive tasks in succession. Using the keyboard involves "tabbing" through each control. If there are a lot of controls this can be frustrating i.e. the inability of being able to go straight to the control. Accelerator keys can be used for buttons but for text entry fields this can not be the case. Voice recognition is not a mature technology - it is still developing. As a result its application is limited. This is due to how this works which was described in an earlier section. However, it was found during this research that voice recognition could be used to enable the student to perform some of the elementary tasks of the assessment interaction. These include: confirming an answer, moving to the next question, in fact in all the situations that required the student to interact with a button. This would remove some of the frustration of interacting with the assessment with a process that was natural and easy to do without demanding too much from the technology. Voice recognition would be ideal for multiple-choice questions. The only interactions required in these types of assessments is the selection of A, B, C, D or E and moving to the next question. Using voice recognition all of this interaction could be spoken. The only requirement of using voice recognition is that each student would have to "train" the computer to understand the student's voice. This requires the student to speak the word or letter and the computer saves it. This would require a small amount of time depending on the volume of commands that would need to be trained. It was found that for five commands this required minutes, which is not too demanding.

The interaction in terms of inputting a step of a solution to problem did present a challenging problem. Within chemistry a lot of theoretical problem solving involves working with chemical equations and the manipulation of these using mathematical \ numerical rules. Other problems involve drawing intermediate molecules in a synthesis of a target molecule, which require dedicated graphics ability. The ability to do sub- and superscripts was a big problem encountered during the course of this research. This was mainly due to the inability of Visual Basic 3.0 Professional to do sub- or superscripts. Using a third party control solved this problem. This control did not enable the direct input of text. This problem was resolved by making all of the input be done through a text box. However, third party controls would display the formatted text depending on which step the student was working on. Despite this initial problem being solved this led to another problem. How would the student decide when they wanted text displayed as a subscript and how they would return to normal formatted text? The solution was to provide as button as found in word processing packages that would enable the student to turn on and off the sub- or superscript, as they required. However, this was found to be irritating because when inputting a chemical equation almost every other character is a subscript. This then led to fundamental question "What is actually being assessed?" If a student was doing a summative assessment this would be important. If the student was given the formulae but typed in a formula wrong this would be marked incorrectly and considered "careless". This is because the student should be developing the ability to reflect on their answer to make sure that it has been input correctly. The same is true for the characters that make up a chemical formula. Ensuring the letters are in the correct case. Therefore, in summative assessment making the student type in all their answer is important. One suggestion to solve the students having to click a button to indicate that they require the text to be sub- or superscripted is to use voice recognition. The following voice commands would be added: subscript, superscript, and normal. This would provide an easy, natural and reduce the irritation of having to format the text using the mouse or the keyboard.

In formative assessment marking the case of characters or whether the correct characters have been sub- or superscripted properly is not as important as assessing the student's method for solving the problem. It was decided to investigate whether the system could produce automatically the correct character case and the sub- and superscript for the student. This would save the irritation for the student of having to use an award interface and avoid some of the errors that are not as important to assess than the student's methodology of solving the problem.

It was found that it was possible for the system to automatically sub- or superscript characters and to insert the correct case in a chemical formula or if this was not the objective of the assessment it was found that the system could be made not to assess these attributes. In conclusion it is proposed that an options panel containing a series of check boxes could be produced for a lecturer. These check boxes provide details of a range of attributes that a lecturer could chose that the assessment would or would not assess for. In other words a lecturer would be able to select those skills that the system would assess.

Another key area that was investigated in this research with respect to student interaction was that of assembling apparatus. The prototype that was developed used photographs of individual pieces of apparatus and assembled apparatus. The student would click on the photographs to identify joints. The system would "join" the apparatus for the student. This prototype was developed to illustrate the principles of such a package in its look and feel. However, the fully developed application would involve the student having to assemble the apparatus by dragging and dropping the apparatus. The apparatus could be rotated. The key functionality of a package of this nature would be the "joining" of apparatus. Each piece of apparatus would have to "merge" seamlessly without looking distorted. This distortion would lower the overall professionalism of the package. The main problem is that photographs of apparatus are essentially rectangular in shape. Therefore when joining two pieces of apparatus proportion of each photograph will overlap that includes background as well as the actual joint. The technology exists that would enable these photographs to merged. It is proposed that for future development in the assembly of apparatus that further work into the merging of photographs is investigated.

A third area for the interaction between the computer and student within chemistry that investigated was that of reaction mechanisms. Despite this application not being prototyped there was strong evidence to be able to conclude that this is possible. Students can interact with the system to draw reaction mechanisms. However, several issues resulted when determining the aim of the assessment itself. These issues were similar to those that were highlighted in the mathematical problem solving earlier - "What is being assessed by the assessment?". With reaction mechanisms the student is required to draw molecules. Two scenarios resulted from this problem. Should the student be given the structure of the molecule or should they be required to draw it? If the student has to draw the molecule this will be irritating for the student especially if the molecule is aromatic. The student would have to be proficient at using the keyboard and mouse. In conclusion the student should only be required to draw a molecule of a compound at a particular stage of their learning. In a reaction mechanism scheme it has to be assumed that the student has mastered these more elementary skills. The assessment should be concerned with the fundamentals of how reaction mechanisms occur. As a result the student should be given a "tool kit" that they can to select molecules and have the computer draw it for them. The student would be required to "fill-in" mechanism steps.

In summary one of the key problems of having the computer used as an electronic notebook is that of student interaction. It has been found that it is possible to use the computer to input each step of a solution to complex problems in a relatively easy manner. The key to the success is that of good student keyboard and mouse skills. Of course this form of interaction may be limited in the future if technology continues to advance.

4.5 The Computer and Assessing Student Process Skills

Once the hurdle of producing an assessment that enables the student to interact with the software in a logical and intuitive manner, the next problem in the extension of computerised assessment was how the system was going to "understand" the student input. This "intelligence" was solved by producing expert systems. As described in the background theory section 1.4 *Expert Systems* these systems use a series of rules that enable the system to start from an initial problem and search through a series of pathways to find a solution to the problem. It is proposed to use this model to describe how this research applied this knowledge to the problem of assessing students. This research started the investigation of several specific problems and produced prototypes that could be used in specific cases. As a result of this research it was found that it was possible to assess students process skills. However, the shortfall of the prototypes developed is that they can be only used for specific cases. The next step to continue on from this research is to develop a generic application that will mark students' solutions to problems no matter what method they use to solve the problem. However, in order to develop these generic systems the research into how the Determination of the Enthalpy of Reaction that was developed will be described. With the analogy to the tree diagram in 1.4 *Expert Systems* the problem to be solved is the root. All the branches are the possible methods that may be used to solve the problem. All of these possible pathways must be identified and coded into the system as rules. In order for the system to interpret the student input the system must be able to understand the student response and compare it to a correct answer.

This was achieved by identifying the rules that a lecturer would use to assess the student answer. From this analysis it was concluded that by providing the system with the rules of what a chemical formula was, what a balanced equation was and what the rules of algebraic manipulation of equations were, it would be possible for the system to mark the student's answer. This research did identify all of these rules and they were coded into the system. It was found that the system could mark each of the steps identified in order to solve the problem. The system was given the rules to determine what the answer was, based on the students input. The purpose of this was to highlight to the student if they made a mistake based on their input, that their actual methodology was incorrect. This was to encourage the student to reflect on their methodology and perhaps increase their understanding of how to solve the problem. The system also calculated what the correct answer should have been. This answer was then compared to the student's answer to determine if it was correct. This feedback was then given to the student. The system was able to calculate the correct answer because in order for this type of system to work it must be given first principles from which it can work out the answer for itself. The first principles are the actual rules and methods that a student would use to solve the problem. For the Determination of the Enthalpy of Reaction using Simultaneous Equations as the method to solve the problem the key concept in the solution to the problem is the elimination of specific compounds from the given equations to produce the target equation. This was achieved by comparing the components of each equation to the target equation. This would identify which components that had to be eliminated. In order for a component to be eliminated from an equation in mathematical terms the component would have a value of zero.

It can be concluded as a result of this research that by virtue of the existence of these prototypes that all of theoretical chemistry problem solving assessments of an objective nature could be computerised.

This research used text boxes and set areas where students would input their answer for mathematical problem solving. These areas tended to be less than the width of the screen. Future work would include producing a "word processing" style work area that students would input their answer. This would reduce the need for having to use scroll bars when the text was longer than the length of the text box and therefore making the input more easy to read.

4.6 Standardisation of Marking Schemes

The next stage in the development process of an assessment was producing the marking scheme. It was found in this research from a sample of lecturers that their marks do vary. This is one of the disadvantages of lecturers marking student scripts even if the question is objective is that lecturer marking scheme vary. In a problem solving question that is objective in nature there should be a standardised marking scheme which reflects those skills that are fundamental to solving a problem. During the Determination of the Enthalpy of Reaction problem a suggested solution was sent to five subject experts. The questionnaire asked specifically to assign marks to this problem or indicate those steps that they would consider important to solving this problem. The question and solution was as follows:

Typical Problem

Calculate the heat of the reaction

$$C_2H_4\,(g) + H_2\,(g) = C_2H_6\,(g)$$

at 298K from the following data:

1) $C_2H_4\,(g) + 3O_2\,(g) = 2CO_2\,(g) + 2H_2O\,(g)$ $\Delta H_{298} = -1395$ kJ

2) $C_2H_6\,(g) + 7/2O_2\,(g) = 2CO_2\,(g) + 3H_2O\,(g)$ $\Delta H_{298} = -1550$ kJ

3) $H_2\,(g) + 1/2O_2\,(g) = H_2O\,(g)$ $\Delta H_{298} = -243$ kJ

Solution

A typical solution to the above would be:

1. the terms which need to be eliminated from the equations 1), 2), 3) above are O_2, CO_2 and H_2O

2. add equations 1) and 3)

$$1)\ C_2H_4\,(g) + 3O_2\,(g) = 2CO_2\,(g) + 2H_2O\,(g)$$

$$+ \quad 3)\ H_2\ (g) + 1/2O_2\ (g) = H_2O\ (g)$$

to produce equation 4)

$$4)\ C_2H_4\ (g) + 3.5O_2\ (g) + H_2\ (g) = 2CO_2\ (g) + 3H_2O\ (g)$$

3. subtract equation 2) from equation 4)

$$4)\ C_2H_4\ (g) + 3.5O_2\ (g) + H_2\ (g) = 2CO_2\ (g) + 3H_2O\ (g)$$

$$-\quad 2)\ C_2H_6\ (g) + 7/2O_2\ (g) = 2CO_2\ (g) + 3H_2O\ (g)$$

to produce equation 6)

$$6)\ C_2H_4\ (g) + H_2\ (g) - C_2H_6\ (g) = 0$$

4. Rearranging 6) produces the desired equation

$$C_2H_4\ (g) + H_2\ (g) = C_2H_6\ (g)$$

5. Repeat the above steps for the actual values of ΔH_{298} for each equation to determine the unknown ΔH_{298}

$$\Delta H_{298\ \text{unknown}} = \Delta H_{298\ \text{eq 1}} + \Delta H_{298\ \text{eq 3}} - \Delta H_{298\ \text{eq 2}}$$

$$\Delta H_{298\ \text{unknown}} = -1395 + (-243) - (-1550) = -88\ \text{kJ}$$

The results of this questionnaire are provided in the table below. It shows the solution steps 1 to 5 above and the mark assigned by each expert to each step.

Table 4.2 Allocation of Expert Marks to a Solution to a Problem

Solution Step	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5
1	2	3	-	-	3 (1.5)
2	2	2	3	2	3 (1.5)
3	2	2	-	2	-
4	1	1	3	1	3 (1.5)
5	3	2	4	5	8 (4.0)
Total	10	10	10	10	20 (10)

It can be immediately seen from these results that experts themselves have differences in opinions about what are the important skills and the weight of marks assigned to each step in the solution. An important observation was that no expert disagreed with this solution in that they did not say that an important step in the solution was missing. It would be interesting to give these experts a solution to this problem again say with only step 5 in the solution and ask them to reassign marks again. One would wonder if they would even mention the omission of steps 1-4 this time. Computerised assessment would mean that the experts in specific subject areas would have to sit down and identify the key skills that a student must show in order to solve a problem. Of course as mentioned by some experts if the answer is correct to a problem the student will receive full marks. The assumption here is that the student's preceding steps to the solution must be correct to have obtained the correct answer. In computerised assessment the computer will mark every step and therefore not assign marks for those steps that the expert has failed to see or has not had time to look for.

4.6.1 Summative Versus Formative Computerised Assessment

The main conclusion that was drawn here was that summative assessment was more difficult to implement than formative. With formative assessment it does not matter if the students ask each other questions or see the next student's monitor, or if the software crashes in the middle of doing the assessment. However, in summative assessment these issues do matter. In summative assessment the student must be confident using the computer and the software for the assessment. In the questionnaires that were returned by the students doing the summative Chemical Information and Retrieval Assessment, in this small sample of students it was found that a third of them still suffered from computer anxiety. If summative computerised assessment is to be implemented this must be reduced to zero if the test is to be considered fair. The main problem / issue of this summative test was that of security. All the advantages of computerised for summative tests were lost during the inability to write to the network because of network security issues. Floppy disks had to be used and therefore this would lead to an administration problem. The other issue here was that if student results were to be recorded say to a database then students would be tempted to cheat. In the small sample of students that were tested there was one instance of cheating. So the system must be robust if summative computerised assessment is to be successful. There needs to be the support of the IT department to help solve these technical issues.

So in conclusion it was found that formative assessment was the easiest to implement and suffers the least problems.

4.7 Multimedia

It was found that the application of multimedia within chemistry enabled the assessment of skills that were not previously possible with the computer that had no multimedia capabilities. These skills were of a visual nature. Video clips can be played and questions can be asked about the clip. This was done for photographs in this research. The students were being questioned on their knowledge of apparatus and chemical literature using photographs. Photographs were also used to build apparatus. This application of multimedia was the main extension of computerised assessment beyond the multiple-choice question. In the former examples that tested student knowledge multimedia only extended the subject of the question i.e. the text-based question could now be replaced by a photograph or a video clip on which the question could be based. The questions were essentially extended multiple-choice questions. The computer still has the inability to let the students enter their own words. The main problem that was found with video was that it is not possible to interact with the video clip in a manner other than stop, play, fast forward or rewind. However, photographs were found to have more interaction capability than video. Photographs were found to have the ability to be interacted with. This increased interactivity resulted in computerised assessment being extended to include the ability to assemble apparatus using colour photographs, which was not previously possible.

The main issue that merged with using photographs on a standard VGA screen 640x480 was that screen space was at a premium. If was found that there was minimum limiting size that photographs could be displayed on the screen before they became unrecognisable. However, there was other screen elements that also required space. The assessments that were developed had the following common screen elements - an area in which to display the question and present the feedback, a work area for the students to input their answer, tools that enabled the student to confirm their answer, manipulate the screen and specialist tools that were essential for the student to input their answer. For example the Fractional Distillation Assembly needed an apparatus tool kit to select the appropriate piece of apparatus, zoom in out buttons for viewing photographs, confirmation buttons to indicate that they had finished and the system could mark their answer, and a Next button that was used when the student had finished the question and would proceed to the next question. All of these screen elements were competing for space. In order to enable the photograph to be seen at its best the photograph display areas had zoom buttons that when pressed would zoom the photograph in or out. When the photograph was zoomed in the display area size did not change. This area would display a zoomed in photograph but would have scroll bars to enable the scrolling of the photograph. It is a recommendation of this research when developing multimedia applications that VDUs should be as large as possible 17” is recommended and use a minimum screen resolution of 800x600.

It was found that producing multimedia images and sound is of high quality and easy to do in simple cases. However, it was found that when the images became technically challenging for the taking of photographs of apparatus specialist knowledge and equipment is vital to producing high quality images. It was also observed that if these images were poor in quality impressions of the software were poor. However, if the quality of the images were high the software was received more favourably. It can be concluded that if the software is to be presented externally to other institutions unless good resources and the expertise are available the production of high quality multimedia should be produced by the experts.

4.7.1 Guidelines for the Development of Computerised Assessment

This research has concluded that it is a possible to extend computerised assessment beyond the multiple-choice question. However, the main issue that this raises "Can a lecturer within chemistry with no real computing skills develop their own computerised assessments?". The answer to this question depends on what type of computerised assessment the lecturer wants to develop. If the assessment is a simple multiple-choice then the lecturer would be able to develop their own computerised assessments with little or no computing experience. This type of assessment can be developed using commercial packages such as Question Mark that enables a wide range of objective questions to be asked. For example, multiple-choice questions (True / False, Yes / No, Agree / Disagree) single and multiple response, push button multiple-choice, hotspots on images. Randomisation of the question is also built-in.

A series of reports can be produced automatically for example a List Report that contains a report on the test including the user name, the score and date of each answer file selected; a Summary Report that produces a report of maximum, minimum and average score and time of completion of selected answer files; a Full Report that produces a report giving details of the answers to each question.

Question Mark can also be used to produce a Question Analysis. This involves producing a report analysing each question as cumulative answer data. This essentially details the number of students that answered the question correctly and incorrectly.

Question Mark does not require the lecturer to have any previous computing experience to produce an assessment. It can be done easily and quickly.

However, for the assessment methods reported in this thesis and the findings that all of the theoretical problem solving assessment within chemistry could be computerised is an enormous and demanding task. This will require detailed computing expertise and really should be undertaken using a low-level language. The work involved will require much investment in resources and time using a professional team of developers. This is looking at the task as a whole. However, this task does consist of many smaller sub-units. Each of these sub units can be coded independently from each other. It is recommended in conclusion of this research that these sub-units should be identified and coded as objects. This would follow a bottom up design as opposed to a top down design. This future work should involve an object-orientated programming approach. This bottom up approach involves identifying each sub-unit. The sub-unit has an existence that is independent from other units. Once these units have been identified and developed they can be used as objects that have certain properties and methods associated with them. The idea is that once they have been coded once they can be used in any application and therefore do not have to be continually re-invented. An example of how these objects would be used in assessment would be for identifying a solution to a problem. Therefore, in the example of the Determination of the Enthalpy of Reaction using simultaneous equations to solve the problem the method developed to solve equations and assess the students solution would become a method of an object. This object could be an object that has methods to assessment certain mathematical / chemical problems. The important point is that the object does not have a scope that is limited to the problem of Enthalpy of Reaction this object method can be re-used in any situation that uses simultaneous equations.

4.7.1.a) Assessment Analysis

This is the most important step of the whole cycle of the development of an assessment. Within this step the learning outcomes are identified that the assessment will be designed to test for. Therefore this analysis determines which assessment method will be used. The questions that must asked within this step include:

- What type of knowledge (i.e. the learning outcome) is being tested?
- What types of assessment are available?
- What type of knowledge does each assessment method test for?

After the answers for these questions have been determined, the assessment method that should be used should be known. It could be that one assessment method is sufficient or that a number of different methods combined are better to test for this learning outcome.

The next step in this analysis should be then to ask the question:

- Could the assessment method be potentially computerised?

  If the answer is "Yes" to this question then proceed to the next stage in these guidelines. If the answer is "No" then the assessment method must be designed and implemented in the traditional manner.

Once it has been identified that the assessment could be computerised then the next question to ask is:

- Is this computerisation based on techniques that already exist?

  If the answer is "Yes" to this question then this assessment could be developed using a software package that has already been designed for computerised assessment purposes.

  If the answer is "No" then this assessment method will have to be designed from scratch probably using a low level programming language to provide the flexibility that is required to design an application to the required specification.

Following the "Yes" path to the question above would mean that using the technology of today (1996) would enable assessments of the multiple-choice format, drag and drop and hotspot question types to be designed. Much of the "programming work" has been removed using programs of this nature to enable a person with low computing experience to develop computerised assessment. This type of package empowers the subject expert to develop assessment applications.

The rest of this section will focus on the development of computerised assessments with respect to the scenario the above question being "No". This would mean that the method that was required would not be in the "mainstream" of computerised assessment techniques.

4.7.1.b) System Analysis

Once it has been decided to develop a computerised assessment method a physical description of what the system is required to do must be done next. Within this description it must be highlighted what it is the system must do with respect of how the student will interact with the software, how the student answers are to be stored and how the application will achieve its functionality. Other issues as to the development environment, how the software is to be implemented what software should be used. It may be that after this analysis that it may not be possible to develop a computerised assessment method. Or that the technical expertise or hardware do not exist in the department. Or the time available or available resources do not warrant further development. Time spent on this analysis is the most important step in the whole process - a thorough analysis would highlight if it is worth further development and therefore saving both time and money.

These various steps can be summarised by the following categorisations:

- Physical Design

  The physical design describes how the system will be implemented, the hardware that will be used etc.

- Logical Design

  The logical design describes how the system will look and function.

With respect to the development of computerised assessment the Logical Design is the more important to this thesis. It will be described further.

4.7.1.c) Logical Design

A description of some the more important elements of the logical design will be described:

- Description

At the start of the analysis and physical description of the system must be done. This will describe exactly the requirements of the system and what it must do.

- Screen Shots

  This will illustrate what each of the intended screens must look like.

- Screen Navigation Diagram

  This diagram will show how all of the screens in the system are linked together. It will show the pathway of how each screen is called; which screens come before it and those that can be called after it (if any).

- Structure Diagrams

  These diagrams show exactly the controls and their names that are on each screen shot.

- Pseudo-Code

  The pseudo-code describes the functional of the system i.e. what happens if the user clicks this button, which controls should be enabled and those that should be disabled.

- Database Design

  The database design describes the structure of the database. It identifies the tables and fields that will be include in the database. The fields will be described in terms of what data type will each field accept and where appropriate the size of the field and a description of the purpose of the field.

If after this analysis it is concluded that it is technically and financially viable then the next stage in the development is the actual development of the system.

4.7.1.d) System Build

In order to develop a novel and new computerised assessment method the authoring tool will have to be a low-level language. This is because the functionality of the system probably will not have been commercially developed and therefore, this functionality will have to be developed from scratch using a language with the power to enable this functionality to be developed in the first instance. For example C++ would be such a development tool. However, this requires technical expertise. Tools like VB are easier to use and make the interface easy to develop however, building the functionality will require knowledge of C++ and the Windows API functions.

The first step in the development of the system is to build a prototype. This step is normally done to provide a "walk through" to give a feel of how the application will look and feel. It is often used to show to a potential funding organisation in order to obtain funding for the project.

The basic stages involved in the development of the application are as follows:

- Requirements Confirmation

  This refers to the time spent in consultation with the subject expert to ensure that the system is being built the way the subject expert has intended.

- System Design

  This refers to the documentation of the system. Any system that is built must have appropriate documentation to describe what each part of the code does. This will help in maintenance. The programmer may not be involved will the application after it is built. Things may go wrong with the code and somebody else may be required to fix the problem. Having good documentation on the system is vital in ensuring that this maintenance is easy and quick to do.

- Code and Test

  This refers to the actual coding and testing of the code as the system is built.

- Error Handling

  The code that is built in the previous step may have errors that are unforeseen. This may cause the system to crash i.e. a fatal error has occurred. It is important that the system can adequately cope with these errors and any other errors that may occur. Time is also needed to ensure that the system is well designed for error handling.

- Help

  Providing help to the user is of utmost importance in order to help them use the application. It is important to note that the time that help would be of most benefit to the user is when the system is first released to them. It is important that good help is developed before the system is released. It must be stressed that the development of good help is a demand task and should not be under estimated.

4.7.1.e) Project Team

In order to develop an assessment a team of experts is required. In the simplest form the project team would consist of a technical expert and a subject expert. However, for multimedia then a multimedia expert would be required to develop the video or the audio or take the photographs i.e. this could expand to several specialist experts. So identification of the skills required for the development of the project is important.

4.7.1.f) Summary

In order to develop computerised assessments the first decision that needs to be made is can computerised assessment be used to assess the intended learning outcomes. The next decision to be made is, is there enough time and resources to develop a new computerised assessment technique. In conclusion in order for a lecturer to develop a computerised assessment with respect to the least learning curve and development time they must use software that is commercially available and designed specifically for assessment purposes.

4.8 Conclusion

As a result of this study it was found that:

- Of the assessment methods currently used in a tertiary level chemistry course 57% could be computerised to some extent;
- Computerised assessment techniques could be greatly advanced beyond the multiple-choice question format;
- Computerised assessment should be implemented in formative mode to ease problems of giving student write access to the network;
- For computerised assessment to be successful it must have the support of all the department;
- Computerised assessment can be increased to assessing not only the students final answer but also their skills;
- The time to develop complex computerised assessments is large relative to the multiple-choice format;
- In order to develop complex computerised assessments a dedicated team of developers would be required;
- In order to develop high quality multimedia assessments specialist skills and dedicated hardware are required;
- Computerised assessment does not (yet) enable all the skills of a chemistry student to be assessed;
- The following areas that favour further development of computerised assessment within chemistry are all chemical mathematical \ numerical problem solving exercises, balancing equations, predicting which reactants will react with each other.

Before this research started the main computerised assessment technique was the multiple-choice format or a hybrid of it. Computerised adaptive tests have been developed that adapts to the ability of a student depending on whether the student got the answer correct or incorrect. Despite this new technique it was still based on the multiple-choice format. This research has contributed to knowledge by extending computerised assessment within chemistry beyond the multiple-choice format to enable the student to input their answer as they would a paper-based test and receive immediate feedback. The student is not prompted by asking them to select their answer from a list. They have complete freedom to input their own solution to a problem and they can show the individual steps of the solution. The student therefore can be assessed on their methodology i.e. process skills as well as their product skills. This removes the main criticism of the multiple-choice question, which is guessing.

There were two areas which this type of computerised assessment were developed as prototypes. These areas were Determination of the Enthalpy of a Reaction and Assembly of Typical Fractional Distillation. These areas are very different. The former deals with mathematical \ numerical problem solving within chemistry. This prototype was designed for a specific problem due to the uncertainty if it was possible in the first instance. Now that it is proven to be possible for one such problem solving exercise it will follow that for most, if not all, of the problems of a similar nature, then it is possible to develop assessments for all these others.

The latter prototype enables students to build a typical fractional distillation apparatus using colour photographs. Again this assessment enables students to make a free choice with out prompting which answer may be correct, therefore eliminating guessing. For the first time students will be able to build a piece of apparatus using real photographs of the apparatus, therefore seeing what the apparatus really looks like. Again student process knowledge will be assessed. Since this prototype has been developed for one instance it will be possible to do this for any other scenario of a similar nature.

A third assessment of this nature i.e. enabling the student to input an answer as they would in a paper-based test has been investigated and outlined with respect to the functionality of how the application would work. A prototype has not been developed. The area of chemistry, which this investigation was under taken, was Reaction Mechanisms.

The computer has now been turned into an *electronic notepad* enabling the student to solve problems and receive immediate feedback in real-time.

The data that supports this hypothesis is the actual prototypes that were developed. The very existence of these applications is proof that it is possible. It must be stated that these prototypes are very basic and are in no way fully developed robust applications.

Appendix I

Learning Outcomes

Learning Outcomes

The following learning outcomes have been compiled from a typical chemistry degree course at tertiary level.

Overview

"Students will gain a knowledge of the principles and techniques of pure and applied chemistry, they will develop a full understanding of observation and interference, they will combine their analytical and creative skills in the investigation of new subject areas and in the design of new processes."

Graduate knowledge

"1) possess a wide factual knowledge in all branches of chemistry.

- know the theories of chemical bonding and apply them to inorganic and organic compounds

- recall the properties of the elements and relate their properties to their position within the periodic table

-explain chemical reactions in terms of kinetic and thermodynamic parameters

- use physico-chemical laws to make predictions about chemical systems

- use systematic nomenclature to name chemical structures

-identify functional groups within organic molecules and predict the reactions of compounds

-know synthetic methods for inorganic and organic compounds

-recognise and classify different types of biologically important compounds and explain their significance in life-processes

2) have a knowledge of the principles and techniques of analytical chemistry.

- consider an analysis as a complete process and apply analytical methodology

- be able to select an appropriate technique for the separation of mixtures

- apply instrumental methods to the analysis and identification of chemical compounds

- describe the use of atomic and molecular spectroscopy for qualitative and quantitative analysis

3) either

understand the chemical and physical principles of industrial chemistry.

- describe the unit operations used in chemical technology

- explain the use of instrumentation in process control

- design plant equipment for industrial processes in which chemical or biological changes take place

- discuss the necessity for modifying synthetic routes when laboratory based reactions are scaled up

- apply the principles of experimental design to process optimisation

or

understand the key features of biochemistry

- demonstrate the relationship between structure and function at the cellular level

- explain the role of enzymes in metabolism and discuss the relationship between the structure of enzymes and their properties

-discuss the roles and biosynthesis of DNA and RNA

- describe the principles of genetic engineering and its role in biotechnology

- understand the uses and mechanisms of pharmacologically active compounds

Practical Skills

1) planing and organising laboratory experiments

- carry out COSHH assessments

- handling and disposing of chemicals in a safe manner

- calibrating analytical equipment

- appreciating the influence of experimental error on quantitative results

- selecting and setting up appropriate apparatus

- choosing representative samples in analytical procedures

- understanding the chemical reactions in the process being considered

- using statistical techniques for the optimisation of yield in chemical reactions

2) executing a chemical reaction or analysis

- acting responsibly in relation to individual and group safety

- carrying out a synthetic reaction

- making and recording observations in a logbook

- recording measurements and experimental data

- weighing accurately and distinguishing between accuracy and precision

- making up solutions of known concentration

- isolating and separating reaction products

-analysing a product using an appropriate analytical technique

- carrying out qualitative and quantitative analyses on a range of samples

3) analysing and presenting data

- presenting and summarising data in a clear manner

- determining confidence limits for parameters

- using regression analysis for determining relationships between variables

- testing hypotheses

- using computer packages in data analysis and the presentation of reports

4) developing ideas in project work

- specifying the project topic from an idea

- searching the relevant literature and collecting related data

- devising appropriate experimental methods, recording results and analysing data

- interpreting the results, and drawing conclusions and making predictions

-producing a written report

Personal Skills

1) thinking and the scientific method

- developing thoughts and ideas through to experimental investigation

- interpreting results and reaching conclusions

- drawing together relevant information from different areas of chemistry in problem solving

- designing experiments to test hypotheses

2) retrieving and handling information

- using library resources, locating primary research literature and carrying out on-line searches

- reading effectively by choosing the appropriate reading material and abstracting key points

- recalling important items of frequently used factual knowledge and knowing the sources of less commonly used material

- recording experimentally determined measurements

- using chemical databases and molecular modelling program

- writing computer programs for data handling

3) communicating orally and in writing

- compiling written reports of practical work

- writing essays and answering examination questions

-selecting, integrating and preparing material from a variety of sources

- developing and expressing personal opinions

- preparing adequately short presentations to a peer group, by choosing quantities of relevant material in the proper order and at the appropriate level

- using audio and visual aids in presentations

4) listening and responding

- concentrating on what is being said and developing attentiveness

- taking appropriate notes

- thinking constructively while listening

- asking questions at the appropriate time

- following up lecture material by reading around the subject

5) co-operating with others in group tasks

- learning to work as part of a team by taking a fair share of the burden

- making compromises when appropriate

- developing leadership skills by taking the role of group leader, motivating the group and taking responsibility for their activities

- obtaining assistance from others when solving complex problems

6) developing independence

- managing time efficiently by advanced planning

- prioritising work in order to optimise productivity

- meeting set deadlines for the submission of coursework

- developing confidence in own ideas

- developing from a passive to an active learner

- accepting criticism and responding positively

- working alone in the laboratory and outside the teaching environment

- developing self-motivation and tenacious working practise

- becoming more resourceful, flexible and adaptable"

Appendix II

INTERFACE DESIGN –

THE GRAPHICAL USER INTERFACE

INTRODUCTION

One of the main problems that computer users have had to endure when running new applications is persevering with the ritual cognition of that particular application's unique syntax, semantics and pragmatics.

Little attention has been paid into the research of the psychological and physiological principles behind screen design in order to produce a universally accepted code of practice.

This has given software companies a vita on the standardisation of interface design in the hope that their software will eventually become the accepted standard.

In recent years, however, there have been a number of screen design guidelines released such as IBM's System Application Architecture Common User Access (SAA-CUA) (1987, 1989a, 1989b). The International Standards Organisation is one of several bodies currently producing standards[80]. These have been produced in an attempt to finally bring about some degree of consistency and guidance in screen design.

With the advent of the graphical user interface (GUI) a number of interfaces have been designed with the aim of providing consistency throughout all the applications within the confines of that particular operating system.

For example Apple's Macintosh, DECwindows which runs on VMS and UNIX operating systems, and Windows which runs on personal computers (PCs)[80].

The main emphasis of this paper is based on the PC's Windows GUI for this is the system with which the author uses. However, these guidelines can be applied to the design of any application.

The philosophy behind the GUI is to try to reproduce one's "desktop" and to make it resemble the "real world" as much as possible. To reduce the learning time when running different applications the interface is common to all of these applications i.e. once the screen elements for one screen have been learnt, they have been learnt for all other screens. Note, however, this does not apply to each applications actions and commands which have to be learnt separately.

In order to reproduce one's desktop each application runs in a "window". A number of windows can be opened at any one time, thus enabling one to switch from one application to another without having to exit that particular application and open a new application. Each window in effect sits on top of each other just like a pile of papers sitting on one's desk. Thus, a "3-D" environment is created which is not possible in the traditional "2-D" textually based applications.

In the real world one symbol can be worth many words and once learnt humans can readily interpret the symbol's meaning. Symbols (or icons as they are referred to), are used in a GUI to minimise the textual information on a screen which could otherwise produce information overload on the user.

SCREEN ELEMENTS

A comparison of the screen elements and techniques commonly found in textual and graphical screens are shown below taken from [80]:

Table 1 Graphical vs Textual Screen Elements

Graphical Screens	**Textual Screens**
Title	Title
Screen ID	Screen ID
Headings	Headings
Captions	Captions
Entry/Data Fields	Entry/Data Fields
-	Function Key Listings

-	Command Fields
Messages	Messages
-	Blinking
Scrolling	Scrolling
High/Low Intensity	High/Low Intensity
Upper/Mixed-Cased Char	Upper/Mixed-Case Char
Normal/Reverse Video	Normal/Reverse Video
-	Underlining
Action/Menu Bars	-
Pull-Downs	-
Pop-up Menus	-
Buttons	-
Window Style	-
Direct Manipulation	-
Indirect Manipulation	-
Radio Buttons	-
Check Boxes	-
Value Sets	-
List Boxes	-
Spin Lists	-
Attached Menu Boxes	-
Mice	-
Icons	-

Multiple-Character Styles	-
Multiple-Character Sizes	-
Thin/Thick/Double Rulings	-
Foreground Colours	-
Background Colours	-
Colour Lightness Differences	-
Proportion	-
Oval Shapes	-
Rectangular	-
Shapes	-
Scalloped Corners	-
Bevelled Edges	-
Drop Shadows	-
Shrinking/Growing	-
Motion	-
Control-Menu Box	-
Toolbar	-

As one can see there are many more associated screen elements with a graphical user interface than a textually based screen. These elements (derived from design principles) provide a rich environment with which to design powerful interfaces.

IN A "2-D" ENVIRONMENT

Obviously one can not produce a real 3-D environment from 2 dimensions, however, we can fool the eye that what it is perceiving is 3-D.

The are various rules for giving the impression of 3-D.

Closer objects should be lighter, clearer, and larger and if they are moving they should move faster. Whereas, distant objects should be darker, blurred, smaller and if the object is moving it should move slowly. Parallel lines always appear to be converging.

For an object or text one can put a drop shadow behind it to create the illusion that the object etc. is "floating" in mid air.

One can easily create a 3-D graph/frame/panel/button/box/text effect. By varying the shade of the shadow (i.e. highlighting or lowlighting) one can produce the effect that the object is above (i.e. "raised") or below (i.e. "inset") the plane of a particular level.

Using a bevelled edge and a drop shadow on a screen element is a really effective method in producing 3-D. A bevelled edge is non-right-angle lines to the border of the screen element [80].

TEXT & LEGIBILITY

It is generally accepted that reading from a VDU screen is slower than reading from a book by up to 30%, less accurate and causes more eye fatigue [80-83].

Typography, colour, size, spacing, number of words on screen all affect text legibility and eye fatigue.

It is recommended that the typeface used for text be plain, non-stylish with clear ascenders and descenders [81]. Ornate fonts should not be used if possible [80,81]. The font generally accepted as the most legible is the "sans serif" font when the point size is below 8 or the VDU is poor [80, 84].

Only one type of font should be used throughout - emphasis can be achieved by varying the style of the text i.e. regular, bold, italics outline, shadowed. However, no more than two styles should be used at once [80] and for highlighting purposes bold or changing the colour should be used as opposed to italics or underlining - these make the text more illegible [82].

No more than three font sizes should be used in any one screen [80] and the font size should create a visual angle greater than 20 minutes of an arc [81].

Once the typeface, style and size have been used they should be used consistently throughout the application - this will help the user to learn the conventions used in the application more readily.

The letter case of the text is also an important factor that affects legibility. It has been found that upper-case text is generally harder to read in long pieces of text than lower-case text by up to 20% [81]. This is because the ascenders and descenders are hidden in upper-case text whereas in lower-case text they establish a word's shape. For the main body of the text it is recommended that lower case should be used, while reserving upper for short captions like main headings.

Limiting the number of words per line to 8 - 15 and 3-5 lines per paragraph has been found to produce the optimum legibility [81]. The number of characters per line should be between 60 – 80 [85].

Legibility can be further increased by using positive contrast (black text on a white background) [81].

COLOUR

Using a colour-coding scheme in an application can be a powerful tool to indicate conventions that are used within it. For example, a further explanation of a word could be achieved through hotwords which would always have the same colour, or a screen could be divided up into key areas of group related fields by that area's colour.

However, colour must be used cautiously because:

- certain colours may have a certain meaning to different cultures already. For example, red on a traffic light means stop. Telling a user to stop in an application using a green circle would be totally confusing to them and therefore increase their learning time of the application;
- too much colour can be distracting, decreases legibility, increases eye strain and fatigue, and also increases the learner's cognition.

It is generally regarded that 4 to 7 colours is the maximum number of colours that should be displayed at any one time[80, 85].

Properties of Colour and the Human Eye

In order to understand why colour can lead to increased eye strain and fatigue, one needs to know a little more about the properties of colour and the human eye.

There are three properties that make up any colour, which are:

1. hue;

   Hue is the colour's wavelength and is the property, which we see as colour.

2. chroma or saturation;

   Chroma (or saturation) is in the range from the brightest shade of the colour to grey. In other words it is the colour's purity.

3. intensity (or value).

   Intensity (or value) is the colour's brightness property e.g. light red or dark red.

In the visible spectrum (wavelength range from 400 - 700 mu) the different colours are separated when they pass through the eye's lens. The longer wavelength colours are focused further back than the shorter wavelength colours. However, since the retina is at a fixed distance from the lens only one colour can be focused on it at any one time. Hence, the other colours (which have a different wavelength than the one being focused on) appear to be out of focus. This continual process of refocusing each colour leads to eye fatigue and strain.

In effect people see red as being further away than blue which can give a 3-D appearance to objects. However, blue can not be focused on the retina; its focusing point is always in front of it - therefore it is impossible to obtain a sharp blue image.

The eye is most sensitive to yellow light. In general the eye is more sensitive to colours in the middle of the spectrum and less sensitive to those at the extreme ends.

The Right Colour

Human Associations

Examples of common associations that we have for colours are:

- Red - Stop, fire, hot, danger
- Yellow - Caution, slow, test
- Green - Go, OK, clear, vegetation, safety
- Blue - Cold, water, calm, sky, neutrality
- Grey - Neutrality
- White - Neutrality
- Warm colours - Action, response required, spatial closeness
- Cool colours - Status, background information, spatial remoteness

Colours are also used to depict emotional situations:

- High illumination - Hot, active, comic situations
- Low illumination - Emotional, tense, tragic, melodramatic, romantic situations
- High saturation - Emotional, tense, hot, melodramatic, comic situations
- Warm colours - Active, leisure, recreation, comic situations
- Cool colours - Efficiency, work, tragic and romantic situations

(Taken from [1] who is citing Marcus, Aaron, Ten Commandments of Colour, Computer Graphics Today, 1986B)

Position

The position of colours on the screen is also an important consideration when using colour.

The eye is most sensitive to red and green in the centre of the retina and to blue, yellow, black and white at the retina's edges. Therefore it is important when the eye is focusing on an object to use these colours correctly to aid legibility.

Foreground and Background

The foreground colour should be a warm colour while the background should be a cool colour such as blue. Text should be taken from the centre of the spectrum such as white, yellow, and green.

While colour can be used for these screen elements, the eye can not distinguish fine detail when colour is used. In order for the eye to resolve fine detail grey scale should be used.

Five shades have been recommended by Marcus, Aaron, Ten Commandments of Colour, Computer Graphics Today, 1986B [1]:

- White - Screen background

  Text located in any black area
- Light Grey - Push-button background area
- Medium Grey - Icon background area

  Menu dropshadow

  Window dropshadow

  Inside area of system icons

  Filename
- Dark Grey - Window border
- Black - Text

  Window title bar

  Icon border

  Icon elements

  Ruled lines

I recommend light grey should also be considered for use as the screen background colour because a white background can be tiring (i.e. it is too bright) to read after long periods of use.

There have been a number of studies conducted to investigate which are good foreground/background combinations. For example, Lalomia and Happ (1987) set out to determine which colour combinations were the most effective for an IBM colour screen [80]. Using 16 foreground and 8 background colours a combination of 120 colours were tested in terms of:

"1) response time to identify characters;

2) subjective preference of users".

They found that:

1) the best foreground colours were bright;

2) low contrasting foreground/background combinations were poor;

3) best background colours were found to be black and blue;

4) poorest background colours were brown and green.

Bailey and Bailey (1989) have summarised the best foreground / background combinations which is based on research such as Lalomia and Happ (1987) [80].

Background	Acceptable Foregrounds
Black	Dark Cyan Light Green Dark Yellow Light Cyan Dark White Light Magenta Light Yellow Light White

Blue	Dark Green Light Green Dark Yellow Light Cyan Dark White Light Yellow Light White
Green	Black Light Yellow Dark Blue Light White
Red	Light Green Light Cyan Light Yellow Light White
Magenta	Black Light Cyan Light Yellow Light White
Yellow	Black Dark Blue Dark Red
White	Black Dark Blue

Aesthetic Screen Properties and the Presentation of Data

When humans look at buildings, shapes or objects they often like or dislike it purely on the basis if it "looks nice". This aesthetic quality applies just as much to the design of screens. How data is presented on the screen can affect they way we perceive it, if it looks good or not, and also the ease with which it is read.

Aesthetically Pleasing Shapes

There are a number of length:width ratios of areas which have been found to be aesthetically pleasing such as those given by Marcus (1988) cited by [80]:

- Square (1:1);

  Square root of two (1:1.414). Has been used as the standard size of paper in many countries such as the British A4 size;
- Golden Rectangle (1:1.618). Used in Greek architecture (5 century BC);
- Square root of three (1:1.732);
- Double square (1:2). Used in Japan. For example, in their tatami mat.

Data Presentation

Start in upper left corner

In Britain we read from left to right and from top to bottom. It is therefore natural enough to assume to that the starting place for data should be in the upper left corner. A study carried out by Sreveler and Wasserman (1984) cited by [80] found that objects located in the upper left corner were found more quickly than objects in the bottom right.

Be consistent - reserve key areas of the screen for certain information

Certain areas of the screen should be reserved only for key pieces of information. For example, in MS Windows the control bar, title, minimum/maximum icons, menu bar, toolbar, scroll bars and information have all been reserved. This means that the user will always find them in the same thus, decreasing the time to learn the interface - once it has been learnt once they will be in the same place for all applications.

Other aesthetically pleasing properties identified include: [80]

Balance (vs. Instability)

If a picture is crooked on a wall we will realign it so that it looks balanced. The same will be true for a screen if it looks unbalanced. "Realignment" in terms of screen design involves making sure the elements are equally weighted to the left and right, and from top to bottom.

Regularity (vs. Irregularity)

Regularity is defining where each line will start on a screen in terms of rows and columns and keeping to this throughout the application.

Symmetry (vs. Asymmetry)

One side is a mirror image of the other i.e. there are symmetry axis. Note symmetry will lead to balance but balance can be achieved without symmetry.

Predictability (vs. Spontaneity)

Ability to predict what another screen will look like from seeing part or all of a screen.

Economy (vs. Intricacy)

Summary of data in a highly efficient manner using few screen elements.

Sequentially (vs. Randomness)

Logical ordering of screen elements to make it easy to read the information.

Unity (vs. Fragmentation)

The grouping of screen element to make them seem as one.

Simplicity (vs. Complexity)

This is the ease with which information can be retrieved and understood by the user. The more times the eye has to realign itself at the start of a line of data, for example, if the start of each line of data is not left justified but starts randomly in any position, the longer it will take to read/search the screen i.e. increased complexity.

Related data should be grouped together, the starting point should be the same for each new line of data, for example, left justified.

A simple method to quantify simplicity/complexity is to count:

1. the number of fields on the screen, plus;
2. the number of horizontal alignment points, plus
3. the number of vertical alignment points[80].

The resulting figure is the complexity of the screen.

Therefore, the less number of alignment points on the screen the easier it will be to read / search.

Groupings

If data is presented line after line, finding data will take longer than if the data was presented into logical groupings that had a caption / heading at the top of the data. This produces a screen that can be more meaningful to the user. They can search these captions quickly and therefore locate the appropriate information much more quickly.

Faster screen searches from the use of groupings has been confirmed by research Dodson and Shields, 1978; Hauber and Neumann, 1986; Tullis, 1983; Triesman, 1982 cited by [80].

It has been suggested that the following should apply to groupings:

> *1) There should be 5-7 rows after which space lines should be included;*
>
> *2) The width of the row should be 11 to 15 characters (this is obviously dependent on the type of information being displayed)* [80].

This has been derived from research and work on visual acuity.

Visual Acuity

Visual acuity is the detail that is in focus when the eye focuses on it. For example, when the eye is focused straight-ahead objects to the side are blurred i.e. not in focus.

> *"... relative visual acuity is approximately halved at a distance of 2.5 degrees from the point of eye fixation (e.g. Bouma, 1970). Therefore, a 5 degree diameter circle centred around an eye "fixation" character on a display has been recommended as the area "near" that character (Tullis, 1983) or the maximum length for a displayed word (Danchak, 1976)."* [80]

It is possible to calculate the number of characters that will lye in this 5 degree circle. Assuming a VDU distance from a user is 19", average character size and spacing etc.

it has been calculated that 88 characters will lye in the circle which works out at 7 rows and the longest line width is 15 characters.

This means that the eye should be able to search a 5 degree grouping with one eye fixation which decreases a screen search time. Groupings which require more than one eye fixation i.e. greater than 5 degrees will take longer to search.

Human Perception

Human perception is another important attribute (as well as visual acuity) which has to be considered in screen design.

Perception is the ability to learn, understand, integrate and apply one's knowledge of the objects all around us. In everyday life one is confronted by many stimuli. Our senses decide which are important (called signals) and which are not (called noise) in order to react to them if necessary.

The following are also perceptual characteristics:

- Proximity - if objects are close to one another we see them as being together;
- Similarity - if objects have the same colour, size, shape etc. we see them as being related to one another;
- Matching Patterns - if objects have the same pattern we will respond to them in the same way no matter if they are larger or smaller;
- Closure - if a shape/object such as a circle is not quite closed we will perceive it as closed anyway;
- Balance - we prefer things balanced to be aesthetically pleasing.

In terms of the grouping of information, these perceptual characteristics can be used to enhance screen layout. (However, remember that sometimes these characteristics may conflict with each other. In these instances common sense must prevail.)

The proximity characteristic is the most frequently used. The rule is to make as much space as possible from the grouping of information i.e. never have no space on the screen.

Using the similarity characteristic each grouping can be emphasised or using a colour-coding scheme, bold etc can highlight related information.

By applying the closure and matching patterns characteristics this will allow related information to be displayed in windows, line borders (do not exceed three line thickness) [80] or using graphics.

Screen Density

Once information is put onto the screen the overall density (i.e. the number of characters on the screen compared to the number which in theory could be placed on it) should be 25 - 30 percent [80].

Density is directly related to complexity in that density is a measure of characters and complexity is a measure of fields. Therefore, as the complexity increases the density should also increase.

Research has shown that as the density increases information retrieval increases and so does the error rate.

Another type of screen density is local density. This is a measure of how tightly packed the visual acuity circle is. Within this circle characters have a weight factor. Adding up all the weight factors will give the local density.

Screen Elements

As mentioned previously there is no standard interface as yet. However all the GUI's all appear to be converging in that the principles of screen design are being used throughout each manufacturers' interface. There are now a growing number of screen elements in the GUI, which are becoming common to all GUI configurations.

Here the emphasis will be on the Windows PC interface.

In Windows the application is event-driven i.e. the application will remain idol until the user does something, for example, clicking the mouse over a button. Every Window has common elements such as:

- Title;
- Control Bar;
- Maximise/Minimise Buttons;
- Menu Bar;
- Tool Bar;
- Scroll Bar;
- Information area;
- Screen body or client area.

Once the user is familiar with these the user will have learnt their purpose for every other Windows application.

In Windows there are functions called dynamic-link libraries (DLLs), object linking and embedding (OLE) and dynamic data exchange (DDE). These features allow users:

1. to get routines which have been written other Windows applications therefore saving the time rewriting the code again;
2. to implant, for example, a drawing which had been done using a painting application into a word processing document;
3. to retrieve and use data from, for example, a database into an application without writing the data out again.

Visual Basic is Microsoft's Windows programming application. This will allow anyone to create their own Windows application which uses the above functions, and provides the user with commonly used objects which are found in the graphical environment such as buttons, text boxes, frames etc. This will save the user a lot of time since they will not have to redraw these objects time and time again.

Objects Commonly used in GUI

Visual Basic 3.0 Professional provides the following objects already drawn for the user [87]:

Table 2 Common Visual Basic Controls

Check Box	Combo box	Command button
Common dialog	Data	Directory list box
Drive list box	File list box	Frame
Grid	Horizontal scroll bar	Image
Label	Line	List box
OLE	Option button	Outline
Picture box	Shape	Spin Button
Text box	Timer	Vertical scroll bar

In addition Visual Basic (Professional) provides standard 3-D controls:

Table 3 Common Visual Basic 3-D Controls

3-D Check box	3-D Command button	3-D Frame
3-D Option button	3-D Panel	

There are also other controls available.

The user now has a wealth of objects with which to create powerful and stunning applications. However, the problem still remains in the utilisation of these objects. Too many graphics in an application can clutter the screen hence, making it hard to use, and if they are not drawn properly can be annoying to look at.

To help avoid appearance of screen clutter lines, borders, frames and panels are used. These objects group the data but there is also space inside the parameters, which satisfies our proximity perception characteristic.

This technique has been applied to menu/action bars. These are drop down menus which have related commands from which the user can chose from. The border should surround the data in a drop-down menu window by a minimum of two spaces [80].

If borders, lines were omitted the screen would be very hard to use. However, when using lines or borders it is important to keep these as thin as possible and minimise the number of borders displayed on the screen at any one time. Thick lines and numerous borders can increase screen clutter and hence, decrease the screen's clarity.

Menu Bar

The menu/action bar is located at the top of the screen below the title of the screen. Depending on the number of choices that are contained within the menu bar will determine the number of rows that it covers. However, the maximum number of choices displayed in the menu bar should be eight[80].

In fact, Microsoft are now concentrating on producing document-centric programs instead of application-centric programs. This will mean the menu choices will be the same in each application (as well as toolbar icons).

Single, mixed cased words should be used to describe the choices. These choices should be arranged from left to right with the most frequently used at the left.

The Help choice should be right justified to the far right hand side.

The first menu choice and Help choice should be 1 space from each of the margins. There should be three spaces between each choice.

The menu bar is separated from the rest of the screen by using a different coloured background or using solid lines across the width of the screen above and below the choices.

The choices in the menu bar can be selected either by the mouse or the keyboard. In order to select the appropriate choice by keyboard mnemonic code keyboard accelerators are used. These should be as unique as possible. Each accelerator is highlighted by underlining the appropriate letter that is to be pressed.

No matter if the menu choice is selected by keyboard or mouse, the choice is highlighted using the same colour as the screen title. The highlight should be 1 space before and after the choice.

Pull-Down Menu

Once the choice has been selected the pull-down menu should be directly below the choice which is aligned with the highlight position of the menu bar choice. The pull-down menu choices are surrounded by a border. There should be 2 spaces between the pull-down choices and the border edges.

The choices should be mixed case. No more than 10 choices should be presented with the most frequently used at the top. Related choices are grouped and separated by drawing a line above and below. The choices should be left aligned.

Scroll bars and instructions are forbidden in pull-down menus.

Keyboard accelerators are used as for menu bar choices. If alternative keys have been assigned to a choice for example <alt> + <insert> they should be right aligned. 3 spaces should be left between the longest choice name and the alternative keys. A "+" should be included in between the keys with no spacing.

Pull-down choices, which lead, to another pull-down menu should have a right-pointing triangle after the choice description. One space should be left between the choice and the triangle. The colour of the triangle should be the same as that for the choices (i.e. black).

For pull-down choices which lead to a pop-up window these should be indicated by have placing three dots (...) (ellipsis) (same colour as choice) immediately after the choice.

Pop-Up Menus

These are menus which "pop-up" on the screen when, for example, the mouse is in a certain area of the screen or the user clicks the right mouse button in a certain area of the screen. A series of choices is displayed to the user. They follow the same rules for pull-down menus in terms of spacing between border and the longest text, keyboard accelerators, text colour, highlighting, justification etc.

However, pop-ups should;

1. appear directly below the area where the pop-up was called from;
2. have a centre-justified title displayed in upper case or if it is in mixed case it should be in a larger font or be bolder than the choices text;
3. have a different but complementary background colour to the screen background.

Buttons

As previously mentioned, buttons are designed to be 3-D deliberately so they can resemble "real" buttons. In the real world people are accustomed to using buttons so it is a logical step to include buttons in a screen interface; people will not be afraid of using them and will be aware that the button is to be "pressed" to select it or make something happen.

Buttons can be of the 2-D square-corned rectangle or rounded-corner rectangle. However, for added screen enhancement 3-D buttons should be used. 3-D buttons will change colour when depressed. This gives the impression that the button has actually been depressed.

Buttons can be located in two positions on the screen:

1. centred at the bottom of the screen arranged horizontally (most frequently used at the left) with a minimum of 2 spaces between each;
2. down the right hand side of the screen arranged vertically below one another if there is no room available at the bottom (most frequently used at the top).

In both cases related buttons should be grouped together where appropriate.

The button caption (one word if possible) should be fully spelt out using mixed case including accelerators, ellipsis (...) where appropriate.

It is essential that once a common action has been assigned a caption that this should be used throughout the application. For example, IBM guidelines recommend [80] that:

- OK - should be used when changes within a window are confirmed and the window is closed.
- CANCEL - should be used when closing a window without the application accepting the changes made within the window.

- HELP - should be used when there is help available to the user when they position the cursor over a screen element.

There are other standard Windows common names for common actions. These should be used when applicable. This is because changing names, with which the user is familiar, will increase their time to learn your interface that will lead to frustration.

Once a button caption has been assigned the button should always occur in the same position on every screen in which it is used.

All buttons should be the same dimensions.

When the buttons are presented on the screen one button should be assigned as the default. This is immediately obvious to the user because the button's border will be bolder than the others. The button which is assigned as the default is, for example, OK for a positive action or CANCEL for a destructive action such as deleting a file.

Windows

A window can be as large or small as possible within the VDU dimensions. It can be resized, moved around, or displayed in conjunction with other windows either "tiled" or "overlapping"/ "cascading" with one another. The common elements of a window have been mentioned earlier. Four types of windows [80] are:

- Application or primary window

  This is the first window that appears when the user enters the application.

- Document, supplemental, or secondary window

  This window is usually the window in which the user types in or processes data.

- Dialog boxes

These are boxes which ask the user to fill in certain parameters, or make further selection of choices, or to set up device settings such as the printer which is attached to the PC.

There are two types of dialogue boxes:

1. Modal: When this box appears on screen the user must fill in the options and close it before they can interact with the rest of the application.
2. Modaless: When this type of box is opened it enables the user to interact with windows and other dialogue boxes without having to close the box first.

- Active windows

  This is the window that is currently active for manipulation. This window should be in the foreground and visually different from other windows to indicate that it is the currently active window. Methods used to differentiate it are:

  - highlighting the title bar;
  - different border or background colour.

There are two types of interaction [80] within the Windows environment:

1. Direct manipulation

   This occurs when the interface is set up just like the user's desk would be. All of the objects on the screen would be familiar to the user. The objects would one would found on the user's desk. Hence, the user would know how to use these objects already. Any unfamiliar objects etc. would be hidden from view of the user so as not to confuse them.

2. Indirect manipulation

   It may be impossible to have a system that is totally directly manipulated. For example, there may be too much information to be displayed on the screen. This then has to be put in the form of a menu bar with pull-down choices. It will probably be too much for the user to remember. When this results the user is using indirect manipulation.

Scroll Bars

Scroll bars contain an area within which a slider box or handle can move to indicate how much area has been scrolled and how much area there is still to scroll. At each end of the scroll area box a directional arrow button is located.

The user has the option to depress the directional arrows to scroll or slide the slider box in the appropriate direction.

Vertical scroll bars should be located at the right-hand side of the window and should be equal to the height of the scrollable area of the window. Horizontal scroll bars should be located at the bottom of the window and should be at least half the width of the scrollable area of the window.

Entry Fields

For single lined entry fields the label or caption should be fully spelt out, in mixed case and located to the left of the text box.

A colon should be placed immediately after the label.

A space should separate the colon and the text box.

For multi-lined:

- data, alphanumeric entry fields the label should be left-justified one line above the text box.
- fixed length fields the label should be centre-justified.
- numeric variable length data fields the label should be right justified.

For a group of different single lined entry fields the labels can be:

- left aligned with one space between the colon of the longest label and the text box;
- right aligned with one space between each colon and the text box.

If the entry fields are arranged in columns there should be a minimum of 5 spaces between the longest text box and the adjacent label.

For vertical arrangements there should be a single space line every five to seven rows.

Selection Fields

These come in two forms. One allows the user only to select one option - the most common is called radio buttons. The other allows the user to select one, two or more parameters - the most common is called check boxes.

Radio Buttons

The graphic representation of a radio button stems from the older style of wirelesses in which the user could only push one button at a time to select a radio station.

The radio button consists of a transparent circle when the button is not selected; when it is selected a smaller solidly filled circle is inserted into the centre of the transparent circle.

A fully spelt out label or caption (mixed case) is located to the right separated by one space. Each radio button is aligned vertically. However, if there is not enough room they can be aligned horizontally with 3 spaces between each option.

The maximum number of radio buttons that can be displayed at one time is eight [80].

Check Boxes

A check box consists of a transparent square box when it has not been selected; when it has been selected a "X" appears in the box to confirm that it is selected.

The same rules apply for the label, separation, alignment and number of boxes displayed as the radio button.

List Boxes

List boxes display choices to the user which they can select using the mouse or keyboard. These choices are arranged in list form.

A maximum of eight choices should be displayed at any one time.

When there are more choices than eight, the list can be scrolled using scroll bars.

The list is arranged alphabetically, in mixed case, usually in black and left justified.

A border is used to group the list and should be in the same colour as the text. There should be one space between the border and the left most characters and the furthermost right character of the longest word.

List boxes have a caption located left justified above the top border of the list box.

Again the caption should be mixed case in the same colour as the list choices text and should be worded the same as the label of an associated text box.

When a list box is chosen the highlighting colour is the same as that for the active title bar colour. The highlight spans from the left border to the right border.

Drop-Down List Box

This is another way to display a list of choices while saving screen space.

The drop-down list box consists of a single selection field. An arrow pointing down is at the right-hand side of the box. Once the user clicks it with the mouse or using the keyboard a list of choices is displayed directly below. Scroll bars are used if there are more choices than the maximum of eight, which can be displayed at any one time.

The text format, highlighting, separation etc. is the same for the list box above.

However, the drop-list box can be the same width as the single selection field.

In both the list box and the drop-down box the last choice selected by the user should become the default which is displayed in the single selection field.

Combination box

When a list box is attached to a text box it becomes a combination box. This means that the user can type text into the text box as well as being able to select an option from the list.

The list box is attached to the text box but the left vertical border of the list box should be indented by one space to the left vertical border of the text box.

Spin button

A spin button is a list of choices that are on a "round wheel" like a fruit machine wheel. The user can then "spin" the wheel to the desired option using the up and down direction buttons.

A spin button consists of a label, text box and up and down arrows to the right of the box. The label and text box follow the same guidelines as if it was a text box and label except that there is always a value within the text box when it is first displayed. The user can input a value from the keyboard or by clicking on the up and down arrows.

Borders

Field Borders

When grouping related screen elements using a 3-D frame from Visual Basic for example, one line should be left between the top and bottom borders and the screen elements.

A minimum of two spaces should be between the left and right borders and the longest screen element including the caption (if it is the longest element).

The caption should be indented by one space from the left border, mixed case and in the top border.

If there is a number of grouped borders, the screen should be balanced by keeping the border widths and the border heights the same (as far as is reasonably possible).

Section Borders

Section borders should have one line between the top and bottom element and the border.

A minimum of four spaces should be left between the left and right element and the vertical borders.

The section border header should be upper case and indented by two spaces from the left border and in the top border.

Maintain balance as much as possible if there are multiple groupings.

Icons

Icons are an example of direct manipulation. They are used to depict objects in everyday life with which the user is familiar. There are three kinds of icons [80]:

1. Icon: this is a true icon i.e. it look likes what it means e.g. calendar icon;
2. Index: this icon may represent different things to different users depending on their past experience;
3. Symbol: this icon meaning must be learnt by the user.

Each icon has three characteristics Marcus, 1984 cited by [80]:

1) syntactics: this is an icon's appearance (colour, size, shape etc.) which could be used to relate icons with similar meanings;

2) semantics: this is an icon's meaning;

3) pragmatics: this is an icon's quality (i.e. is the icon sufficiently well drawn on the screen for a user to see the illustration clearly).

These characteristics determine the icon's effectiveness and usability.

Icons should be designed so that they are familiar to the user. If standards have been created for symbols such as the International standards Organisation (ISO) these should be adhered to.

Labels should be located beneath the icon. This will aid in the comprehension and learning of the icons meaning.

The maximum number of symbols should be 20. These should be arranged in a logical way. However, the choice should exist that if the user wants to arrange them to their own configuration they can do so.

Appendix III

General Computing Questionnaire

(Circle your response where appropriate)

Q1 Do you suffer from computer anxiety? Y / N

33 % Yes

67% No

Q2 Have you ever used a personal computer before? Y / N

(Excluding playing games)

79% Yes

21% No

If YES go to Q4

Q3 Before starting university have you ever heard of a

i) Visual Display Unit (VDU)	Y / N
ii) floppy disk drive	Y / N
iii) hard disk drive	Y / N
iv) CD-ROM	Y / N
v) keyboard	Y / N
vi) mouse?	Y / N

i) 17% Yes 21% No

ii) 33% Yes 4% No

iii) 33% Yes 4% No

iv) 21% Yes 21% No

v) 42% Yes 0% No

vi) 38% Yes 0% No

58% no response

4% didn't understand question

Go to Q6

Q4 Where have you used a personal computer?

a) home	**42%**
b) school	**50%**
c) university	**33%**
d) work	**25%**

Q5 Do you own a personal computer? (Not game consoles) Y / N

33% Yes

50% No

17% No response

Q6 Have you ever heard of Windows? Y / N

92% Yes

8% No

If NO finish

Q7 From what source have you heard about it?

a) friends 46%
b) magazine 21%
c) school 46%
d) university 38%

8% no response

Q8 Have you ever used Windows? Y / N

83% Yes

8% No

8% No response

Q9 Where have you used Windows?

a) home 29%
b) school 46%
c) university 50%
d) work 9%

17% No response

Q10 Did/do you find Windows easy to use? Y / N

63% Yes

13% No

25% No response

Q11 Excluding textual or numeric input do you prefer to use

a) keyboard 25%

b) mouse 75%

when using Windows?

Q12 When typing using a keyboard do you find it easy to use?

1	2	3	4	5
17%	**25%**	**29%**	**21%**	**0%**

(1 = very easy 5 = very difficult)

8% No response

Appendix IV

Chemical Information and Retrieval Assessment Questionnaire

Q1 Were you more anxious about:

a) using a computer **25%**

b) doing the test? **63%**

8% No response

4% didn' t understand question

Q2 In general did you find the package easy to use?

1	2	3	4	5
8%	**29%**	**42%**	**17%**	**4%**

(1 = very easy 5 = very difficult)

Q3 Did you have to concentrate more on how to use the package as opposed to the answering of the questions?

Yes	No
50%	**50%**

Q4 Did you use the tutorial?

Yes	No
42%	**54%**

4% No response

If NO go to Q6

Q5 Did you find the tutorial adequately explained how to use the package?

1	2	3	4	5

(1 = very adequately 5 = very poorly)

0% 13% 29% 0% 0%

42% No response

Q6 Would you prefer a lecturer explain how to use the package before the test?

Yes No

50% 50%

Q7 Would you want separate sessions for familiarisation with the package and doing the test?

Yes No

54% 46%

Q8 Could you find help easily within the program on your queries?

Yes No

67% 29%

4% No response

Q9 List any queries that you could not find help on within the program's help section.

How to start questions and stop going round and round in circles

92% No response

Q10 Did you think the photographs were acceptable in terms of quality?

Yes	No
38%	58%

4% No response

Q11 Do you think the photographs were useful for illustrating what popular chemical literature look like when you go to the library?

Yes	No
75%	21%

4% didn't understand question

Q12 Do you like to know if you get an answer correct or incorrect instantly?

Yes	No
96%	4%

Q13 Do you like to have detailed feedback explaining why your answer is wrong?

Yes	No
96%	4%

Q14 Did you find the feedback was adequate and useful in this test?

Yes	No
92%	4%

4% No response

Q15 Was there enough information available to help you select your answer?

Yes	No
38%	58%

4% No response

Q16 Did you use a sheet of paper for taking rough notes?

Yes	No
4%	96%

Q17 Did you print out the summary notes at the end of the test?

Yes	No
46%	**42%**

13% No response

Q18 Do you prefer

a) this method of testing or

b) the traditional method in which you write you answers down and hand them in at the end of the test and wait a few weeks for the results?

a) 83%

b) 8%

8% No response

Q19 Is there anything you do not like about the package?

No just not confident with it

difficulty in printing

too slow

Q20 Is there anything you do not like about this method of testing?

Obtained a low percentage - not happy

No apart from it is still a test

Fairness, unfamiliarity with type of package

Requires instant think - can not go back to the previous question

Can not explain your answers

Appendix V

Sample Code to Draw an Arrow Through 360^{O}

The following code is provided to illustrate how it is possible to enable an arrow representing electron flow can be drawn through 360^{O} if the arrow if anchored in one position i.e. the mouse left-hand button is pressed and held down an the mouse is then moved in any direction. The basic principle can be illustrated using the following diagram:

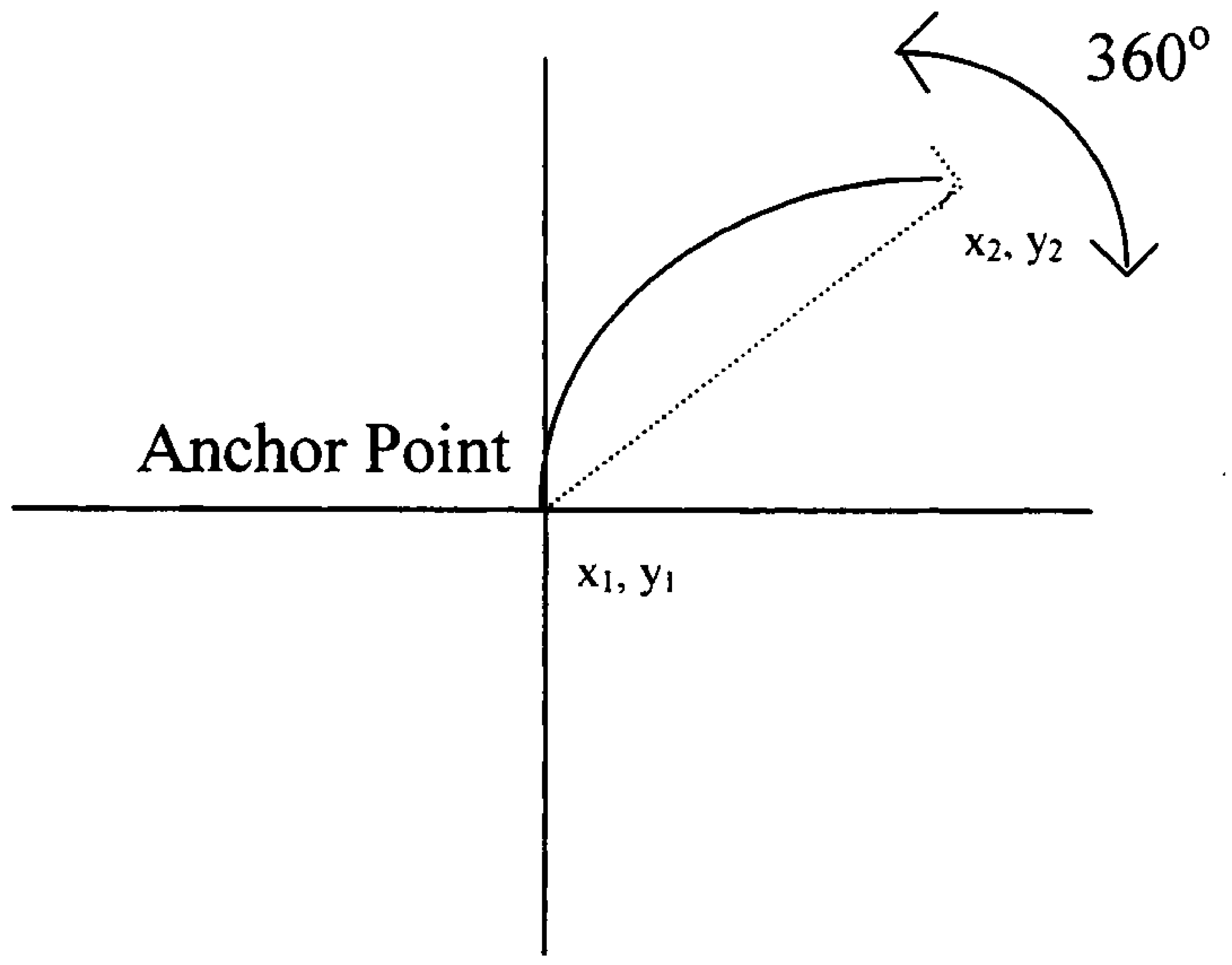


Figure 1 Illustration of the Drawing of an Arrow

In effect when drawing a curly arrow using the mouse a straight line is being drawn between the co-ordinates (x_1, y_1) and (x_2,y_2). Within VB the circle function enables an arc to be drawn. By passing the (x_1, y_1) and (x_2,y_2) values and the *aspect ratio* it is possible to draw an elliptical arc. The *aspect ratio* determines if a circle is non-elliptical or elliptical. The circle function works in radians. One the arc is drawn the arrowhead itself has to be drawn. This arrowhead consists of two lines which, as the arc moves through 360° must be also recalculated.

```
Private Sub arrow()

    If x1 = x2 And y1 < y2 Then

        Circle (x1 + ((x2 - x1) / 2), y1 + ((y2 - y1) / 2)), (y2 - y1) / 2, , 1.57,
        4.71

        Line (x2, y2)-Step(-0.3 * ((y2 - y1) / 2), 0.3 * ((y2 - y1) / 2))

        Line (x2, y2)-Step(-0.3 * ((y2 - y1) / 2), -0.3 * ((y2 - y1) / 2))

    End If
```

```
If x1 = x2 And y1 > y2 Then
        Circle (x1 + ((x2 - x1) / 2), y2 + ((y1 - y2) / 2)), (y1 - y2) / 2, , 4.17,
        1.57
        Line (x2, y2)-Step(0.3 * ((y1 - y2) / 2), -0.3 * ((y1 - y2) / 2))
        Line (x2, y2)-Step(0.3 * ((y1 - y2) / 2), 0.3 * ((y1 - y2) / 2))
End If
If x1 < x2 And y1 = y2 Then
        Circle (x1 + ((x2 - x1) / 2), y1 + ((y2 - y1) / 2)), (x2 - x1) / 2, , 3.14, 0
        Line (x2, y2)-Step(-0.3 * ((x2 - x1) / 2), 0.3 * ((x2 - x1) / 2))
        Line (x2, y2)-Step(0.3 * ((x2 - x1) / 2), 0.3 * ((x2 - x1) / 2))
End If
If x1 > x2 And y1 = y2 Then
        Circle (x2 + ((x1 - x2) / 2), y1 + ((y2 - y1) / 2)), (x1 - x2) / 2, , 0, 3.14
        Line (x2, y2)-Step(-0.3 * ((x1 - x2) / 2), -0.3 * ((x1 - x2) / 2))
        Line (x2, y2)-Step(0.3 * ((x1 - x2) / 2), -0.3 * ((x1 - x2) / 2))
End If
'north east
If x2 > x1 And y1 > y2 Then
        Circle (x1 + ((x2 - x1) / 2), y2 + ((y1 - y2) / 2)), (Sqr((x2 - x1) ^ 2 +
        (y1 - y2) ^ 2)) / 2, , (Atn((y1 - y2) / (x2 - x1))), 3.14 + (Atn((y1 - y2) /
        (x2 - x1)))
        'left line
        If (Atn((y2 - y1) / (x2 - x1))) < -0.785 And (Atn((y2 - y1) / (x2 - x1))) >
        -1.57 Then
```

```
        Line (x2, y2)-Step((-0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
        2)) * (3.14 - (Atn((x2 - x1) / (y2 - y1)))), (-0.05 * ((Sqr((x1 -
        x2) ^ 2 + (y1 - y2) ^ 2)) / 2)) * (6.28 - (Atn((y2 - y1) / (x2 -
        x1)))))

End If

If (Atn((y2 - y1) / (x2 - x1))) > -0.785 And (Atn((y2 - y1) / (x2 - x1))) <
0 Then

        Line (x2, y2)-Step((-0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
        2)) * ((Atn((x2 - x1) / (y2 - y1)))), (-0.05 * ((Sqr((x1 - x2) ^ 2 +
        (y1 - y2) ^ 2)) / 2)) * (6.28 - (Atn((y2 - y1) / (x2 - x1)))))

End If

If (Atn((y2 - y1) / (x2 - x1))) < -0.785 And (Atn((y2 - y1) / (x2 - x1))) >
-1.57 Then

        Line (x2, y2)-Step((-0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
        2)) * (4.71 - (Atn((x2 - x1) / (y2 - y1)))), (0.05 * ((Sqr((x1 - x2)
        ^ 2 + (y1 - y2) ^ 2)) / 2)) * (3.14 - (Atn((y2 - y1) / (x2 - x1)))))

End If

If (Atn((y2 - y1) / (x2 - x1))) > -0.785 And (Atn((y2 - y1) / (x2 - x1))) <
0 Then

        Line (x2, y2)-Step((-0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
        2)) * (4.71 - (Atn((x2 - x1) / (y2 - y1)))), (-0.05 * ((Sqr((x1 -
        x2) ^ 2 + (y1 - y2) ^ 2)) / 2)) * (1.57 + (Atn((y2 - y1) / (x2 -
        x1)))))

End If

'right line

If (Atn((y2 - y1) / (x2 - x1))) = -0.785 Then

        Line (x2, y2)-Step(-0.3 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
        2), 0)
```

```
            Line (x2, y2)-Step(0, -0.3 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2))
            / 2))
        End If
    End If
    'south east
    If x2 > x1 And y2 > y1 Then
        Circle (x1 + ((x2 - x1) / 2), y1 + ((y2 - y1) / 2)), (Sqr((x2 - x1) ^ 2 +
        (y2 - y1) ^ 2)) / 2, , 6.28 - (Atn((y2 - y1) / (x2 - x1))), 1.57 + (Atn((x2 -
        x1) / (y2 - y1)))
        If (Atn((y2 - y1) / (x2 - x1))) > 0.785 And (Atn((y2 - y1) / (x2 - x1))) <
        1.57 Then
            Line (x2, y2)-Step((0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
            2)) * (3.14 - (Atn((x2 - x1) / (y2 - y1)))), (-0.05 * ((Sqr((x1 -
            x2) ^ 2 + (y1 - y2) ^ 2)) / 2)) * (6.28 - (Atn((y2 - y1) / (x2 -
            x1)))))
        End If
        If (Atn((y2 - y1) / (x2 - x1))) < 0.785 And (Atn((y2 - y1) / (x2 - x1))) >
        0 Then
            Line (x2, y2)-Step((-0.04 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
            2)) * (3.14 - (Atn((x2 - x1) / (y2 - y1)))), (-0.04 * ((Sqr((x1 -
            x2) ^ 2 + (y1 - y2) ^ 2)) / 2)) * (6.28 - (Atn((y2 - y1) / (x2 -
            x1)))))
        End If


        If (Atn((y2 - y1) / (x2 - x1))) > 0.785 And (Atn((y2 - y1) / (x2 - x1))) <
        1.57 Then
```

```
            Line (x2, y2)-Step((0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
            2)) * (4.71 - (Atn((x2 - x1) / (y2 - y1)))), (0.05 * ((Sqr((x1 - x2)
            ^ 2 + (y1 - y2) ^ 2)) / 2)) * (3.14 - (Atn((y2 - y1) / (x2 - x1)))))
        End If
        If (Atn((y2 - y1) / (x2 - x1))) < 0.785 And (Atn((y2 - y1) / (x2 - x1))) >
        0 Then
            Line (x2, y2)-Step((0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
            2)) * (4.71 - (Atn((x2 - x1) / (y2 - y1)))), (-0.05 * ((Sqr((x1 -
            x2) ^ 2 + (y1 - y2) ^ 2)) / 2)) * (1.57 + (Atn((y2 - y1) / (x2 -
            x1)))))
        End If
    End If
    'north west
    If x1 > x2 And y1 > y2 Then
        Circle (x2 + ((x1 - x2) / 2), y2 + ((y1 - y2) / 2)), (Sqr((x1 - x2) ^ 2 +
        (y1 - y2) ^ 2)) / 2, , 6.28 - (Atn((y2 - y1) / (x2 - x1))), 3.14 - (Atn((y2 -
        y1) / (x2 - x1)))
        If (Atn((y2 - y1) / (x2 - x1))) > 0.785 And (Atn((y2 - y1) / (x2 - x1))) <
        1.57 Then
            Line (x2, y2)-Step((0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
            2)) * (3.14 - (Atn((x2 - x1) / (y2 - y1)))), (-0.05 * ((Sqr((x1 -
            x2) ^ 2 + (y1 - y2) ^ 2)) / 2)) * (6.28 - (Atn((y2 - y1) / (x2 -
            x1)))))
        End If
        If (Atn((y2 - y1) / (x2 - x1))) < 0.785 And (Atn((y2 - y1) / (x2 - x1))) >
        0 Then
            Line (x2, y2)-Step((-0.04 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
            2)) * (3.14 - (Atn((x2 - x1) / (y2 - y1)))), (-0.04 * ((Sqr((x1 -
```

```
x2) ^ 2 + (y1 - y2) ^ 2)) / 2)) * (6.28 - (Atn((y2 - y1) / (x2 -
x1)))))
End If
If (Atn((y2 - y1) / (x2 - x1))) > 0.785 And (Atn((y2 - y1) / (x2 - x1))) <
1.57 Then
Line (x2, y2)-Step((0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
2)) * (4.71 - (Atn((x2 - x1) / (y2 - y1)))), (0.05 * ((Sqr((x1 - x2)
^ 2 + (y1 - y2) ^ 2)) / 2)) * (3.14 - (Atn((y2 - y1) / (x2 - x1)))))
End If
If (Atn((y2 - y1) / (x2 - x1))) < 0.785 And (Atn((y2 - y1) / (x2 - x1))) >
0 Then
Line (x2, y2)-Step((0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
2)) * (4.71 - (Atn((x2 - x1) / (y2 - y1)))), (-0.05 * ((Sqr((x1 -
x2) ^ 2 + (y1 - y2) ^ 2)) / 2)) * (1.57 + (Atn((y2 - y1) / (x2 -
x1)))))
End If
End If
'south west
If x1 > x2 And y2 > y1 Then
Circle (x2 + ((x1 - x2) / 2), y1 + ((y2 - y1) / 2)), (Sqr((x1 - x2) ^ 2 +
(y2 - y1) ^ 2)) / 2, , (Atn((y2 - y1) / (x1 - x2))), 3.14 + (Atn((y2 - y1) /
(x1 - x2)))
If (Atn((y2 - y1) / (x2 - x1))) < -0.785 And (Atn((y2 - y1) / (x2 - x1))) >
-1.57 Then
Line (x2, y2)-Step((-0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
2)) * (3.14 - (Atn((x2 - x1) / (y2 - y1)))), (-0.05 * ((Sqr((x1 -
x2) ^ 2 + (y1 - y2) ^ 2)) / 2)) * (6.28 - (Atn((y2 - y1) / (x2 -
x1)))))
```

```
        End If
        If (Atn((y2 - y1) / (x2 - x1))) > -0.785 And (Atn((y2 - y1) / (x2 - x1))) <
        0 Then
            Line (x2, y2)-Step((-0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
            2)) * ((Atn((x2 - x1) / (y2 - y1)))), (-0.05 * ((Sqr((x1 - x2) ^ 2 +
            (y1 - y2) ^ 2)) / 2)) * (6.28 - (Atn((y2 - y1) / (x2 - x1)))))
        End If
        If (Atn((y2 - y1) / (x2 - x1))) < -0.785 And (Atn((y2 - y1) / (x2 - x1))) >
        -1.57 Then
            Line (x2, y2)-Step((-0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
            2)) * (4.71 - (Atn((x2 - x1) / (y2 - y1)))), (0.05 * ((Sqr((x1 - x2)
            ^ 2 + (y1 - y2) ^ 2)) / 2)) * (3.14 - (Atn((y2 - y1) / (x2 - x1)))))
        End If
        If (Atn((y2 - y1) / (x2 - x1))) > -0.785 And (Atn((y2 - y1) / (x2 - x1))) <
        0 Then
            Line (x2, y2)-Step((-0.05 * ((Sqr((x1 - x2) ^ 2 + (y1 - y2) ^ 2)) /
            2)) * (4.71 - (Atn((x2 - x1) / (y2 - y1)))), (-0.05 * ((Sqr((x1 -
            x2) ^ 2 + (y1 - y2) ^ 2)) / 2)) * (1.57 + (Atn((y2 - y1) / (x2 -
            x1)))))
        End If
    End If
End Sub
```

References

References

[1] Brown G.A. and Pendelbury, M., *Assessing Active Learning*, USDTU Effective Learning in Teaching in Higher Education Module 11, 1992

[2] Gibbs G., Habeshaw S., Habeshaw T., *53 Interesting Ways to Assess Your Students*, 1988

[3] Bull J., *Using Technology To Assess Student Learning*, 1993

[4] *Editorial*, CTI Active Learning Number Dec. 1994, 1, 2

[5] Bunderson, C., Inouye, D., & Olsen, J. *The Four Generations of Computerized Education Measurement*, (1989 In Robert Linn (Ed.) Educational measurement (3rd edition), 367-407. NY: American Council of Education / MacMillan

[6] Dempster, J.A., *Review Using Technology to Assess Student Learning*, CTI Active Learning Number Dec. 1994, 1, 54

[7] Stephens, D., *Using computer assisted assessment: time saver or sophisticated distraction?,* CTI Active Learning Dec 1994, 1, 11-15

[8] Griffiths, F., *Freedom and flexibility in the higher education examination process*, CTI Active Learning Dec 1994, 1, 22

[9] Bull, J., *Computer based assessment: some issues for consideration*, CTI Active Learning Dec 1994, 1, 18-21

[10] Sandals, L.H., *An overview of the uses of computer based assessment and diagnosis*, Canadian Journal of Educational Communication 1992, 21, (1), 67-78

[11] De Cicco, E., *Assessment - Ways Forward*, National Council for Educational Technology, ALT-N 8, 3

[12] Eaves, C., R., *Educational Assessment in the United States*, 1984-85, Diagnostique 10, 5-39

[13] Peel, A., *Computer aided assessment through hypermedia*, CTI Active Learning Dec 1994, 1, 35-37

[14] Doughty, G., Magill, J., & Turner, I., *Interactive multimedia computer based assessment* CTI Active Learning Dec 1994 1, 38

[15] Dempster, J.A., *Question Mark Designer for Windows*, CTI Active Learning Dec 1994, 1, 47-50

[16] Gagan, J.M.F., *Computers in Assessment,* Research in Assessment V Computers in Assessment RSC 1987, 33-49

[17] Williams C. S., *Multiple-choice tests and the computer*, J. Furth. High. Educ., Autumn 1984, 8, (3), 52-73

[18] Harris, T. C.; Edouard L. J., *Multiple-choice questions: fact, fallacy and fancy,* Int. J. Math. Educ. Sci. Technol., Nov-Dec 1983, 14, (6), 685-691

[19] Turnbull, G., *Multiple-choices: testing in the United States*, Education, Jul 1992, 180, 2, (10), 33

[20] Wise, S.L., Plake, B. S., *Computer-Based Testing in Higher Education,* Measurement and Evaluation in Counselling and Development, Apr 1990, 23, (1), 3-10

[21] Binder, B., Maffat, T. M., *Microcomputer generated interactive repeatable general chemistry examinations (CS),* J. Chem. Educ. 1988, 65, 342

[22] Pickering Sir G., *Against multiple-choice questions,* Med. Teach., Mar-Apr 1979 1, (2), 84-86

[23] Bauer H., *Sore finger items in multiple-choice tests,* System, 1991, 19, (4), 453-458

[24] Roitberg, J., Cherkas B. M., *Humanizing the multiple-choice test with partial credit*, International Journal of Mathematical Education in Science and Technology, Nov-Dec 1993, 24, (6), 799-812

[25] Atkinson, G. F.; Doadt, E. G., *New computer-graded examination allowing partial credit,* J. Chem. Educ. 1982, 59, 213

[26] Friel S., *The assessment of partial knowledge in multiple-choice chemistry examinations*, Thesis (Ph.D.) - Glasgow, 1981

[27] Tamir P., *Positive and negative multiple-choice items: how different are they?* Studies in Educational Evaluation, 1993, 19, (3), 311-325

[28] Tamir P., *Justifying the selection of answers in multiple-choice items,* International Journal of Science Education, Oct-Dec 90, 12, (5), 563-573

[29] Dewey, M. E., Slade P. D., *Role of grammatical clues in multiple-choice questions: an empirical study*, Med. Teach., Oct 1983, 5, (4), 146-148

[30] Cassels, J. R. T.; Johnstone, A. H., *The effect of language on student performance on multiple-choice tests in chemistry,* J. Chem. Educ. 1984, 61, 613

[31] Hodson D., *How important is question sequence?* Educ. Chem., 1987, 24 (1)

[32] Hodson D., *Multiple-choice tests - are four options as good as five?* Educ. Chem., May 1986, 23, (3)

[33] Johnstone A.H., Friel S., *Does the position matter?* Educ. Chem., Nov 1979, 16, (6)

[34] Kolstad, R. A., Kolstad R. K., *Multiple-choice achievement tests are improved with the option "none of these"*, Scientia Paedagogica Experimentalis, 1991, 28, (1), 69-78

[35] Pollard G.H., *Scoring to remove guessing in multiple-choice examinations,*

International Journal of Mathematical Education in Science and Technology, 1989, 20, 429-432

[36] Taylor J. F., *An innovative approach to data capture in automated assessment*

Program. Learn. Educ. Technol., 1992, 19, (4), 300-307

[37] Doadt, E., Atkinson G.F., *Matrexam : a new form of computer-graded examination,* Assess. High. Educ., 1980, 5, (2), 112-121

[38] Joorabchi B., *Construct problem-solving MCQs*, Med. Teach., 1981, 3, (1), p9-13

[39] Yoong, P., Jeffcoat C., *CPS : a self-assessment scheme*, Assess. High. Educ.; 1979, 4, (3), 180-186

[40] Dodd, B. G., And Others, *Computerized Adaptive Testing Using the Partial Credit Model: Effects of Item Pool Characteristics and Different Stopping Rules,* Educational and Psychological Measurement, 1993, 53, (1), 61-77

[41] Wise, S. L., And Others, *A Comparison of Self-Adapted and Computerized Adaptive Tests,* Journal of Educational Measurement, 1992, 29, (4), 329-39

[42] De Ayala, R. J., *A Comparison of the Nominal Response Model and the Three-Parameter Logistic Model in Computerized Adaptive Testing,* Educational and Psychological Measurement, 1989, 49, (4), 789-805

[43] Binder, B., Maffat, T. M., *Microcomputer generated interactive repeatable general chemistry examinations (CS),* J. Chem. Educ. 1988, 65, 342

[44] Vogel, G., *Microcomputer-based system for filing test questions and assembling examinations (CS),* J. Chem. Educ. 1985, 62, 1024

[45] Hartley, J.R., *Multimedia views of science education,* Studies in Science Education, 1994, 23, 75-87

[46] Dwyer, C., *Multimedia in education*, Educational Media International, 1993, 30, (4), 193-198

[47] Pedersen, H.-J., *Multimedia-interactive media*, Educational Media International, 1993, 30, (4), 219-220

[48] Wragg, E.C., *Multi-media in education: bane or boon?* Journal of Educational Television, 1993, 19, (2), 73-79

[49] Sealey, M., *Not like a book*, Educational Computing and Technology; 1994, 15, (2), 9-10,12

[50] Rocklin, T., O'Donnell, A. M., *Self-Adapted Testing: A Performance-Improving Variant of Computerized Adaptive Testing,* Journal of Educational Psychology, 1987, 79, (3), 315-19

[51] Perzylo, L., *The application of multimedia CD-ROMS in schools*, British Journal of Educational Technology, 1993, 24, (3), 191-197

[52] Atkins, M.J., *Theories of learning and multimedia applications: an overview*, Research Papers in Education, 1993, 8, (2), 251-271

[53] Rae, A., *Self-paced learning with video for undergraduates: a multimedia Keller Plan*, British Journal of Educational Technology, 1993, 24, (1), 43-51

[54] Hiroo, S., *Are we ready enough to learn from interactive multimedia*, Educational Media International, 1992, 29, (3), 181-188

[55] Barnard, J., *Multimedia and the future of distance learning technology*, Educational Media International, 1992, 29, (3), 139-144

[56] Koumi, J., *Getting it together: the co-production, involving planning, co-ordinating and control, of a multimedia course,* Educational Media International, 1992, 29, (2), 99-106

[57] Shir, N., Movshovitz-Hadar, N., Mevarech Z. R., *Is more always better?: the separate and combined effects of a computer and video programme on mathematics learning,* British Journal of Educational Psychology, 1992, 62, (1), 106-116

[58] International Council for Educational Media International Conference, *Multimediality 1989*, Educational Media International, 1990, 27, (1), 7-47

[59] Tamir, P., *Some issues related to the use of justifications to multiple-choice answers*, Journal of Biological Education, 1989, 23, (4), 285-292

[60] Galliani, L., *A pedagogic model of multi-mediality*, Educational Media International, 1989, 26, (3), 133-137

[61] Giacomantonio, M., *Interactive and multi-media systems for individual open learning,* Educational Media International, 1989, 26, (3), 138-143

[62] Paton, G., *Aspects of multi-media production in the United Kingdom,* Educational Media International, 1989, 26, (3), 157-160

[63] Pring, I., *Multi-media in the making*, Computer Education, 1989, 62, p19-21

[64] Megarry, J., *Hypertext and compact discs: the challenge of multi-media learning,* British Journal of Educational Technology, 1988, 19, (3), 172-183

[65] Barker P. G., *A practical introduction to authoring for computer-assisted instruction: part 8: multi-media CAL,* British Journal of Educational Technology, 1987, 18, (1), 25-40

[66] Barker, P. G., *Multimedia CAL,* Computer Education, 1986, 52, 20-23

[67] Peach, L., *The adult tutor in a multi-media society,* Adult Education, 1986, 58, (4), 336-338

[68] Whitnell, R. M et al, *Multimedia chemistry lectures,* J. Chem. Educ. 1994, 71, (9), 721-725

[69] Riley, F., *Understanding IT: A Review of Hypermedia Authoring Packages,* ITTI

[70] Heath, J., *When interactive multimedia is not truly interactive,* CTI Active Learning, Dec. 1995, 3, 7-9

[71] Jennings, R., Discover Windows™ 3.1 Multimedia, Que 1992

[72] Davies P., *Guidelines for Multimedia Courseware Developers in Higher Education,* Vol 1 Delivery, Production and Provision, University of Nottingham

[73] Williams, J., *The Use and Capture of Images For Computer-Based Learning,* ITTI

[74] Firebaugh, M.W., *Artificial Intelligence A Knowledge-Based Approach,* 1989

[75] Holme, B.A., Pierce, T.H., *Expert System Applications in Chemistry,* Symposium Series 1989

[76] Jurs, P.C., *Computer Software Applications in Chemistry,* 1987

[77] Cartwright, H.M., *Applications of Artificial Intelligence in Chemistry* 1993 Oxford Science Publications

[78] CTI No 10, Theme: Exploiting CD-ROM Technology

[79] Edwards J., Howe, G., Smith, F., Advice on Choosing an Authoring Package, ITTI

[80] *User-Interface Screen Design* Wilbert O. Galitz 1993

[81] Engineering the Human-Computer Interface

[82] ToolBook Ideas (Educational Edition V 1.53)

[83] Times Higher Education Supplement Feb 14th 1994 Multimedia Supplement ppxii Added Attraction, Brian Austin, University of Liverpool.

[84] CTI Centre for Chemistry, University of Liverpool

[85] TLTP University of Kent *Guidelines for the production of Hypertext Documents*

[86] Microsoft Windows Software Development Kit

[87] Visual Basic 3.0 Professional Visual Design Guide

[88] McMurry J., *Organic Chemistry*, Brooks / Cole

Bibliography

Bibliography

[1] Knight P., *Swifter Assessment*, Lancaster University 1994

[2] Gibbs G., *Assessing Student Centred Courses*, 1995

[3] Gibbs G., Ed., *Improving Student Learning Through Assessment and Evaluation*, 1994

[4] Research In Assessment XII *Science Ain't Easy*, RSC Assessment Subject Group, 1995

[5] Gibbs G., *Assessing More Students*, PCFC, 1993

[6] Brown S., Rust C., and Gibbs G., *Strategies for Diversifying Assessment in Higher Education*, Oxford Centre for Staff Development 1994

[7] Darby J., and Martin J., Ed., *Teaching With Multimedia*, CTI Active Learning 1995, 3

[8] Brailsford, T., *Examine: a self-assessment system for Windows and Macintosh*, CTI Active Learning 1994, 1, 16

[9] Greaves, R., Synopse report *Advanced solutions to courseware development,* CTI Active Learning 1994, 1, 65

[10] Pople, A., *Microsoft Visual Basic: a way to quick and easy courseware?* CTISS File 1992, 13, 13-15

[11] Miller, P., *Question Mark Professional and the multimedia extensions to it,* CTISS File 1992, 13, 30-33

[12] Ford, P., *Using Question Mark to evaluate your teaching,* CTISS File 1992, 13, 34

[13] Riley, F., *Visual Basic Professional 3.0,* ITTI, Understanding IT: A Review of Hypermedia Authoring Packages June 1994, pp 13-20,

[14] Bar, S., *Approaches to student assessment. Is computerised student assessment reliable?* CTI Active Learning 1994, 1, *69*

[15] Green, B. F., *Construct validity of computer based tests,* 1988, In H. Wainer & H.I. Braun (Eds) Test Validity Hillsdale N.J. Erlbaum

[16] Seddon, G.M., Hubbard J.I., *Changes in the marking standard and reliability of successive assessments of practical skills in science,* British Educational Research Journal, 1989, 15, (1), 53-60

[17] Seddon, G.M., *The validity of reliability measures,* British Educational Research Journal, 1988, 14, (1), 89-97

[18] Hammersley, M., *Some notes on the terms "validity" and "reliability"*, British Educational Research Journal, 1987, 13, (1), 73-81

[19] Jamieson, J., *Reliability of a computerized scoring routine for an open-ended task,*

[20] System, 1993, 21, (3), 305-322

[21] Townsend, M.A.R., *Headings within multiple-choice tests as facilitators of test performance*, British Journal of Educational Psychology; 1990, 60, (2), 153-160

[22] Hodson, D., *Some effects of changes in question structure and sequence on performance in a multiple-choice chemistry test*, Research in Science and Technological Education, 1984, 2, (2),177-185

[23] Akeroyd, F. M., *Progress in multiple-choice scoring methods, 1977-81*, J. Furth. High. Educ.; 1982, 6, (3), 86-90

[24] Abu-Sayf, F.K., *Recent developments in the scoring of multiple-choice items*, Educ. Rev., 1979, 31, (3), 269-279

[25] Elton, L.R.B., Chelu C.J., *An item bank for multiple-choice questions,* Phys. Educ., 1977, 12, 263-267

[26] Bridgeman, B., *A Comparison of Quantitative Questions in Open-Ended and Multiple-Choice Formats,* Journal of Educational Measurement, 1992, 29, (3), 253-71

[27] Burns, E., *Multiple-Choice Computer Diagnostic Test Designs,* Educational Technology, 1990, 30, (1), 49-53

[28] Choppin, B., *Principles of Item Banking,* Evaluation in Education: An International Review Series, 1985, 9, (1), 87-90

[29] Lloyd, A., Holden, C., *Computer aided assessments in pharmaceutical microbiology,* CTI Active Learning 1994, 1, 27-28

[30] *Computer assisted assessment in the Teaching ad Learning Technology Programme (TLTP)*, CTI Active Learning 1994, 1, 43-45

[31] Nduka E.C., *Students' degree project as an efficient test discriminator*, International Journal of Mathematical Education in Science and Technology, 1991, 22, (3), 395-401

[32] Treagust, D. F., *Development and the use of diagnostic tests to evaluate students' misconceptions in science*, International Journal of Science Education, 1988, 10, 2, 159-169

[33] Birenbaum M., *"How", beyond the "what", towards the "why": a rule-assessment approach to achievement testing*, Studies in Educational Evaluation, 1986, 12, 2, 159-168

[34] Cresswell, M.J., *Examination grades: how many should there be?* British Educational Research Journal, 1986, 12, 1, 37-54

[35] Eaves, R.C., Smith, E., *The Effect of Media and Amount of Microcomputer Experience on Examination Scores*, Journal of Experimental Education, 1986, 55, (1) 23-26

[36] Dempsey, J.V., Wager, S.U., *A Taxonomy for the Timing of Feedback in Computer-Based Instruction,* Educational Technology, 1988, 28, 10, 20-25

[37] Wulfson, S., Ed. *Software Reviews,* Science and Children, 1987, 25, 3, 50-51

[38] Birenbaum, M., Tatsuoka, K.K., *Effects of "On-Line" Test Feedback on the Seriousness of Subsequent Errors,* Journal of Educational Measurement, 1987, 24, 2, 145-55

[39] Brohn, D. M., *The Use of Computers in Assessment in Higher Education,* Assessment and Evaluation in Higher Education, 1986, 11, 3, 231-39

[40] Perry, D., *Classroom Testing on the Microcomputer,* Business Education Forum, 1987, 41, 4, 22-23

[41] Waugh, M.L., *Effects of Microcomputer-Administered Diagnostic Testing on Immediate and Continuing Science Achievement and Attitudes,* Journal of Research in Science Teaching, 1985, 22, 9, 793-805

[42] Mizokawa, D.T., Hamlin, M. D., *Guidelines for Computer-Managed Testing,* Educational Technology, 1984, 24, 12, 12-17

[43] Ebner, D.G., *Current Issues in Interactive Videodisc and Computer-Based Instruction,* Instructional Innovator, 1984, 29, 3, 24-29

[44] Brightman, H., J., *et al, Constructing and Using Computer-Based Formative Test,* Educational Technology, 1984, 24, 6, 36-38

[45] Wedman, J.F., Stefanich, G.P., *Guidelines for Computer-Based Testing of Student Learning of Concepts, Principles, and Procedure,* Educational Technology, 1984, 24, 6, 23-28

[46] Bejar, I.I., *Educational Diagnostic Assessment,* Journal of Educational Measurement, 1984, 21, 2, 175-89

[47] Johnson, S., Maher, B., *A Thesaurus-Linked Science Question-Banking System,* British Journal of Educational Technology, 1984, 15, 1, 14-23

[48] David, C.W., *Computer-assisted testing: In technical subjects (CS),* J. Chem. Educ. 1992, 69, 396

[49] Fischbach, F.A., Sell, N. J., *A structures approach to teaching applied problem solving through technology assessment,* J. Chem. Educ. 1986, 63, 522

[50] Myers, R.L., *Computerized grading of freshman chemistry laboratory experiments (CS),* J. Chem. Educ. 1986, 63, 507

[51] Russell, A.A., Staskun, M.G., Mitchell, B. L., *The use and evaluation of videodiscs in the chemistry laboratory,* J. Chem. Educ. 1985, 62, 420

[52] May, P.M., Murray, K., Williams, D. R., *CAPE: A computer program to assist with practical assessment (CS),* J. Chem. Educ. 1985, 62, 310

[53] Mihkelson, A., Willett, G.D., Williams, A.J., *The application of computer technology in the assessment of large undergraduate classes*, J. Chem. Educ. 1984, 61, 631

[54] Bath, D.A., Hughes, B.G., *Computerized quizzes with instant grading and response analysis (CS),* J. Chem. Educ. 1983, 60, 734

[55] Levkov, J.S., Thaker, U., *Computer-generated safety quiz (CS),* J. Chem. Educ. 1982, 59, 599

[56] Johnson, J.K., Van Willis, W., Seely, O., Moore, J.W., *Computer-assisted test construction in chemistry (CS),* J. Chem. Educ. 1981, 58, 177

[57] Bland, W., O'Donovan, J., *Is that a beaker?* Educ. Chem., 32, May 95, p 77-78

[58] Meester, M.A.M., Maskill, R., *The practical side of chemistry*, Educ. Chem., 30, 1993, p156-159

[59] Hills H., *Moving to multimedia*, IT Training, 1994, 6, 3, 17-23

[60] Vaughan, W., *Multimedia: how do they do that?* Training Technology and Human Resources, 1994, 7, 3, 3-4

[61] King, T., Barker P., *Evaluating interactive multimedia courseware - a methodology,* Computers and Education, 1993, 21, 4, 307-319

[62] Atkins, M., Yildiz R., *Evaluating multimedia applications,* Computers and Education, 1993, 21, 1-2, 133-139

[63] Hall, W., Layman J., *Applications of hypermedia in education,* Computers and Education, 1991, 16, 1, 113-119

[64] Brown D., *Multi-media?: why bother?: or let's extract the HYPE,* Information Technology and Learning, 1991, 13, 3, 110-112

[65] Ananda A.L., *Design of an intelligent on-line examination system,* Computers and Education, 1989, 13, 1, 45-52

[66] Singer R., *Optical mark reading technology,* Media in Education and Development, 1988, 21, 4, 153-156

[67] Manji, K., Barker P., *Device interfacing for multi-media CAL: part 2,* Computer Education; No.58: Feb 88 p7-9