

UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE QUÍMICA

ESTE EXEMPLAR CORRESPONDE A REDAÇÃO FINAL DA TESE,
DEFENDIDA POR PAULO SÉRGIO DE SOUZA, E APROVADA
PELA COMISSÃO JULGADORA.

Campinas, 18 de Abril de 1994



Dr. Celio Pasquini

DESENVOLVIMENTO DE UM PROGRAMA GERENCIADOR
DE BIBLIOTECA DE ESPECTROS E DE UM PROGRAMA
ESPECIALISTA PARA IDENTIFICAÇÃO DE GRUPOS
FUNCIONAIS DE COMPOSTOS ORGÂNICOS A
PARTIR DE ESPECTROS DE INFRAVERMELHO

TESE DE DOUTORADO
PAULO SÉRGIO DE SOUZA t 89

ORIENTADOR: PROF. DR. CELIO PASQUINI t

- 07/04/1994 -

*À minha querida esposa M. Madalena
e à minha querida filha Mariana, pelo
amor, carinho, compreensão e espera.*

*Aos meus pais José Ignácio e
Maria Alves (in memoriam), que
me deram a oportunidade de
existir, aprender e realizar.*

ÍNDICE

Página

Resumo	i
Abstract	iii
Introdução	1

Capítulo I

- Sistemas Especialistas em Química Analítica	6
I.1 - Sistemas Especialistas Aplicados em Cromatografia	8
I.2 - Métodos Automatizados para Elucidação de Estruturas	12
I.2.1 - Técnicas de Busca em Bibliotecas.....	12
I.2.2 - Métodos de Reconhecimento de Padrões	13
I.2.3 - Sistemas Especialistas	14
I.2.4 - Métodos Híbridos	15
I.3 - Programas para Elucidação de Estruturas Químicas	17

Capítulo II

- Construção de Sistemas Especialistas	25
II.1 - Linguagens Computacionais utilizadas em Programas Especialistas	29
II.1.1 - Sintaxe da Linguagem Turbo Prolog	30

Capítulo III

- Programa Gerenciador de Biblioteca de Espectros - PROGEIV	39
III.1 - Parte Experimental	40
III.2 - Descrição do programa PROGEIV	41
III.3 - Aquisição de Dados Espectrais	45
III.3.1 - Aquisição de Dados com a Mesa Digitalizadora	45
III.3.2 - Aquisição de Dados com um Espectrofotômetro de IV	51
III.4 - Manipulações dos Espectros da Biblioteca	58
III.4.1 - Sub-Rotinas para Manipulações de Espectros	60
A - ShowCurrentSpectrum	60
B - CompareCurrentSpectrum	61
C - ShowTwoSpectrum	61
D - HouseKeeping	62
III.5 - Resultados e Discussões	63

continua na próxima página Índice

Capítulo IV	
- Programa Especialista para Identificação de Grupos Funcionais de Compostos Orgânicos - EXPERTIV	70
IV.1 - Descrição do Programa EXPERTIV	72
IV.1.1 - Módulo de Leitura	76
IV.1.2 - Módulo de Interpretação	93
IV.2 - Resultados e Discussões	127
- Compostos carbonilados	132
- Compostos não carbonilados	148
IV.3 - Perspectivas	173
Conclusões	175
Referências Bibliográficas	177
Palavras Chaves com as suas respectivas referências	186

ÍNDICE DOS APÊNDICES

A - Mesa Digitalizadora - HM	187
A.1 - Parte Mecânica	187
A.2 - Parte Eletrônica	189
A.2.1 - Transdutor Analógico de Posição	189
A.2.2 - Porta do Usuário para Microcomputadores IBM-XT ou AT	191
A.2.3 - Circuito Decodificador de Endereços	191
A.2.4 - Conversor Dinâmico Analógico/Digital de dois Canais	194
A.3 - Resultados e Discussões	196
B - Base de Conhecimento " <i>dba_links</i> " do programa EXPERTIV	198
C - Relação dos compostos orgânicos interpretados pelo programa EXPERTIV através de seus espectros de I.V.	206

ÍNDICE DE FIGURAS

	Página
II.1 - Função do engenheiro de conhecimento	27
II.2 - Componentes de um Sistema Especialista ou sistema baseado em conhecimento	28
III-1 - Diagrama de blocos do programa PROGEIV	43
III-2 - Saída de vídeo para ajudar o usuário a selecionar o tipo de monitor adequado	44
III-3 - Menu principal do programa PROGEIV	44
III-4 - Diálogo originado da opção número 1 do menu principal (Atributos da Escala para o Espectro de I.V.)	46
III-5 - Parte do diálogo originado pela opção número 2 do menu principal (Inicia a Digitalização)	47
III-6 - Diagrama de blocos das sub-rotinas encarregadas para aquisição de dados espectrais	49
III-7 - Parte do arquivo "spe001.spt" referente ao espectro de I.V. do ácido 3-bromopropiônico	50
III-8 - Fluxograma da sub-rotina "PERKIN"	53
III-9 - Espectro de I.V. do hexanaldeído obtido da biblioteca do espectrofotômetro Perkin Elmer, modelo 1600	57
III-10 - Diagrama de blocos das sub-rotinas de manipulações dos espectros da biblioteca	59
III-11 - Espectros de I.V. do ácido dibromoacético	64
III-12 - Tela do monitor mostrando o espectro de I.V. corrente e um outro espectro da biblioteca selecionado pelo usuário, para comparação	67
III-13 - Tela do monitor mostrando dois espectros de I.V. presentes na biblioteca, selecionados pelo usuário, para comparação	68
III-14 - Partes dos arquivos "spec.spt" e "spec.ltd" do álcool 2-pentílico	69
IV-1 - Componentes do programa EXPERTIV	75
IV-2 - Desenvolvimento do módulo de leitura READMOD para obtenção dos dados espectrais	77
IV-3 - Espectro de I.V. original da biblioteca SADTLER do 2-butanona	78
IV-4 - Arquivo DOS "spec.ltd" em formato de lista Prolog referente ao espectro de I.V. original da biblioteca SADTLER do 2-butanona	78
IV.5 - Espectro de I.V. do 2-butanona com os fatos que são armazenados no arquivo DOS "points.lec"	83
IV-6 - Formatos das bandas de absorção mais comumente encontradas num espectro de I.V.	84

continua na próxima página

	Página
IV-7 - Arquivo DOS "bandlist.lec" para o espectro de I.V. original do 2-butanona, mostrando separadamente as bandas de absorção extraídas	85
IV-8 - Arquivo DOS "realband.lec" para o espectro de I.V. original do 2-butanona	86
IV-9 - Bandas de absorção de uma vibração de estiramento O-H e de uma carbonila mostrando a largura a $\frac{3}{4}$ e a $\frac{1}{4}$ da altura	88
IV-10 - Bandas de estiramento O-H de um ácido carboxílico e de um álcool com os seus arquivos DOS "highreg.lnk"	90
IV-11 - Fatos da base de dados interna "dba_bands", referente a leitura do espectro de I.V. do 2-butanona	92
IV-12 - Componentes do módulo interpretador INTERMOD	94
IV-13 - Parte do arquivo DOS "dataiv.dba" onde estão inseridos os fatos referentes a base de conhecimento "dba_links", apenas para o grupo funcional ácido carboxílico	97
IV-14 - Parte do arquivo DOS "scanned1.lnk" referente a interpretação do espectro de I.V. do ácido 3-bromopropiônico	101
IV-15 - Espectro do ácido 3-bromopropiônico e parte do arquivo DOS "scanned2.lnk" da base de dados interna "segunda_varredura"	104
IV-16 - Comportamento das posições da banda de carbonila extraídas de 52 espectros de I.V. de ácidos carboxílicos	108
IV-17 - Comportamento das intensidades da banda de carbonila extraídas de 52 espectros de I.V. de ácidos carboxílicos	110
IV-18 - Comportamento das larguras das bandas de carbonila extraídas de 52 espectros de I.V. de ácidos carboxílicos	110
IV-19 - Parte do arquivo DOS "scanned3.lnk" relacionando os fatos inseridos na base de dados interna "terceira_varredura", referente a interpretação do espectro do ácido 3-bromopropiônico	112
IV-20 - Fatos espectroscópicos extraídos de maneira simbólica pelo INTERMOD, até a terceira varredura feita em um espectro de I.V.	113
IV-21 - Arquivo DOS "probable.lnk" relacionando os fatos inseridos na base de dados interna "povavel_grupo_funcional", referente a interpretação do espectro de I.V. do ácido 3-bromopropiônico	115
IV-22 - Fatos da base de dados interna "tipo_especifico", salvos no arquivo DOS "specific.lnk", referente a interpretação do espectro de I.V. do anidrido de ácido crotônico	123
IV-23 - Parte do arquivo DOS "dataiv.dba", onde estão armazenados os fatos da base de conhecimento "dba_links" para o grupo nitro	125

IV-24 - Desempenho do programa EXPERTIV na interpretação de cento e sessenta espectros de I.V. de compostos carbonilados, para identificar os grupos funcionais mais prováveis	132
IV-25 - Desempenho do programa EXPERTIV nas interpretações de dezoito espectros de I.V. de anidridos de ácidos	135
IV-26 - Desempenho do programa EXPERTIV nas interpretações de vinte e oito espectros de I.V. de ácidos carboxílicos	138
IV-27 - Desempenho do programa EXPERTIV nas interpretações de quarenta e quatro espectros de I.V. de ésteres	141
IV-28 - Desempenho do programa EXPERTIV nas interpretações de vinte e quatro espectros de I.V. de compostos cetônicos	143
IV-29 - Desempenho do programa EXPERTIV nas interpretações de vinte e oito espectros de I.V. de aldeídos	145
IV-30 - Desempenho do programa EXPERTIV nas interpretações de dezoito espectros de I.V. de haletos de acila	147
IV-31 - Desempenho do programa EXPERTIV nas interpretações de cento e noventa espectros de I.V. de compostos não carbonilados	149
IV-32 - Desempenho do programa EXPERTIV nas interpretações de vinte e cinco espectros de I.V. de compostos alcoólicos	151
IV-33 - Desempenho do programa EXPERTIV nas interpretações de vinte e seis espectros de I.V. de compostos olefinicos	154
IV-34 - Desempenho do programa EXPERTIV nas interpretações de vinte e seis espectros de I.V. de compostos aromáticos	156
IV-35 - Fórmulas estruturais de alguns compostos orgânicos da tabela IV-18	163
IV-36 - Fórmulas estruturais de alguns compostos orgânicos da tabela IV-19	167
IV-37 - Fórmulas estruturais de alguns compostos polifuncionais da tabela IV-20, que foram identificados pelo programa EXPERTIV	170

Apêndices

A-1 - Vista em perspectiva da parte mecânica da mesa digitalizadora.....	188
A-2 - Detalhe do contato elétrico deslizante utilizado no digitalizador.....	189
A-3 - Circuito transdutor de posição horizontal e vertical do cursor da mesa digitalizadora	190
A-4 - Diagrama do circuito decodificador de endereços.....	192
A-5 - Circuito do conversor dinâmico analógico/digital com dois canais	195

ÍNDICE DE TABELAS

	Página
I.1 - Programas para elucidação automática de estruturas químicas	24
IV-1 - Relação dos valores adotados para as larguras e intensidades das bandas de absorção para os espectros de I.V.	96
IV-2 - Relação dos grupos funcionais com suas vibrações características	103
IV-3 - Resultados obtidos da interpretação do espectro de I.V. do composto ácido 2,2-dimetilpropionico, indicando como seria o retorno para o usuário, sem as atribuições dos valores dos "scores" no INTERMOD.....	128
IV-4 - Relação dos grupos funcionais analisados pelo programa EXPERTIV	130
IV-5 - Resultados obtidos das interpretações dos espectros de I.V. para cada grupo funcional analisado pelo programa EXPERTIV, indicando a sua porcentagem de identificação correta.....	131
IV-6 - Resultados obtidos do programa EXPERTIV nas interpretações de cento e sessenta espectros de I.V. de compostos carbonilados, relacionando os grupos funcionais mais prováveis	133
IV-7 - Resultados obtidos nas interpretações dos espectros de I.V. de dezoito anidridos de ácidos	135
IV-8 - Resultados obtidos nas interpretações dos espectros de I.V. de vinte e oito espectros de I.V. de ácidos carboxílicos	138
IV-9 - Resultados obtidos nas interpretações dos espectros de I.V. de quarenta e quatro espectros de I.V. de ésteres	140
IV-10 - Resultados obtidos nas interpretações dos espectros de I.V. de vinte e quatro espectros de I.V. de compostos cetônicos	142
IV-11 - Resultados obtidos nas interpretações dos espectros de I.V. de vinte e oito espectros de I.V. de aldeídos	145
IV-12 - Resultados obtidos nas interpretações dos espectros de I.V. de dezoito espectros de I.V. de haletos de acila	147
IV-13 - Resultados obtidos do programa EXPERTIV nas interpretações de cento e noventa espectros de I.V. de compostos não carbonilados, relacionando os grupos funcionais mais prováveis.....	148
IV-14 - Resultados obtidos nas interpretações dos espectros de I.V. de vinte e cinco compostos alcoólicos	151
IV-15 - Resultados obtidos nas interpretações dos espectros de I.V. de vinte e seis compostos olefinicos	154
IV-16 - Resultados obtidos nas interpretações dos espectros de I.V. de vinte e seis compostos aromáticos	156

continua na próxima página

IV-17 - Relação dos compostos acetilênicos não identificados pelo programa EXPERTIV, com o retorno dos grupos funcionais identificados como mais prováveis.....	157
IV-18 - Relação de grupos funcionais de compostos carbonilados, com os seus tipos específicos identificados pelo programa EXPERTIV.....	162
IV-19 - Resultados obtidos nas interpretações dos espectros de I.V., retornando os grupos funcionais, os tipos específicos e outros grupos identificados como mais prováveis	165
IV-20 - Resultados obtidos nas interpretações dos espectros de I.V. de alguns compostos polifuncionais, identificando os grupos funcionais, os seus tipos específicos e outros grupos como mais prováveis	169
IV-21 - Resultados obtidos nas interpretações de três espectros de I.V. com o programa de busca em biblioteca e o programa EXPERTIV	172

RESUMO

Este trabalho pode ser dividido em duas partes principais que se complementam com o objetivo comum de utilização de computadores de pequeno porte na obtenção, armazenamento, gerenciamento e interpretação de dados espectrais de infravermelho (I.V.).

Na sua primeira parte é descrito um programa elaborado em linguagem Quick-Basic, versão 4.5, capaz de adquirir dados de espectros de I.V., através de digitalização com instrumentos que empregam transdutores de posição do tipo denominado "mesa digitalizadora". Um destes instrumentos foi desenvolvido para possibilitar o andamento inicial deste trabalho e sua construção se encontra aqui descrita. Os dados de espectros de I.V. podem ser acessados, também, através de tradução de arquivos originados em discos flexíveis através do microcomputador do espectrofotômetro da linha Perkin Elmer.

Os dados espectrais originam uma biblioteca de espectros, de baixo custo e de utilidade específica do usuário, que pode, com o auxílio do programa, gerenciá-la de forma a visualizar graficamente, apagar e/ou incluir espectros e compará-los de forma fácil e efetiva com o uso dos recursos gráficos da linguagem empregada.

Os espectros presentes na biblioteca podem ser preparados para serem interpretados, através de sistemas especialistas desenvolvidos em linguagem Turbo Prolog. A interface que permite a interação entre os dois aplicativos, está também descrita.

Na segunda parte deste trabalho, é descrito um sistema especialista elaborado em linguagem Turbo Prolog, versão 2.0, executável em um microcomputador tipo IBM-PC/AT-386. O sistema foi desenvolvido com o objetivo de identificar, automaticamente, grupos funcionais de compostos orgânicos através de dados obtidos de espectros de infravermelho (I.V.). Uma vez que o espectro digitalizado é colocado na forma de uma lista (formato requerido pelo Prolog), o programa pode ser utilizado para interpretação do espectro através de regras baseadas em conhecimento de um especialista humano. Esta interpretação possibilita a identificação dos grupos funcionais presentes no composto que deu origem ao espectro. O programa especialista é constituído de dois módulos principais.

O primeiro, denominado módulo de leitura, atua sobre a lista Prolog, que contém o espectro digitalizado, produzindo como resultado a extração de informações relevantes para a interpretação tais como: posição, intensidade, largura e formato do pico. O segundo módulo é encarregado da interpretação propriamente dita, atuando sobre o conjunto de fatos amostrados na etapa anterior e utilizando inúmeras regras de inferência. Neste módulo, também, foram implementadas ferramentas de lógica difusa, com o objetivo de tornar o banco de conhecimento do programa especialista mais completo e de direcionar algumas conclusões de forma menos determinística. Uma avaliação do sistema especialista revela um índice médio de acerto maior do que 90%, na identificação de grupos funcionais, quando presentes nos compostos analisados.

ABSTRACT

"Development of an Spectra Library Manager Program and of an Expert System for Identification of Functionals Group of Organic Compounds using Infrared Spectra"

The present work is divided in two main parts which complement themselves in order to make use of low capacity computers for acquisition, storage, management and interpretation of spectral data in the infrared (I.R.) region.

The first part describes a program written in Quick-Basic version 4.5 aiming I.R. spectral data acquisition using digitising devices named "digitising tables". One of such devices has been developed in the early stages of this work and its construction is, therefore, herein described. Also, I.R. spectral data can be retrieved from standard files generated by Perkin Elmer spectrophotometers.

Spectral data that has been acquired originate low-cost, user tailored library of spectra. By making use of the program, the user can, manage the library showing graphically the spectra, remove/include spectra and compare those spectra in a user friendly environment.

Spectra in the library can be prepared to be interpreted through the use of Expert System wrote in Turbo Prolog. The interface allowing such interaction is also described.

The second part describes an Expert System (E.S.) developed in Turbo Prolog, version 2.0. The E.S. has been developed for automatic identification of organic functional groups, exploiting data obtained from infrared spectra (I.R.). The expert program is constituted of two main modules. The first, named reading module, works on Prolog lists containing the digitised spectrum, providing the relevant information for interpretation such as: position, intensity, band width and peak shape. The second, makes the interpretation, working on the set of facts gathered in first step and is constituted of many production rules. Fuzzy characteristics have been imposed to the system aiming to avoid extremely deterministic conclusions. The interpretation module of the expert system, achieves correct functional group identification in more than 90 % of the cases when they actually exist in the investigated organic compounds.

INTRODUÇÃO

Nas últimas décadas, a Química Analítica vem buscando caminhos alternativos para os métodos tradicionais de análises. Este direcionamento está relacionado principalmente com a introdução dos métodos instrumentais e dos microcomputadores. Desde então, iniciou-se uma mudança significativa no desenvolvimento da instrumentação moderna de laboratório, que não causou apenas uma simplificação da operação de alguns instrumentos comuns mas, também, forneceu meios para o desenvolvimento de instrumentos baseados em medidas mais complexas, como por exemplo, os espectrômetros de infravermelho com transformada de Fourier (FT/IR), de ressonância magnética nuclear (RMN) e de massas. O microcomputador acoplado a um instrumento de análise permite uma maior velocidade na aquisição de dados, obtenção de resultados mais confiáveis e maior capacidade de armazenamento e tratamento de dados, o que o torna útil em trabalhos de rotina mais intensos.

O impacto advindo do emprego de microcomputadores para a compilação e avaliação de dados químicos tornou-se atualmente muito mais claro. Não é apenas o desenvolvimento de computadores cada vez mais sofisticados que está crescendo. O entendimento em relação ao tratamento da base de dados químicos também tem sido aperfeiçoado. Hoje, encontram-se disponíveis no comércio microcomputadores de baixo custo, que apresentam um desempenho satisfatório para muitas aplicações na área da Química Analítica.

O desenvolvimento de uma nova área em Química Analítica, normalmente referida como Quimiometria, é resultado da crescente aplicação de microcomputadores nesta disciplina. Esta nova área pode ser vista como a integração da Ciência da Computação e da Estatística na Química. Através do computador, conceitos matemáticos e estatísticos complexos (por exemplo, aqueles envolvidos nas técnicas de reconhecimento

de padrões (R.P.)^a, análise de componentes principais (PCA)^b, regra do vizinho mais próximo (KNN)^c, etc.), são empregados para a resolução de problemas encontrados na Química Analítica. Alguns dos objetivos importantes em Química Analítica são a extração de informações, tão completas e confiáveis quanto possível, de uma determinada técnica analítica e, também, a obtenção de dados finais da melhor qualidade possível. Em geral, a obtenção destes objetivos resulta principalmente de uma tomada de decisão, onde a resolução do problema torna-se muito dependente de pessoas especializadas. Um procedimento envolvendo conceitos puramente estatísticos não produz, para estes casos, resultados satisfatórios. Este processo pode ser substancialmente melhorado através da implementação de Sistemas Especialistas (S.E.)^d [1].

Um S.E. pode ser definido como a incorporação num programa de computador de um componente baseado no conhecimento de um especialista humano, buscando a solução de um determinado tipo de problema. Desta maneira, o sistema deve oferecer orientação inteligente ou tomar uma decisão inteligente no processamento de uma determinada função[2].

Nos últimos dez anos, a potencialidade dos S.E. tem sido explorada na área de Química Analítica. Apesar de terem sido obtidos excelentes resultados iniciais, em alguns casos, esses não foram estimulantes e o principal motivo foi o entendimento incompleto sobre o que são S.E. e do porque estes nem sempre podem ser aplicados com sucesso. Com a evolução da área de processamento da linguagem natural e robótica, o campo de desenvolvimento de S.E. passou a ser considerado aquele de maior domínio dentro da disciplina conhecida por Inteligência Artificial (I.A.)^e[3]. Se um programa ou sistema é descrito como sendo artificialmente inteligente, espera-se que ele tenha condições

Do Inglês: ^a Pattern Recognition, ^b Principal Component Analysis, ^c K-nearest neighbor, ^d Expert Systems e, ^e Artificial Intelligence.

de fornecer respostas detalhadas para uma ampla faixa de problemas, como por exemplo, a identificação, em qualquer situação, de grupos funcionais e/ou estruturas de compostos através de dados espectrométricos. No entanto, nos últimos dez anos, têm sido desenvolvidos vários S.E. que atuam em domínios específicos da química, buscando a solução de problemas bem definidos. Alguns exemplos dessa natureza são descritos no capítulo I deste trabalho.

Os dois campos da Química nas quais os S.E. apresentam uma maior penetração são o planejamento de sínteses e a elucidação de estruturas de compostos orgânicos[1]. No planejamento de sínteses, o programa especialista elabora o melhor caminho para a síntese de um determinado composto. Normalmente, o programa atua numa retro-direção (*encadeamento para trás*), isto é, encontra aqueles compostos que reagem fornecendo o produto desejado. Este encadeamento para trás implica em um mecanismo de retrocesso, onde pode ser iniciada uma nova busca, para que a meta (neste caso, o melhor caminho da síntese) seja bem sucedida.

Com relação à elucidação de estruturas de compostos, muitas técnicas têm sido desenvolvidas, para se obter informação sobre a composição e estrutura de um composto, dentre as quais, a espectroscopia vibracional no infravermelho (I.V.)[4]. Esta técnica constitui-se em uma ferramenta indispensável para a identificação de grupos funcionais de compostos orgânicos.

Através dos anos em que a espectroscopia no I.V. tem sido utilizada como técnica de elucidação estrutural de compostos, um conhecimento muito significativo foi acumulado. Esse conhecimento é utilizado para estabelecer a correlação entre a presença de um certo arranjo de átomos e a ocorrência de um fato espectroscópico. Também, sabe-se que as propriedades físicas e químicas, assim como as atividades biológicas e farmacêuticas, estão diretamente relacionadas com a estrutura do composto. Portanto, a elucidação de estruturas de compostos, é uma área muito importante de pesquisa em Química Analítica.

A interpretação de espectros no I.V., objetivando a extração das informações relacionadas à estrutura do composto que o originou, é uma tarefa que requer bastante experiência e conhecimento. Como para outras técnicas da Química Analítica, um vasto número de espectros no I.V. encontra-se atualmente disponível, enfatizando assim, a importância das técnicas auxiliadas por computadores particularmente quanto ao processamento e interpretação de dados.

Recentemente, a necessidade de metodologias automatizadas de interpretação espectral tornou-se ainda mais importante, em vista, principalmente, da obtenção de uma vasta quantidade de dados. Algumas técnicas híbridas como a de cromatografia líquida ou gasosa, que utilizam a detecção espectrométrica no I.V. produzem uma tal quantidade de espectros, que a interpretação dos mesmos sem o auxílio de programas computacionais, tornar-se-ia praticamente impossível.

Em vista das tendências até aqui descritas, este trabalho propõe:

1. O desenvolvimento de um programa gerenciador de biblioteca de espectros e aquisição de dados a partir de uma mesa digitalizadora de espectros, de baixo custo, construída neste laboratório, ou através de dados gerados em discos flexíveis pelo microcomputador de um espectrofotômetro de I.V..
2. O desenvolvimento de um programa especialista, capaz de identificar automaticamente grupos funcionais presentes no composto orgânico que deu origem a um espectro de I.V., bem como estabelecer outras características específicas da molécula.

Este trabalho está organizado da seguinte forma: O capítulo I descreve o uso de S.E. em química, mostrando as suas aplicações em Química Analítica, os métodos automatizados para elucidação de estruturas e os principais programas desenvolvidos para tais objetivos. O capítulo II descreve a construção de S.E. e as linguagens computacionais comumente empregadas no seu desenvolvimento. O capítulo III relata o desenvolvimento do programa gerenciador de biblioteca de espectros (coleta, armazenamento e manipulação dos dados espectrais). No capítulo IV, está descrito o programa especialista desenvolvido

e a avaliação do seu desempenho quanto a identificação de grupos funcionais de compostos orgânicos. Em seguida, são apresentadas as conclusões extraídas deste trabalho. Finalmente, no apêndice A, estão descritas as partes mecânica e eletrônica da mesa digitalizadora de espectros construída, no apêndice B, a base de conhecimento do programa especialista e no apêndice C, a relação dos compostos orgânicos interpretados pelo programa especialista através dos seus espectros de I.V..

CAPÍTULO I

SISTEMAS ESPECIALISTAS EM QUÍMICA ANALÍTICA

Na literatura constata-se a descrição de vários S.E. empregados na área de Química Analítica, principalmente dirigidos para algumas técnicas de análises instrumentais[1,5,6]. Por exemplo, espectrometria de massas[7-9], de RMN[10], no I.V., no ultravioleta (U.V.)(11), de análises de ativação por nêutrons, de difração de raio-X[12-14], etc.. Também foi avaliado o desempenho de S.E. em cromatografia líquida e espectrofotômetros de massas quadrupolar[15].

Um sistema denominado MAX (Sistema Especialista para Determinação de Umidade) foi desenvolvido, com o objetivo de selecionar a melhor localização de pontos de amostragem em plataformas petrolíferas do Mar do Norte. Este sistema permite a uma determinação correta dos pontos de amostragem, conduzindo a uma determinação mais confiável dos teores de água em gás natural e glicóis encontrados nas plataformas de extração de petróleo[2,16].

Um S.E. foi desenvolvido com finalidade de preparação de amostras dissolvidas com o auxílio de microondas[17]. Mesmo com os avanços da instrumentação, a preparação das amostras continua uma etapa crítica em Química Analítica. Todavia, a medida instrumental recebe frequentemente, uma maior atenção do que a preparação da amostra. Esta última exige alguns procedimentos como dissolução, diluição e separação, os quais são trabalhosos e consomem muito tempo. A base de conhecimento deste sistema escrito em linguagem Lisp, contém regras para equiparar a matriz da amostra do usuário com uma amostra padrão da SRM (Standard Reference Materials). A base de dados contém informações detalhadas sobre a composição elementar da amostra padrão. Inicialmente, o usuário interage com o S.E. para selecionar a melhor equiparação entre as amostras através da geração de uma série de menus. Quando o sistema adquire informação suficiente do

usuário retorna com o nome da amostra padrão mais semelhante ao da amostra descrita pelo usuário, estabelecendo o melhor procedimento para dissolução da amostra.

Uma revisão da literatura relaciona o uso de S.E. para o auxílio nas decisões quanto ao controle do meio ambiente, como por exemplo, nas situações de controle de resíduos perigosos[18].

O objetivo do S.E., quando aplicado a interpretação de espectros de I.V., é o de gerar uma base de conhecimento constituída das regras, normalmente empregadas pelo especialista humano, na tarefa de analisar e interpretar o espectro. Um tópico que tem recebido muita atenção é o uso da Inteligência Artificial relacionando as informações espectrais para a identificação de compostos. Um exemplo desta aplicação é a interpretação computadorizada de espectros de I.V., RMN-¹³C e de massas, acompanhada com a ajuda de um pequeno S.E.. A base de dados deste sistema é fundamentada nas informações contidas em tabelas de correlações da literatura. Pela combinação destas três técnicas espectroscópicas e o uso de regras de inferência, é possível eliminar algumas interferências espectrais. Conseqüentemente, resultados significativos têm sido obtidos para os compostos conhecidos. O método é capaz de elaborar previsões sobre moléculas ausentes na base de dados, indicando que este sistema pode ser usado para identificação de novos compostos[19].

Conforme os exemplos citados pode-se afirmar que a tecnologia dos S.E. está avançando rapidamente na área de Química Analítica. Algumas pesquisas estão sendo direcionadas visando principalmente os programas de interpretação espectral, através do auxílio na seleção de métodos baseados em técnicas cromatográficas. No item I.2 deste trabalho, estão descritos os principais sistemas desenvolvidos para a elucidação de estruturas de compostos, através das técnicas espectroscópicas, com o auxílio de computadores.

I.1 - SISTEMAS ESPECIALISTAS APLICADOS EM CROMATOGRAFIA

Na aplicação prática da cromatografia líquida (CL), muitos fatores afetam a escolha do modo de separação (fase reversa, troca iônica, etc.), da fase estacionária, da fase móvel e das condições experimentais. Em geral, requisitos como a resolução, sensibilidade e tempo de eluição, variam para os diferentes tipos de amostra. No desenvolvimento de métodos em CL, essas decisões são muito complexas, e normalmente são feitas pelo especialista humano. Entretanto, este procedimento nem sempre é eficiente, e, nem sempre, um especialista está disponível. Estes fatos levaram vários pesquisadores ao desenvolvimento de metodologias automáticas em CL, tentando incorporar a experiência de um especialista da área, dentro de um S.E.[20].

Um S.E. foi desenvolvido para orientar a extração em fase sólida de drogas presentes em amostras aquosas ou de plasma. Esta operação frequentemente é necessária como uma etapa que antecede o uso da cromatografia líquida ou gasosa. A seleção do eluente é baseada em certas características da droga, bem como, nos parâmetros subsequentes da análise, na concentração do analito presente na amostra e na detectibilidade (presença de certas características no analito direcionam a escolha do sistema de detecção a ser empregado). Os parâmetros estudados foram: seleção apropriada do sorvente e dos solventes a serem utilizados para a adsorção, lavagem e eluição[21].

Praticamente, a maioria dos S.E. aplicados à técnica da cromatografia líquida de alta eficiência (CLAE) está relacionada com a otimização das condições operacionais e experimentais. Alguns exemplos destes sistemas estão resumidamente descritos, abaixo.

O projeto ESPRIT (Programa Estratégico Europeu para Pesquisa e Desenvolvimento na Tecnologia de Informação)[22], iniciado em 1987, para estimular a troca de conhecimento entre a universidade e indústria, premiou o projeto ESCA (Aplicação de Sistemas Especialistas para Análises Químicas), o qual tem sido aplicado na Química Analítica em geral, principalmente no tocante às técnicas cromatográficas. A

proposta principal do projeto ESCA é demonstrar a possibilidade da utilização do S.E. em CLAE, aplicada às análises de produtos farmacêuticos e comparar as ferramentas e os métodos a serem utilizados. A razão da escolha deste domínio deve-se ao crescimento do número de compostos na indústria farmacêutica. A princípio, cada composto tem seu próprio método de análise, de modo que a maioria é analisada usando alguma forma de cromatografia, principalmente a CLAE. No momento, o desenvolvimento do método é feito pelo especialista humano, que algumas vezes utiliza os computadores para processar os resultados de seus experimentos. Porém, os computadores ainda não estão totalmente integrados no processo. Neste caso, um S.E. pode acelerar este desenvolvimento, tornando-o mais consistente e melhor documentado e, ainda, oferecendo informações de especialistas, que normalmente não estão disponíveis em todos os laboratórios farmacêuticos[23,24].

Um S.E. foi desenvolvido para otimizar as condições operacionais e instrumentais em CLAE. Em geral, fornecidas as fases móvel e estacionária e os valores para os parâmetros físico-químicos, tais como temperatura e pH, pode-se otimizar uma separação através da variação dos parâmetros cromatográficos. Estes incluem a dimensão da coluna, parâmetros da amostra (volume de injeção, fator de diluição), condições de operação (vazão do eluente, atenuação) e instrumentação (célula do detector, constante de tempo). A otimização desses parâmetros implica em encontrar um melhor conjunto de valores possíveis, como melhor separação e sensibilidade, dentro do menor tempo possível de análise. Nesse caso, o S.E. permite a otimização do procedimento cromatográfico a ser empregado em separações isocráticas. Inicialmente, um cromatograma experimental não otimizado é solicitado pelo sistema, o qual retorna um conjunto completo de parâmetros cromatográficos. Estes representam o ótimo global, dentro dos limites da resolução exigida e razão sinal/ruído especificados pelo usuário[25-27].

O programa denominado ECAT (Equipe Especialista de Assistência Cromatográfica) escrito em linguagem Lisp, auxilia no desenvolvimento de métodos em CLAE[28]. Este programa tem as seguintes metas: a) formulação de experiências no campo geral de S.E., b) desenvolvimento de programas especialistas para solução dos problemas da área de cromatografia e, c) determinação da relação entre S.E. e programas de computação convencionais, focalizando sobre os programas atualmente aplicados na solução de problemas de cromatografia. No ECAT, a representação do conhecimento em CLAE é organizado numa estrutura hierárquica (módulos). Uma das tarefas no desenvolvimento de métodos em CLAE é escolher um método promissor, requerendo especificação do modo de operação, da coluna, da fase móvel e do detector. O primeiro módulo se encarrega destas recomendações, baseando-se nas entradas feitas pelo usuário quanto a descrição das propriedades fundamentais da amostra sendo analisada. No segundo módulo, um método de otimização é utilizado na seleção da composição versus os parâmetros de tempo, para a fase móvel. Finalmente, a maioria das amostras reais contém analitos dentro de uma matriz, requerendo alguma forma de pré-tratamento da amostra antes da análise por CL. Um módulo de preparação da amostra, auxilia o usuário na seleção destes métodos. Desta forma, cada módulo é confinado para uma área podendo operar independentemente e produzir um programa amplo usando entre 300 a 500 regras.

O ESP (Programa Especialista para Separação) auxilia como um consultor de tratamento e diagnóstico para o químico analítico não experiente com CLAE[29]. Ainda que a capacidade para fazer deduções lógicas sobre o método de separação apropriado é reservada aos especialistas, o ESP tenta selecionar um método adequado e sustenta esta orientação com explicações razoáveis. A linguagem Prolog foi escolhida para este projeto desde que, sua natureza descritiva de alto nível permite a tradução direta dos fatos e regras de inferência dentro do código. Esta linguagem utiliza as regras de produção para selecionar e explicar o sistema de separação caracterizado pela coluna, sistema de solvente, tampão e detector.

Um estudo detalhado foi feito do procedimento de retenção de 43 esteróides utilizando CL de fase reversa e fase normal[30]. Embora haja algumas expressões quantitativas para relacionar os parâmetros de otimização cromatográfica para soluto, eluente e propriedades da fase estacionária, na prática a maioria dos especialistas em cromatografia confiam na experiência e informações da literatura. Todavia, tentativas tem sido feitas para explicar tais informações pelo uso de tabelas de seleção ou diagramas de seleção. O S.E. desenvolvido, foi escrito em linguagem Prolog, na tentativa de simular o caminho pelo qual um especialista em cromatografia deve atacar um problema de separação. Nesta aplicação, o problema é selecionar um sistema de separação apropriado para um determinado número de compostos, podendo ser reduzido num gráfico AND/OR que prescreve uma estrutura dirigida por meta (encadeamento para trás), o movimento de uma junção à outra, ao longo de uma ramificação particular, pode exigir dados de entrada do usuário, o qual propõe restrições na busca. Esta estratégia de dados dirigidos objetiva estreitar a busca e assim, reduzir o tempo necessário para alcançar uma solução. Cada junção é uma regra a qual pode ser chamada se a parte condicional da regra é bem sucedida. Vários fatos estão associados com estas regras considerando as propriedades dos esteróides, eluente, fase estacionária e detector e, regras adicionais em que são propostas as restrições para a busca.

A aplicação dos S.E. em técnicas cromatográficas ainda continua sob estudo, entretanto, a tecnologia de S.E. quando comprovada, será muito significativa para os cromatografistas.

I.2 - MÉTODOS AUTOMATIZADOS PARA ELUCIDAÇÃO DE ESTRUTURAS QUÍMICAS

A resolução de problemas é, certamente, o campo mais importante da aplicação da Inteligência Artificial em Química Analítica. Os problemas nesta disciplina surgem como por exemplo, na seleção do método analítico mais apropriado, na otimização automática de um instrumento analítico para a obtenção de resultados precisos e confiáveis num tempo razoável, etc.. No entanto, numa revisão recente da literatura sugere que a área mais explorada pela I.A. dentro da Química Analítica, refere-se à elucidação, auxiliada por computador, de estruturas de moléculas complexas[31].

Os métodos automáticos para elucidação de estruturas, utilizando principalmente a técnica de espectrometria de I.V., podem ser divididos em três categorias: a) técnicas de busca em bibliotecas, b) métodos de reconhecimento de padrões e, c) métodos baseados em conhecimento ou Sistemas Especialistas.

I.2.1 - TÉCNICAS DE BUSCA EM BIBLIOTECAS

Estas técnicas comparam o espectro desconhecido com aqueles presentes em uma base de dados constituída de centenas ou mesmo milhares de espectros de substâncias puras conhecidas. Alguns aspectos importantes, com respeito a esta técnica, são as representações dos dados, as similaridades das medidas utilizadas e o algoritmo de busca aplicado. Todos estes fatores influenciam na velocidade de recuperação e nos processos de comparações. A representação dos dados determina a dimensionalidade de espaço necessário para o armazenamento da coleção dos espectros de referência. Atualmente, sistemas computadorizados de busca e comparação em bibliotecas ainda constituem os mais comumente empregados e, estão disponíveis comercialmente, na maioria dos espectrômetros no I.V..

O sucesso da utilização desta técnica depende da qualidade desta base de dados, e da presença, entre os seus milhares de espectros, de um ou mais que sejam semelhantes ao espectro desconhecido. A inexistência de tais espectros, o que não é de todo improvável, principalmente em se tratando de compostos recém sintetizados ou isolados, apresenta-se como uma limitação desta técnica[15].

A técnica de busca em bibliotecas é utilizada efetivamente para a identificação e interpretação de espectros. O crescente interesse em instrumentos tais como cromatografia gasosa acoplada a espectrômetros de I.V. com transformada de Fourier (FT/IR), conduziram ao desenvolvimento de algoritmos para técnicas de busca mais versáteis e eficientes[32-36].

1.2.2 - MÉTODOS DE RECONHECIMENTO DE PADRÕES

Uma divisão pode ser feita em métodos de R.P. não supervisionados e supervisionados. Os métodos não supervisionados são usados para encontrar classes de espectros similares. Em muitos casos, uma função matemática é aplicada para definir um modelo abstrato de cada classe. Estes modelos são utilizados durante o R.P. supervisionado, a fim de classificar o espectro de um composto desconhecido. Uma vantagem do método de R.P. é que não necessita de nenhuma suposição que estabeleça correlação da informação espectral utilizada para a discriminação entre as classes[4].

Os métodos de R.P. são utilizados por um grande número de grupos de pesquisas, que os empregam na busca de interpretações de dados espectrais[37,38]. Esses métodos tentam encontrar uma relação entre um conjunto característico de propriedades do objeto (por exemplo a presença ou não de um grupo funcional particular), sem utilizar o conhecimento químico ou pré-julgamento químico e também, requerem a disponibilidade de uma coleção de espectros digitalizados. Caso o problema seja determinar se um composto

particular contém ou não o ácido carboxílico como grupo funcional, é necessário uma coleção representativa de espectros de ácidos e de não ácidos. Um estudo empregou um conjunto de dados consistindo de 2600 espectros (200 em cada das 13 classes utilizadas, tais como: ácido, álcool, aldeído, etc.). A capacidade de previsão frequentemente estenderam entre 70 e 90 % de exatidão. Esses resultados são considerados satisfatórios, porém, um número maior de classes de compostos teria que ser avaliado. Em geral, a adição de mais classes implica em aumento no tamanho do conjunto de dados e, portanto, muito da velocidade do processamento e da facilidade do gerenciamento do R.P. são perdidos. Portanto, o uso dos métodos de R.P. como uma ferramenta para resolver os problemas atuais dos espectroscopistas se mostra, às vezes limitado[15]. Todavia, durante a próxima década, os S.E. poderão auxiliar os métodos de R.P. fornecendo recursos, que limitam o tamanho do conjunto de dados a ser processado para o reconhecimento de uma determinada função.

1.2.3 - SISTEMAS ESPECIALISTAS

Ao contrário dos métodos de R.P. e técnicas de busca, os quais requerem grandes bibliotecas de espectros de referência, os métodos baseados em conhecimento ou Sistemas Especialistas (S.E.) utilizam uma base de conhecimento que contém o conhecimento químico de um especialista para resolver um problema de domínio limitado. Estes sistemas compreendem um avançado tipo de programa computacional, que permite adquirir o conhecimento de especialistas, para ser codificado e utilizado de forma flexível na resolução de problemas, simulando o pensamento humano.

A maioria dos S.E. atuais é baseada em conhecimento, sendo limitados e dirigidos para a solução de um problema específico da área. Atualmente, um S.E. é "especialista" apenas na sua capacidade de solucionar o problema de seu domínio

específico, com informação e perícia localizadas em sua base de conhecimento. A combinação da capacidade de raciocínio, capacidade de indagação, habilidade de fornecer explicações e conveniências de uso, tornam estes sistemas poderosos e fáceis de utilizar. Eles podem funcionar como programas independentes, ou mesmo serem integrados com programas convencionais, para torná-los mais inteligentes e convenientes para a sua utilização[39].

1.2.4 - MÉTODOS HÍBRIDOS

Na tentativa de suprir certas deficiências de um método de elucidação de estruturas em relação a outro, foram desenvolvidos alguns trabalhos, onde mais de um método, dos citados anteriormente, são sequencialmente empregados.

Normalmente, na junção dos métodos de R.P. com a técnica de busca em biblioteca, é realizada uma análise dos componentes principais (PCA), objetivando reduzir os dados de uma biblioteca de espectros e o tempo computacional necessário para a busca. A PCA calcula um novo conjunto de eixos e coordenadas, que reduzem a dimensão dos dados no espaço original. Deste modo, os espectros são projetados dentro de um plano principal, onde são representados por um único ponto, num espaço bi-dimensional[40,41].

No caso da junção de métodos de R.P. com S.E., existe um programa denominado SPECTRE, para a análise de dados quimiométricos, que permite a modelagem e a otimização de experimentos multivariados. O uso adicional do S.E. define as regras (entrada) para estabelecer qual procedimento estatístico é mais adequado na resolução do problema[42].

Um outro exemplo dessa natureza é a integração da metodologia de otimização com S.E., na avaliação de 20 drogas básicas e suas 60 misturas sintéticas, utilizando a técnica de CL. Este S.E. usa regras de produção, contendo o conhecimento e a experiência

desenvolvida no próprio laboratório para os métodos de seleção de CL, nas análises de drogas. O sistema fornece orientação sobre o detector e a fase móvel e, foi aplicado a mais de 50 preparações de drogas, com sucesso em 87 % dos casos. Além disso, inclui um módulo de planejamento experimental, ou mais especificamente, uma otimização por planejamento fatorial. A proposta do referido trabalho foi investigar a possibilidade de interagir experimentalmente a metodologia de otimização dentro de um S.E. e, não para estabelecer o melhor método de otimização[43].

Um S.E. baseado em PCA foi desenvolvido para a identificação de estruturas moleculares, através de espectros de I.V. na fase de vapor. PCA tem sido utilizada para desenvolver um conjunto de regras para S.E., fornecendo um meio eficiente para descrever de maneira quantitativa uma regra, que pode distinguir a presença de um grupo funcional num composto a partir de seu espectro. Um conjunto de treinamento é selecionado para duas classes de compostos, ou seja, aqueles contendo o grupo funcional de interesse e aqueles sem o mesmo. Os "scores" das amostras no primeiro, segundo (ou às vezes terceiro) componente principal devem mostrar uma separação distinta das duas classes de compostos. Cada regra desenvolvida neste sistema estabelece se ou não a presença de um certo grupo funcional no composto[44].

I.3 - PROGRAMAS PARA ELUCIDAÇÃO DE ESTRUTURAS QUÍMICAS

Em 1969, na Universidade de Stanford, os computadores foram aplicados pela primeira vez, com a finalidade de automatizar as investigações sobre as estruturas de compostos, através das interpretações de espectros de massas[31].

O sistema original DENDRAL[45] foi um dos primeiros programas a implementar um S.E. capaz de interpretar dados de espectrometria de massas e de RMN-¹³C. Este sistema utilizava dados de massas para prever a estrutura de moléculas desconhecidas, usando regras em torno de possíveis estruturas e processos de fragmentações no espectrômetro. Estas regras foram obtidas de um especialista, ou derivadas por um programa associado, denominado Meta-Dendral[46], o qual analisava o espectro de massas produzido a partir de um composto de classe particular, para sugerir prováveis regras de fragmentações.

Com o sucesso do projeto DENDRAL, outros centros de pesquisas foram motivados a desenvolver diferentes sistemas automatizados para a elucidação de estruturas de compostos. Desde então, foram construídos alguns programas complexos de computador, com a capacidade de interpretar diferentes tipos de dados espectrais, gerando estruturas moleculares, e calculando espectros teóricos[47], como por exemplo: CASE[48], STREC[49], ESSESA[50-52], ASSIGNER[53], CHEMICS[54-56], SEAC[57], PAIRS[37,58-71], EXPERTISE[72] e EXSPEC[47,73-77].

Estes sistemas possuem certas características comuns em termos de Inteligência Artificial. Um exame dos métodos de resolução de problemas em química analítica, indica que a maioria envolve a idéia da busca por tentativa e erro, isto é, os problemas são resolvidos buscando-se uma solução num espaço de possíveis soluções.

Na elucidação de estruturas, é comum obter-se um conjunto de estruturas candidatas, que podem ser geradas de sub-estruturas, detectadas durante a análise espectral da molécula. Todavia, vários procedimentos heurísticos (por exemplo, uma base de dados

específica para o sistema), podem ser aplicados para reduzir o espaço da solução, diminuindo assim, o tempo computacional necessário para a identificação de sub-estruturas[31].

A maioria dos sistemas existentes utiliza procedimentos dirigidos por modelos, isto é, o sistema cria uma hipótese para uma parte da estrutura do composto, e passa a testá-la, de modo que ela pode ser obtida por diferentes caminhos. Um destes é atravessar uma lista de unidades estruturais pré-definidas como: CH₃, CH₂, CH, C, OH, O, etc.. Este procedimento é usado pelos sistemas CHEMICS, CASE, STREC e SEAC. Uma outra possibilidade é o procedimento utilizado pelo programa PAIRS, o qual armazena as unidades estruturais de uma maneira ramificada. Neste caso, a primeira etapa é verificar a presença ou a ausência de um número de unidades estruturais menores. Então, dependendo dos resultados, são processadas as unidades maiores, que são extensões das unidades encontradas[47]. Por exemplo, a presença de um grupo aromático tem que ser estabelecida antes que seja feita a análise da substituição neste grupo.

Como uma terceira opção, pode-se armazenar as sub-estruturas numa rede neural, desenvolvida inicialmente, como um modelo do cérebro humano. As sub-estruturas (ligações) entre os nós na rede, contêm o conhecimento da hierarquia das unidades estruturais. Por exemplo, um fenol é definido como sendo aromático, e também um álcool. Deste modo, inicia-se na raiz da rede a geração de hipóteses e apenas aqueles nós são verificados. Neste caso, todos os nós encontrados anteriormente, estarão possivelmente presentes no composto de interesse. Este procedimento é utilizado pelos sistemas EXPERTISE[72] e EXSPEC[47].

O sistema computadorizado ASSIGNER foi planejado para auxiliar na elucidação da estrutura de compostos orgânicos. Ele pode operar em linha ou não, produzindo uma lista de possíveis grupos funcionais, consistentes com os dados experimentais e com a informação química e espectroscópica armazenada. Os resultados da interpretação são obtidos com base na comparação das propriedades determinadas

experimentalmente e dos fragmentos estruturais particulares, com aqueles armazenados na memória do computador. As propriedades dos grupos funcionais, incluindo a informação química inerente e, a requisitada para a redução da lista, são armazenadas nas bases de dados chamadas de matrizes. As matrizes dessas propriedades contêm os dados espectrais de I.V., RMN de ^{13}C e de ^1H . Os grupos funcionais são fragmentos de uma estrutura química, definidos pelas suas propriedades. Em alguns casos as propriedades adequadas são essenciais para deduzir a presença de um grupo funcional existente no composto. A maior parte desse sistema foi escrita em linguagem FORTRAN com algumas partes em linguagem ASSEMBLER. Os testes realizados com o sistema mostraram que o número de grupos identificados como possivelmente presentes, normalmente é duas vezes o número de grupos verdadeiramente presentes. Porém, nenhum dos grupos verdadeiramente presentes é perdido[53].

O sistema CHEMICS foi desenvolvido para analisar vários dados espectrais de uma amostra desconhecida e fornecer uma estrutura compatível com estes dados. Entretanto, a maior parte das informações existentes nos dados espectrais de I.V., não pode ser utilizada pelo sistema original, uma vez que as sub-estruturas originais do sistema foram adaptadas para o processamento de dados de RMN- ^{13}C e ^1H . Para auxiliar o sistema CHEMICS, como uma opção independente do processamento de dados de I.V., foi elaborado o programa IRASSL (Sistema de Análises de Infravermelho baseado em Lógica Simbólica), onde a informação dos espectros de I.V. é transformada em símbolos, com a finalidade de formular o procedimento analítico. Estes símbolos são tratados em operações aritméticas simples, de acordo com as regras de lógica simbólica. Como resultado, o sistema indica a possível existência de sub-estruturas e sua conexão com os compostos desconhecidos. Quanto às sub-estruturas, 30 grupos funcionais foram classificados e as tabelas foram preparadas para correlacionar uma certa sub-estrutura que poderia estar ligada a grupo funcional. A introdução da informação da sub-estrutura fornecida pelo

IRASSL no CHEMICS, como uma restrição no processo de elucidação, produz uma diminuição significativa no número de estruturas candidatas[54-56].

O sistema SEAC (Elucidação da Estrutura Auxiliada por Computador) foi desenvolvido para elucidar a estrutura de compostos orgânicos contendo até 40 átomos de carbono na molécula e átomos de H, N, O, S, Cl, Br e/ou I. As técnicas de Inteligência Artificial são exploradas para interpretação das correlações estrutura/espectro (processamento heurístico) diretamente dos espectros de U.V. e RMN-¹H. A identificação da estrutura é baseada em parâmetros calculados através de programas computacionais específicos, para tratamento matemático dos espectros. Os parâmetros se diferenciam conforme o espectro em análise. Por exemplo, para os espectros de I.V. consistem na posição da banda (cm^{-1}), intensidade (%T) e largura a meia-altura (cm^{-1}), enquanto que, para os espectros de RMN, consistem na localização (τ , ppm), intensidade (%T), e constantes de acoplamento e multiplicidade (Hz/100). O espectro de U.V. especificado pela posição da banda (nm) e intensidade (A), também é usado apenas para verificar os resultados obtidos. O sistema é baseado no princípio de processamento heurístico de correlações espectro/estrutura, executado em tres estágios sucessivos. No primeiro estágio, na base de dados espectral e informação preliminar, onde são identificadas as sub-estruturas provavelmente presentes no composto. No segundo estágio, é verificada a exatidão da identificação. No último estágio, um conjunto multi-elementar de fragmentos estruturais é selecionado e, então, é gerado um conjunto completo de estruturas que são igualmente válidas ao da informação dada[57].

O programa PAIRS (Programa para Análises de Espectros de Infravermelho) foi construído para auxiliar o químico em determinar quais grupos funcionais podem estar presentes num composto desconhecido[58]. Embora tenha sido originalmente destinado para interpretar espectros de I.V. de espécies na fase condensada (líquida ou sólida), recentemente, foram implementadas regras para espécies na fase de vapor[59]. A saída da interpretação do espectro analisado é feita através da linguagem FORTRAN, explicitando a

probabilidade de alguns grupos funcionais presentes na molécula responsáveis pelo espectro. As regras utilizadas pelo PAIRS não são codificadas em linguagem FORTRAN, mas sim expressas numa linguagem similar ao inglês chamada CONCISE (Avaliação de Espectros de Infravermelho Orientada por Computador)[60-62].

Na maioria dos sistemas desenvolvidos foi empregada na interpretação de espectros de compostos puros. No entanto, recentemente, alguns trabalhos têm sido publicados com a atenção voltada para análises de espectros de misturas de compostos, como por exemplo o programa PAWMI (Programa para Interpretação de Misturas de Resíduos), que foi aplicado em amostras de resíduos tóxicos, através da espectroscopia no I.V.[62-65]. O programa PAIRS é utilizado pelo PAWMI como uma sub-rotina de interpretação do espectro. Mas, duas limitações são apresentadas pelo PAWMI. Uma vez definido um conjunto de treinamento, consistindo de uma biblioteca de espectros de referência, as regras para o mecanismo de inferência (PAIRS), eram geradas manualmente[58,59,66-69]. A segunda limitação diz respeito a identificação de compostos somente através do uso da posição do pico. Para automatizar as regras de interpretação, foi desenvolvido o MISIP (Pacote para Interpretação Espectral de Infravermelho da Merck), que consiste em uma versão do programa PAIRS[62]. Com o objetivo de identificar compostos de misturas, alguns pesquisadores incluíram programas com árvore hierárquica no programa PAIRS[37], dirigidos por tabelas[70] e métodos de R.P.[71].

O sistema EXPERTISE é baseado em conhecimento e foi desenvolvido para avaliação de espectros de I.V.. A base de conhecimento é dividida em conhecimento da estrutura e da sub-estrutura. O conhecimento da estrutura consiste de uma biblioteca de espectros e as fórmulas químicas associadas do composto. O conhecimento da sub-estrutura, implica as propriedades espectrais e fórmulas das sub-estruturas. As regras do S.E. para encontrar características de sub-estruturas de um espectro desconhecido são obtidas por um procedimento de R.P., evitando assim, a aquisição de conhecimento feita manualmente. Um algoritmo para geração da estrutura liga, a sub-estrutura particular aos

chamados super-átomos (um elemento com todas as suas ligações no composto). Alguns operadores de conjuntos difusos ("fuzzy sets") foram utilizados para considerar inexatidões nas posições e intensidades dos picos. O desempenho do sistema foi testado com um conjunto de 900 sub-estruturas geradas de 100 compostos acíclicos. As sub-estruturas dos compostos do conjunto de treinamento foram todas reconhecidas corretamente e as sub-estruturas dos compostos testes atingiu um reconhecimento de 95 %. A geração de estrutura foi testada com 20 compostos acíclicos, dos quais 12 foram identificados corretamente, de forma que para os demais, o sistema encontrou uma estrutura química maior. Isto foi devido principalmente as sub-estruturas encontradas em suas buscas, que produziram muitos grupos $-CH_2$ [72].

O sistema EXSPEC, baseado em conhecimento, é capaz de interpretar dados espectrais combinados (por exemplo, I.V. e massas). Atualmente, o sistema consiste de três partes principais: um gerador de regras[73], um interpretador de espectros e um gerador de estruturas[74]. A aquisição do conhecimento é fundamentada em bases de dados espectrais e estruturais, e não em tabelas de correlações da literatura. A base de dados spectral consiste de espectros de I.V. e de massas, numa forma reduzida, isto é, tabelas de picos de números de ondas, ou valores massa/carga (m/z) e intensidades. A base de dados estrutural converte a estrutura molecular de cada composto num formato de lista. Cada lista consiste das ligações e unidades estruturais presentes, por exemplo: CH_3 , CH_2 , CH , C , OH e O , mas, pode ser estendida para incluir outros elementos e tipos de ligações (simples, dupla, tripla e aromática). A linguagem Prolog foi escolhida para manipular a informação simbólica. A interpretação do espectro procede através da seleção de uma sub-estrutura "S" de uma rede. Todas as funcionalidades originais de "S" são extraídas de uma rede e a sua presença ou ausência no composto é estabelecida. Por exemplo, para estabelecer a presença de um grupo fenólico num composto, este tem que ser aromático e conter um grupo alcoólico ($-OH$)[47,75-77].

Todos os sistemas acima descritos diferem no tamanho, no tipo de espectros que podem ser manuseados (I.V., massas, RMN- ^{13}C e ^1H , U.V., etc.) e na quantidade de informação que é extraída e utilizada dos dados experimentais. Alguns pesquisadores manifestam-se favoráveis ao sistema EXSPEC e PAIRS, porque estes são capazes de explicar os resultados obtidos durante uma interpretação.

A tabela I-1 mostra esses sistemas utilizados para elucidação de estruturas e relaciona os dados espectrais que são manipulados.

Nos últimos anos, a aplicação das redes neurais nas interpretações de espectros, tem despertado uma atenção considerável. Uma rede neural é um modelo simplificado do cérebro humano, consistindo de algumas camadas de neurônios, que trocam sinais entre si, dependendo dos sinais recebidos. A rede pode ser treinada para fornecer uma saída, em termos de informação estrutural, a partir do espectro utilizado como entrada. Esta área de pesquisa abre perspectivas para novas conclusões com relação à natureza do processo de elucidação de estruturas de compostos[4,78]. Na literatura, alguns trabalhos têm sido publicados, onde a rede neural é utilizada em conjunto com S.E., para a interpretação de espectros de I.V.[79-82].

No Brasil, ainda há poucos grupos de pesquisadores desenvolvendo S.E. direcionados a resolução de problemas em química. Um sistema descrito na literatura é o SISTEMAT[83-85] aplicado à Química de Produtos Naturais. Esse sistema foi originalmente escrito em linguagem FORTRAN e, atualmente, está sendo reescrito em linguagem PASCAL. Uma das suas metas é a compactação de todas as informações químicas obtidas da fórmula química de um composto. Neste caso, o sistema deve ser capaz de extrair funções químicas, fórmula bruta, peso molecular, índice de oxidação, etc., através de uma matriz topológica. Essa matriz é um recurso matemático para representar a interligação de vértices (tipos de ligações químicas) aos nós (correspondentes aos átomos). A compactação é uma técnica de redução de vetor, isto é, uma relação sequencial das posições (linhas e colunas da matriz) e os valores dos elementos não nulos da matriz. Desta

forma, o método de codificação desenvolvido deve conduzir a um vetor menor possível, com todas as informações necessárias para a criação da matriz topológica, conseqüentemente, desenhando a molécula na tela do computador. Uma meta posterior do SISTEMAT refere-se à exploração das informações botânicas contidas nos bancos de dados, com finalidades taxonômicas. O armazenamento de dados espectrais RMN-¹³C das substâncias de origem natural, juntamente com as informações botânicas (família, gênero, espécie) e a classe do produto natural (terpenóide, alcalóide, etc.), permitem que o próprio sistema encontre regras de classificação para as estruturas destes compostos.

TABELA I.1 - Programas para elucidação de estruturas químicas, com os tipos de dados espectrais e linguagens computacionais utilizadas.

SISTEMA AUTOMATIZADO	DADOS ESPECTRAIS	LINGUAGEM COMPUTACIONAL
CHEMICS e IRASSL	I.V., RMN- ¹³ C e ¹ H	FORTRAN 77
ASSIGNER	I.V., RMN- ¹³ C e ¹ H	FORTRAN e ASSEMBLER
SEAC	I.V., RMN- ¹ H e U.V.	FORTRAN
EXPERTISE	I.V.	FORTRAN
PAIRS	I.V.	FORTRAN e CONCISE
SISTEMAT	RMN- ¹³ C	FORTRAN e PASCAL
EXSPEC	I.V. e Massas	PROLOG

CAPÍTULO II

CONSTRUÇÃO DE SISTEMAS ESPECIALISTAS

Um problema em química pode ser adaptado para ser solucionado através de um S.E., se há um especialista apropriado ou uma fonte de conhecimento. Um especialista é alguém perito em um campo particular, capaz de fornecer orientações úteis à resolução de um problema de sua área. Então, por analogia, um programa de computador especialista deve preencher a mesma função. Num contexto geral, é praticamente impossível avaliar um S.E. ou um especialista humano, isto é, a avaliação de ambos deve ser feita após uma soma de resultados acumulados com o tempo. Desta maneira, os critérios a serem adotados na avaliação, assim como os indicadores de sucesso, necessitam ser especificados claramente no início da elaboração do sistema[2].

A pessoa responsável para extrair o conhecimento do especialista é denominado "engenheiro de conhecimento". Ele atua seletivamente de forma a obter do especialista o conhecimento relevante dentre todas as informações que este dispõe e transformá-las numa forma útil para a computação. A sua função é mostrada esquematicamente na figura II-1.

Um S.E. consiste de três componentes principais: a) uma base de conhecimento, b) um mecanismo de inferência, e c) uma interface homem-computador[39]. A figura II-2 ilustra os componentes e o uso de um sistema baseado em conhecimento.

A base de conhecimento consiste de fatos atualmente conhecidos e conhecimentos heurísticos representados na forma simbólica, de maneira que seu conteúdo possa ser alterado. Os fatos compreendem um conjunto de informações, que normalmente estão de acordo com o especialista da área. O conhecimento heurístico consiste de regras de uso comum que caracterizam a tomada de decisão do especialista no seu domínio de atuação. Esta base de conhecimento também armazena as regras de inferências, que são

geralmente fixas. O desempenho de um S.E. é uma função importante do tamanho e da qualidade de sua base de conhecimento.

Uma vez que a primeira versão de um S.E. foi construída, é necessária a avaliação do conhecimento formalizado, através de um mecanismo de inferência e alguma forma de interface "homem-computador" para habilitar o usuário a comunicar-se com o computador.

O mecanismo de inferência é a parte de raciocínio de um sistema especialista, que atua sobre a informação armazenada (conhecida), buscando respostas para as questões. Ele contém os componentes necessários para interpretar os símbolos e fornecer as técnicas de resolução do problema. Para isso, utiliza regras, fatos existentes, e informações obtidas do usuário, no sentido de deduzir novos fatos. O interpretador executa a lógica necessária para comparar a informação na base de conhecimento com fatos do problema sob consideração, seguindo as funções principais: a) determinar quando e onde iniciar uma consulta, b) implementar o procedimento de busca, e c) determinar que uma conclusão foi alcançada ou que nenhuma pode ser atingida.

Uma vez que o mecanismo de inferência foi depurado, ele não necessita ser depurado toda a vez que o sistema é modificado ou estendido. Isto não procede para os programas desenvolvidos usando algoritmos de linguagens convencionais (de procedimentos como FORTRAN, PASCAL e BASIC), os quais devem ser depurados repetidamente em cada modificação executada.

A interface usuário-computador facilita o diálogo entre o usuário e o S.E.. As duas características do sistema, resultante da interação entre a base de conhecimento e o mecanismo de inferência, são a capacidade para explicar ao usuário "PORQUE" uma questão particular está sendo indagada, e "COMO" uma solução foi alcançada. Estas características, quando combinadas, formam uma poderosa e eficiente interface com o usuário.

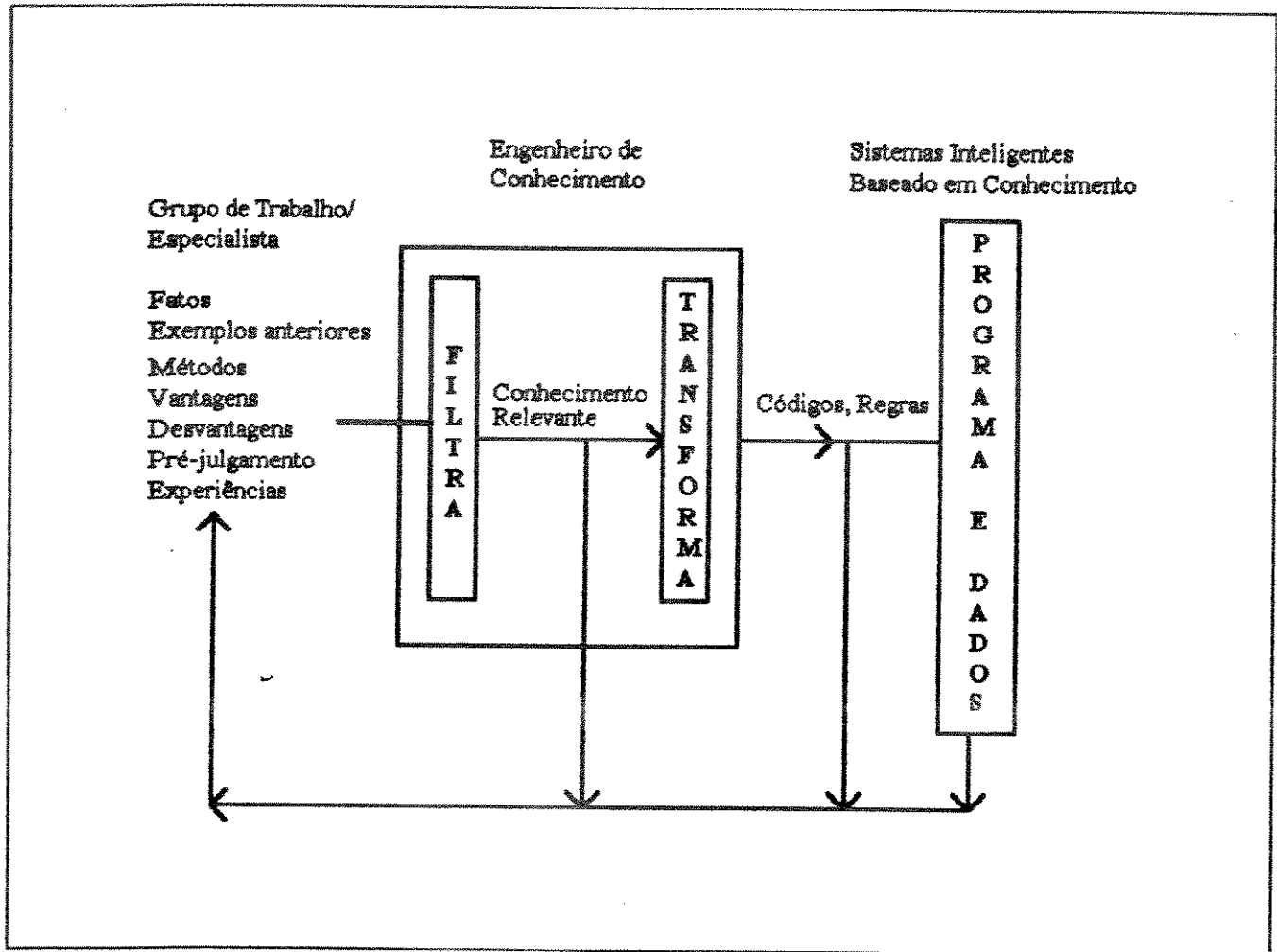


FIGURA II.1 - Função do engenheiro de conhecimento.

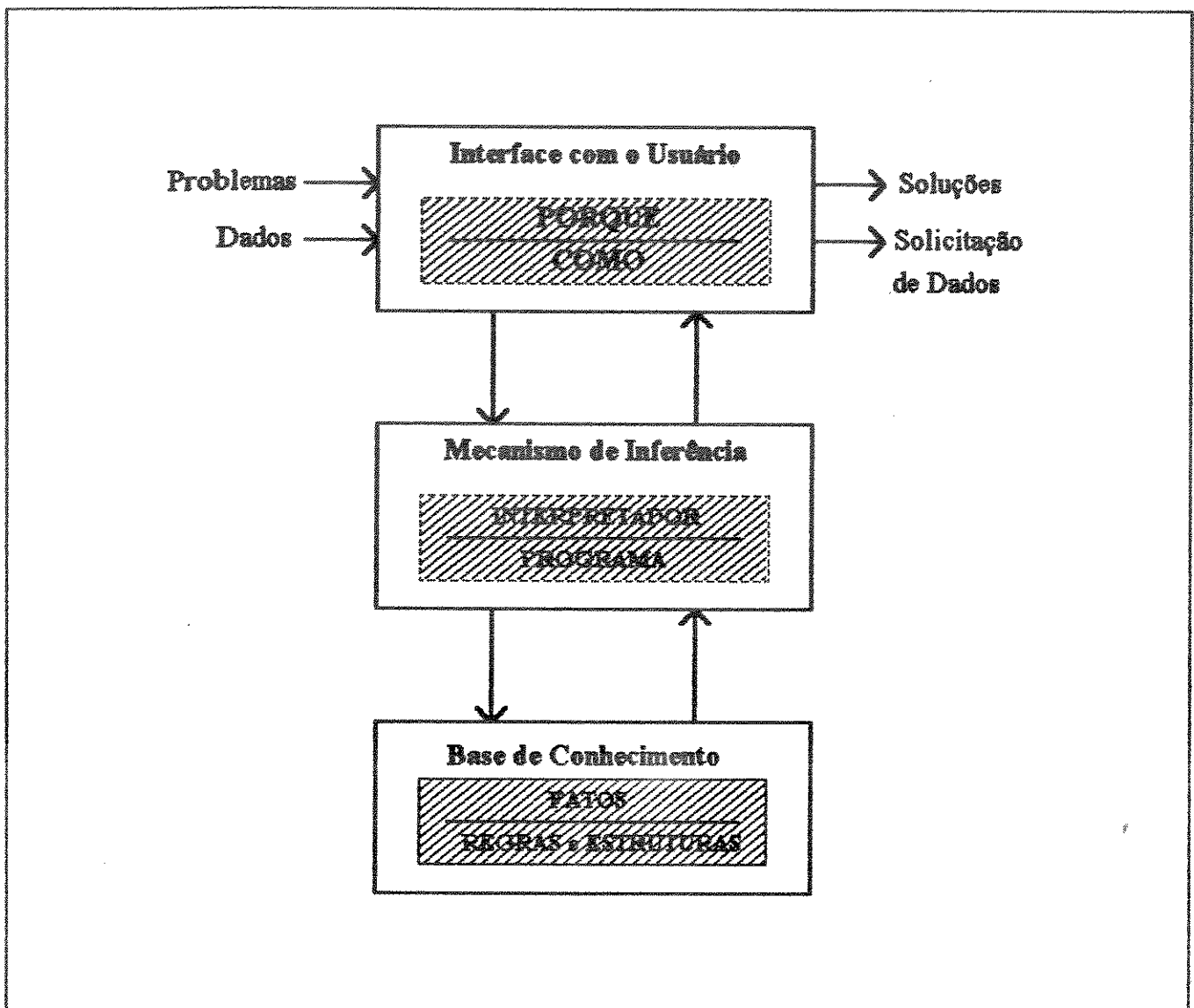


FIGURA II.2 - Componentes de um Sistema Especialista ou sistema baseado em conhecimento. A base de conhecimento contém os fatos e as regras de produção representados na forma simbólica, o mecanismo de inferência controla as informações da base de conhecimento e a interface com o usuário facilita o diálogo entre o usuário e o sistema, fornecendo informações sobre o porque uma questão está sendo formulada e como o sistema chegou a solução do problema.

II.1 - LINGUAGENS COMPUTACIONAIS UTILIZADAS EM PROGRAMAS ESPECIALISTAS

Um S.E. deve ser construído de modo que ele possa ser atualizado se assim exigido. Para isto, a programação baseada em regras é melhor adaptada do que os programas escritos com estilos convencionais como BASIC, FORTRAN e PASCAL. As duas principais linguagens utilizadas em programas especialistas são LISP (Processamento de Listas)[86] e PROLOG (PROgramando em LÓGica)[87-91].

A linguagem Lisp é muito adequada e é considerada a linguagem proeminente da inteligência artificial. Ela fornece um ambiente amplo para programação e depuração. A principal desvantagem apontada em escrever diretamente em Lisp é a velocidade lenta de desenvolvimento do sistema. Além disso, Lisp não possui recurso gráfico direto e, por isso, deve chamar programas escritos em outras linguagens a partir do programa Lisp[15].

Em 1972, na França, o Grupo de Inteligência Artificial de Marseille dirigido por Alain Colmerauer, criou a linguagem Prolog, na tentativa de incorporar numa linguagem de computador, o desenvolvimento mais recente da lógica matemática. Hoje, os interpretadores Prolog estão disponíveis para microcomputadores tais como Apple-II, Macintosh, IBM-PC[92].

Tipicamente, um programa Prolog não é uma sequência de ações, mas uma coleção de fatos juntamente com regras, para obter conclusões destes fatos. A linguagem Prolog é declarativa, ou seja, dados os fatos e regras necessários, ela usa o raciocínio dedutivo para resolver seus problemas de programação, ao contrário das linguagens de computador tradicionais. Nas linguagens de procedimentos, o programador deve fornecer instruções passo a passo, para informar o computador exatamente como resolver um determinado problema, isto é, o programador deve conhecer como resolver o problema antes de executá-lo no computador. O programador da Prolog necessita apenas fornecer uma descrição do problema e as regras fundamentais para resolvê-lo.

Atualmente, Prolog é uma ferramenta muito importante na programação das aplicações de Inteligência Artificial. Ela revelou ser bem adequada para processamento de linguagem natural, bases de dados interativos, resolução de problemas, equiparação de padrão, e o desenvolvimento de S.E. em Química[93-99].

As linguagens Lisp e Prolog foram avaliadas e comparadas, como ferramenta para a construção de S.E.. Em geral, elas apresentam os seus prós e contras, e cada uma tem um grupo de entusiásticos seguidores. Os da linguagem Lisp estão localizados principalmente nos Estados Unidos mas, e os da Prolog encontram-se na Europa e Japão[98,99].

II.1.1 - SINTAXE DA LINGUAGEM TURBO PROLOG - VERSÃO 2.0

- SENTENÇAS: FATOS E REGRAS

Em Prolog, um programador define primeiro objetos e relações. Daí, podem-se definir regras, quando estas relações são verdadeiras. Por exemplo, a sentença:

Vibração de estiramento O-H detectada em álcool

mostra uma relação entre os objetos "Vibração de estiramento O-H" e "álcool", sendo que "detectada" é a relação.

O nome simbólico de uma relação é chamado de predicado. Os objetos que ele relaciona são chamados de argumentos. Uma relação entre objetos é chamada um fato e, é simbolizada por uma sentença. Na lógica de predicado que Turbo Prolog utiliza, uma relação é resumida numa frase simples, um fato, que consiste do nome da relação seguido pelo objeto ou objetos (incluídos em parênteses). Por exemplo:

detectada(vibração_de_estiramento_OH)).

onde "detectada" é o predicado e "vibração_de_estiramento_OH" é o objeto. Ambos são identificadores e a combinação dos dois é um termo. Qualquer número de argumentos pode ser usado com um predicado.

A regra é uma conclusão que é conhecida como verdadeira se uma ou mais conclusões ou fatos são verdadeiros. Uma regra pode definir a sentença citada como verdadeira: Estiramento O-H detecta um álcool se o espectro é de infravermelho e tem vibrações de estiramento O-H e C-O.

Uma regra em Prolog constitui de três partes: a cabeça, o corpo e um símbolo hífen (-) separando os dois, como por exemplo:

```
detecta(estiramento_OH,alcool) :-  
    espectro(infravermelho),  
    tem_vibracoes([estiramento_O-H,estiramento_C-O]).
```

onde a cabeça da regra é o predicado "detecta" e o corpo da regra, após o símbolo hífen, consiste de outras dois sub-metas ("espectro" e "tem_vibracoes"). As sub-metas são separadas por vírgulas (*and*), sendo que a última sub-meta é terminada por um ponto (.). É importante salientar que cada sub-meta chama outro predicado Prolog, que é testado para verificar a sua veracidade. Em seguida, se a sub-meta corrente (tem_vibrações) foi satisfeita, a chamada é retornada e o processamento continua até atingir o ponto final da última sub-meta (banda_da_carbonila) com o mesmo sucesso. Se assim ocorrer, a regra foi bem sucedida. Neste exemplo, simplificado, a regra pode ser utilizada para descrever uma regra geral para a identificação de um determinado grupo funcional a partir de dados espectrais.

- VARIÁVEIS

Uma variável em Turbo Prolog deve começar com uma letra maiúscula, depois pode-se utilizar qualquer número de letras (maiúscula ou minúscula), dígitos ou o carácter de travessão (`_`), como por exemplo:

`Nome_do_Composto`, `Intensidade`, `Posição_1`, `Largura_da_Banda`, `Absorção1e2`

A variável anônima é representada por um sinal de travessão ("`_`"), e empregada apenas quando se necessita determinada informação de uma questão do programa, ignorando o seu valor.

- SEÇÕES BÁSICAS DE UM PROGRAMA

Basicamente, há dois tipos de frases em linguagem Prolog, isto é, uma frase pode ser um fato ou uma regra. Estas frases são conhecidas como cláusulas. Um programa Turbo Prolog inclui as seguintes seções básicas: a) domínios, b) predicados, c) cláusulas e d) metas. Outras seções comumente utilizadas num programa são: a seção da base de dados, a seção de constantes e as várias seções globais.

A - Seção de Domínios

Os objetos numa relação (os argumentos para um predicado) pertencem aos domínios. Portanto, o domínio declara o tipo de argumento que o predicado vai utilizar. Estes podem ser domínios padrões ou domínios especiais, ou seja, domínios que são especificados pelo usuário. Alguns exemplos de domínios padrões são: (inteiro, real, sequência de caracteres, símbolo), etc.. Um exemplo de domínio especial é mostrado abaixo:

`banda_identificada(string, string, integer, string)`

o qual, após a solução de um determinado problema, pode retornar a seguinte resposta:

`banda_identificada("cetona", "estiramento_C=O", 1715, "cm-1")`

B - Seção de Predicados

A seção de predicados, é onde cada predicado é relacionado, mostrando os tipos (domínios) de seus argumentos. O nome de um predicado deve começar por uma letra minúscula, seguido de uma sequência de letras, dígitos e travessões, até 250 caracteres, se necessário. Por exemplo: `estado_da_amostra`, `interpretação_de_Espectros`, `espectros_de_RMN_e_de_Massas`, `analisa_1_Dado`, etc..

C - Seção de Cláusulas

Nesta seção, colocam-se os fatos e regras que Turbo Prolog pode operar quando tenta atingir com sucesso a meta do programa. As cláusulas para um determinado predicado devem ser colocadas juntas na seção de cláusulas; uma sequência de cláusulas definindo um predicado é chamado de procedimento.

D - Seção de Metas

É a seção onde se coloca a meta que se deseja alcançar. Esta é a seção que informa o que tenta descobrir ou o que o computador deve fazer com as informações fornecidas nas outras seções.

E - Seção de Constantes

Nessa seção pode declarar e usar constantes simbólicas, indicadas pela palavra chave "*constants*", seguida pela declaração desejada, usando a seguinte sintaxe:

<Id> = <Macro definition>, onde <Id> é o nome da constante simbólica, e <Macro definition> é o que define como constante. Por exemplo:

```
constants
```

```
    intensidade_forte = 25
```

```
    derivada_minima = 0.08
```

F - Seção da Base de Dados

Um programa Turbo Prolog é uma coleção de fatos e regras. Algumas vezes, enquanto o programa está em execução, deseja-se atualizar (alterar, remover ou adicionar) alguns destes fatos. Em tais casos, os fatos constituem uma base de dados dinâmica ou interna. A palavra chave "*database*" declara a seção da base de dados, na qual pode adicionar fatos, mas não regras.

Os predicados padrões da base de dados interna são: a) "*assert*", que insere um novo fato na base de dados, b) "*retract*", que remove o primeiro fato na base de dados, quando equipara com este fato, c) "*consult*", faz a leitura num arquivo (Nome do arquivo), contendo fatos declarados num seção da base de dados, e insere-os no programa no final da base de dados apropriada, e d) "*save*", que salva os fatos de uma determinada seção da base de dados para um arquivo.

- RETROCESSO OU ENCADEAMENTO PARA TRÁS (BACKTRACKING)

O retrocesso ou encadeamento para trás é o mecanismo que fornece instruções onde a Prolog deve procurar a solução para o problema no programa. Este processo determina a capacidade de busca, através de todos os fatos e regras conhecidos para uma solução. Para se obter uma regra com sucesso, Prolog deve satisfazer todas as sub-metas existentes. Se uma sub-meta falhar, volta e procurar alternativas nas sub-metas anteriores, e então procede a busca novamente para a frente, com diferentes valores das variáveis. Há quatro princípios básicos para o retrocesso:

- a) as sub-metas devem ser satisfeitas (sucedidas) em ordem, do topo à base do programa,
- b) as cláusulas dos predicados são testadas na ordem que eles aparecem no programa, isto é, do topo a base,
- c) quando uma sub-meta equipara a cabeça de uma regra, o corpo daquela regra deve ser sucedido em seguida. Então, o corpo da regra constitui um novo conjunto de sub-metas para serem atingidas,
- d) uma meta é atingida quando é encontrada uma equiparação do fato para cada extremidade (ramificação) da árvore meta.

- CONTROLE DA BUSCA PARA A SOLUÇÃO DO PROBLEMA

O mecanismo de retrocesso de Prolog pode resultar em buscas desnecessárias, e por causa disto, pode surgir ineficiências. Neste caso, Turbo-Prolog fornece duas ferramentas que permitem controlar o mecanismo de retrocesso: o predicado **fail**, que é utilizado para forçar o retrocesso, e o predicado **cut** (indicado por !), que é empregado para evitar o retrocesso.

Quando Turbo Prolog começa retroceder é porque a chamada de um predicado falhou. Em certas situações, é necessário forçar um encadeamento para trás a fim de encontrar alternativas para a solução do problema. Então, utiliza-se o predicado *fail*, o qual sempre falha, isto é, nunca está satisfeito.

O predicado *cut* (!) é inserido num programa da mesma forma como é colocada uma sub-meta no corpo de uma regra. Uma vez que um *cut* é alcançado, não é possível retroceder para as sub-metas colocadas anteriormente na cláusula que está sendo processada.

- LISTAS

Em Prolog, uma lista é um objeto que contém um número arbitrário de outros objetos dentro dela. Ela corresponde a uma dimensão (array) em linguagens de procedimentos, mas, ao contrário da dimensão, uma lista não exige que o seu tamanho seja declarado antes de usá-la no programa.

Uma lista consiste de duas partes: a *cabeça* e a *cauda*. A cabeça é o primeiro elemento e a cauda compreende todos os elementos subsequentes da lista. Portanto, a cauda de uma lista é sempre uma lista, enquanto a cabeça de uma lista é um elemento. Uma lista vazia não pode ser dividida em cabeça e cauda.

A lista contendo os números 1780, 3020 e 1190, referentes as posições de alguns picos de um determinado composto, é escrita como [1780,3020,1190], onde cada item contido na lista é conhecida como um elemento. Neste caso, a cabeça da lista é o número 1780, os demais correspondem à cauda. Uma maneira de separar a cabeça da cauda da lista é usar um (). Para formar uma estrutura de dados no formato de lista, os elementos são separados com vírgulas e inseridos entre colchetes. Exemplos:

[álcool,éter,cetona],

["vibração_de_estiramento_O-H", "vibração_de_deformação_C-H"],

[0.15,0.35,0.42,0.38,0.05,0.52]

Os elementos nessas listas são símbolos, strings e números reais, respectivamente. Para declarar o domínio numa lista de números inteiros, deve-se incluí-los em sua seção da seguinte maneira: `integer_list = integer*`,

onde o sinal de asterisco (*) indica "lista de", isto é, `integer*` significa lista de inteiros. É importante destacar que os elementos numa lista podem ser de quaisquer tipos; devem, porém, pertencer ao mesmo domínio.

- COMENTÁRIOS

Constitui-se um bom estilo de programação a inclusão de comentários no programa, para explicar detalhes que nem sempre são evidentes ao usuário. Isto torna o programa mais fácil para se entender. Contudo, a escolha apropriada dos nomes das variáveis, predicados e domínios, exigirá poucos comentários.

Para os comentários em linhas múltiplas, deve-se iniciá-los com os caracteres `/*` (barra, asterisco) e finalizá-los com os caracteres `*/` (asterisco, barra). Quando o comentário é colocado em uma única linha, pode-se utilizar os mesmos caracteres ou começar com um sinal de porcentagem (%). Alguns exemplos de comentários estão abaixo descritos:

```
/* Isto é um exemplo de comentário que */
```

```
/* utiliza mais de uma linha */
```

```
/* Este exemplo está precedido e sucedido pelos caracteres de comentários (/* e */). */
```

```
% Este exemplo utiliza somente o sinal de porcentagem, por estar em uma linha.
```

- EXEMPLO DE UM PEQUENO PROGRAMA ESPECIALISTA

A construção de um programa especialista para identificar a presença das vibrações que caracterizam o éster em um composto desconhecido, onde há uma base de dados contendo as informações necessárias para uma equiparação, com os fatos obtidos do espectro de I.V. sob análise. Algumas seções descritas anteriormente são mostradas abaixo:

domains

```
lista_de_vibracao = string*           % a (*) significa lista de caracteres
```

predicates

```
identificar_grupo_funcional
grupo_funcional(string)
tem_vibracao(lista_de_vibrações)
```

database - dba_links

```
tipo_vibracao(string,string_list)
```

meta

```
identificar_grupo_funcional
```

clauses

```
identificar_grupo_funcional :-           % cabeça da regra com um predicado sem argumento
```

```
% a regra acima tem 1 sub-meta ou corpo da regra, onde os argumentos são as variáveis
```

```
tipo_vibracao(Grupo,Lista_de_Vibracao),           % informações da base de dados
confirmar_vibracao(Grupo,Lista_de_Vibracao,VibracoesCaracteristicas),
write(VibracoesCaracteristicas), nl,
fail, !.                                           % write, nl, fail e ! são predicados padrões
```

```
identificar_grupo_funcional.                       % condição de parada.
```

```
% sub-meta da cláusula anterior, que agora é outra cláusula recursiva (até lista ficar vazia)
```

```
confirmar_vibracao("éster",[PrimeiraVibracao|Lista_de_Vibracoes],
    [PrimeiraVibracao|ListaVibracoesCaracteristicas]) :-
confirmar_vibracao("éster",Lista_de_Vibracoes,ListaVibracoesCaracteristicas), !.
```

```
confirmar_vibracao(, [], []) :- !.               % condição de parada
```

CAPÍTULO III

PROGRAMA GERENCIADOR DE BIBLIOTECA DE ESPECTROS

Os usuários da espectroscopia no I.V. dispõem atualmente de diversos programas comerciais que incorporam bibliotecas de espectros. Porém, muitas vezes este se encontra na condição de ser obrigado a adquirir uma vasta biblioteca de espectros, da qual ele utiliza apenas uma pequena parte do seu extenso banco de dados. Além disso, o usuário não poderá, frequentemente, acrescentar nesta biblioteca, os espectros obtidos experimentalmente, durante o desenvolvimento de sua pesquisa.

Basicamente, um pesquisador necessita apenas de um número relativamente pequeno de espectros, pertencentes a um certo grupo de compostos de características definidas, que apresentam interesse especial, por estarem diretamente ligados ao seu tópico de pesquisa. O acesso a estes espectros deve ser fácil e de preferência, feito através dos nomes dos compostos. Alguns recursos são necessários para que um espectro de uma substância desconhecida possa ser visualmente comparado com espectros existentes na biblioteca.

Diante desta situação pode-se justificar a construção de uma mesa digitalizadora de espectros, de baixo custo, capaz de transcrever o espectro analógico obtido num instrumento convencional para um banco de dados de um microcomputador. Esta mesa digitalizadora pode ser empregada na construção de bibliotecas de espectros de uso específico do pesquisador, o que demanda menos memória ou espaço de armazenamento para a sua utilização. Neste caso, os espectros de interesse existentes na literatura podem ser digitalizados e armazenados para uma posterior busca e comparação. Além disso, o mesmo instrumento pode ser utilizado na aquisição não somente do espectro, mas, sob controle de um programa adequado, extrair dos dados espectrais, os fatos necessários para

uma interpretação através de um S.E. (por exemplo, para a identificação de grupos funcionais presentes no composto, responsáveis pelo espectro obtido).

Com base nestes aspectos e em manifestações de diversos pesquisadores que se utilizam da espectroscopia no I.V. foi desenvolvido, um programa denominado "Programa Gerenciador de Biblioteca de Espectros de Infravermelho (PROGEIV)". Este programa pode ser utilizado na construção, elaboração e manutenção de uma biblioteca de espectros, através do uso de uma mesa digitalizadora desenvolvida neste laboratório[100], ou de mesas digitalizadoras comerciais[101], ou mesmo diretamente de espectrofotômetros I.V., logo após a obtenção do espectro. Também, este programa inclui uma interface entre os dados analógicos contidos no espectro e os fatos necessários para a atuação do S.E. desenvolvido.

III.1 - PARTE EXPERIMENTAL

O programa PROGEIV foi escrito em linguagem Quick-Basic, versão-4.5 da Microsoft[102] e executado em um microcomputador modelo IBM-PC-AT/386 dotado de coprocessador matemático, utilizando um monitor policromático de resolução VGA. Este microcomputador possui capacidade de memória volátil (RAM) de 4 Mbytes, uma unidade de disco rígido (Winchester) de 80 MBytes, e os dispositivos de armazenamento permanente de dados, que são constituídos por uma unidade de disco flexível de 3.½" e um de 5¼". Uma impressora matricial tipo EPSON, modelo LX-810 de 80 colunas foi utilizada para a obtenção de cópias em papel do conteúdo gráfico da tela do monitor do microcomputador. O PROGEIV, em sua forma executável, ocupa 73 Kbytes de memória.

Os espectros de I.V. de compostos orgânicos foram obtidos da biblioteca SADTLER Research Laboratories Inc.[103] (na forma gráfica em papel), que faz parte do

acervo da biblioteca do IQ/Unicamp. Os espectros oriundos do espectrofotômetro Perkin Elmer, modelo 1600, na forma gráfica em papel ou gravados em discos flexíveis, gerados pelo microcomputador acoplado ao instrumento, foram cedidos por alguns colegas deste instituto e utilizados principalmente para testar a interface do PROGEIV/PERKIN.

A ferramenta para a transcrição de dados espectrais constitui-se de uma mesa digitalizadora, construída neste laboratório, referida como HM (descrita no apêndice A) e/ou uma tipo comercial, como a modelo Digigraf DaVinci, existente num laboratório deste instituto. Com a mesa HM, foram digitalizados 86 espectros de compostos orgânicos, quantidade suficiente para testar o funcionamento das partes mecânica e eletrônica deste instrumento. Através da mesa DaVinci foi possível digitalizar outros 396 espectros de compostos orgânicos. Essa mesa digitalizadora foi utilizada pelo PROGEIV, confirmando a eficácia das sub-rotinas de aquisição de dados.

III.2 - DESCRIÇÃO DO PROGRAMA PROGEIV

O programa PROGEIV apresenta as seguintes funções: 1) aquisição de dados espectrais, 2) construção, manipulação e manutenção da biblioteca de espectros, e 3) interface com o S.E.. O programa PROGEIV permite a introdução ou a remoção de espectros no banco de dados e o acesso do usuário às suas diferentes sub-rotinas é realizada através de menus. Estes menus foram elaborados no programa com opções claras e objetivas permitindo ao usuário a seleção inequívoca da opção desejada e o direcionamento do fluxo do programa. Na figura III-1 está ilustrado um diagrama de blocos do programa PROGEIV.

Inicialmente, quando coloca-se o programa PROGEIV em execução exibe-se na tela do monitor uma relação de tipos de monitores com uma pergunta ao usuário. Assim, pode-se selecionar qual o tipo de monitor que está acoplado no microcomputador. Isso procede porque a parte gráfica do programa está adaptada para diferentes tipos de monitores, como: monocromático ou policromático de resoluções CGA, EGA e VGA. No caso de dúvidas, o usuário pode pressionar a tecla "F1" (HELP), implementada para ajudá-lo quanto à escolha correta do tipo de monitor, condição esta apresentada na figura III-2. Após essa escolha, as mensagens na tela do monitor são retiradas, e imediatamente aparece o menu principal do PROGEIV, como mostra a figura III-3. Este menu oferece ao usuário a possibilidade de selecionar várias opções quanto ao fluxo do programa. Também, pode-se pressionar a tecla "F1", que informa como deve selecionar corretamente a opção desejada. As opções do menu principal são numeradas de 1 a 0 e estão distribuídas da seguinte maneira: a) as opções 1, 2 e 3 direcionam para aquisição de dados espectrais, b) as opções 4, 5 e 6 permitem manipulações dos espectros da biblioteca, c) a opção 7 é uma interface com o S.E. desenvolvido, d) a opção 8 torna corrente um espectro da biblioteca, e) a opção 9 remove espectros não mais desejados na biblioteca, e f) a opção 0 é para modificar as escalas dos espectros a serem digitalizados ou ampliá-las a partir de 2000 cm^{-1} .

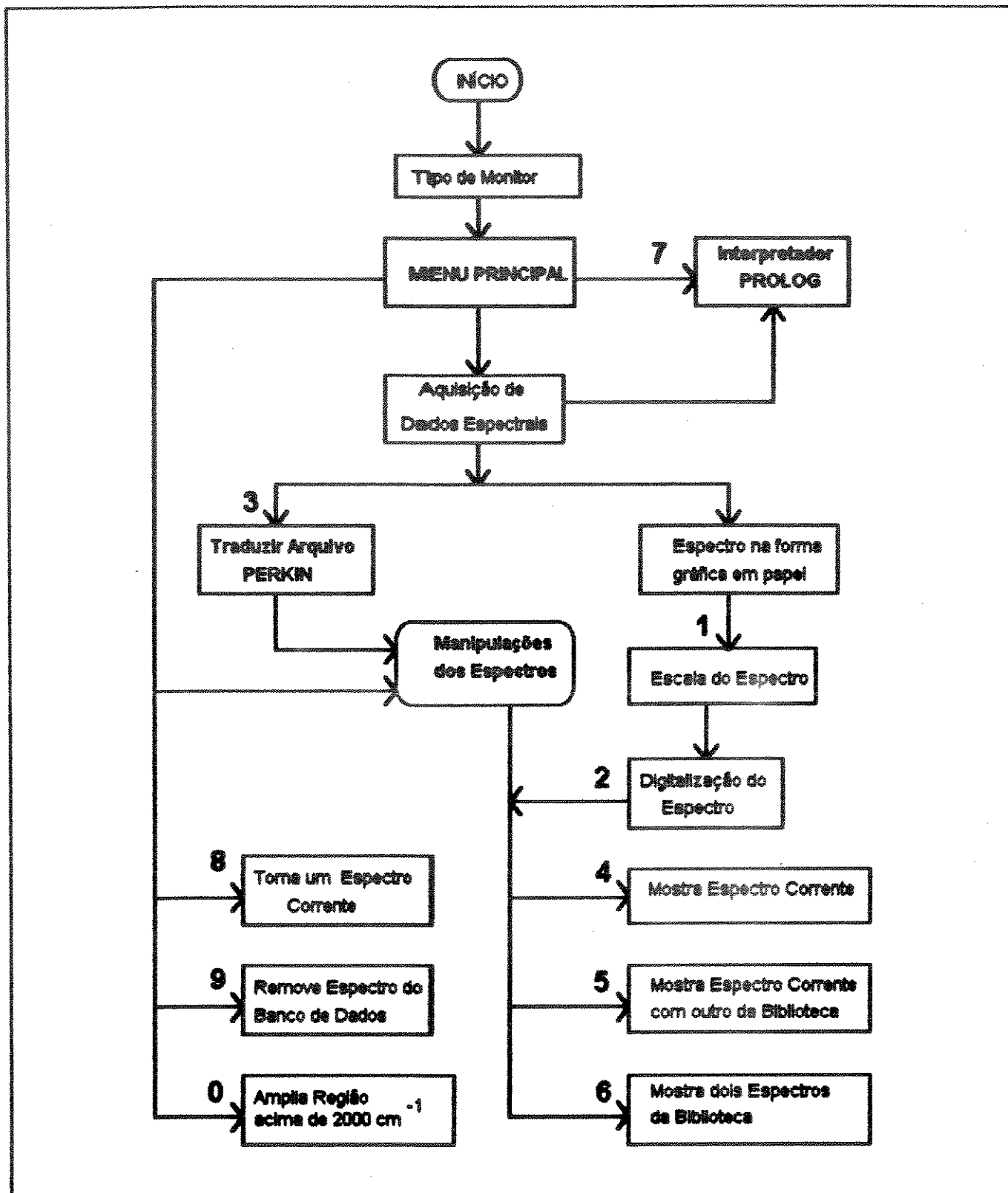


FIGURA III-1 - Diagrama de blocos do programa PROGEIV. Os números representam as opções do menu principal.

<< TELA DE SELEÇÃO DE VÍDEO PARA O PROGEIV >>

(Sistema de Manipulação de Espectros de Infravermelho)

Como o sistema pode funcionar para vídeos de tipos mais comuns:

- 1. CGA
- 2. EGA
- 3. VGA

Selecione o número correspondente ao seu tipo de monitor

Deste modo, será utilizado o máximo de capacidade de resolução existente no vídeo.

Pressione qualquer tecla para sair do Help

FIGURA III-2 - Saída de vídeo para ajudar o usuário a selecionar o tipo de monitor adequado.

<< MENU PRINCIPAL >>

1. ATRIBUE ESCALA PARA O ESPECTRO DE I.V.
 2. INICIA DIGITALIZAÇÃO
 3. TRADUZ ARQUIVOS PERKIN ELMER
 4. MOSTRA O ESPECTRO CORRENTE DA BIBLIOTECA
 5. COMPARA O ESPECTRO CORRENTE COM UM OUTRO DA BIBLIOTECA
 6. COMPARA DOIS ESPECTROS DA BIBLIOTECA
 7. INTERPRETA ESPECTRO CORRENTE DA BIBLIOTECA (EXPERTIV)
 8. TORNA CORRENTE UM ESPECTRO DA BIBLIOTECA
 9. REMOVE ESPECTRO DA BIBLIOTECA
 0. AMPLIA A CONFIGURAÇÃO DO ESPECTRO
- <ESC>. FIM DO PROGRAMA PRINCIPAL <F1>. HELP

FIGURA III-3 - Menu principal do programa PROGEIV.

O controle do programa quanto à aquisição dos dados espectrais e às manipulações dos espectros da biblioteca estão descritos abaixo.

III.3 - AQUISIÇÃO DE DADOS ESPECTRAIS

A aquisição de dados espectrais pode ser feita de duas maneiras: 1) através da mesa digitalizadora de espectros como a HM (descrita no apêndice A), ou uma tipo comercial como a DaVinci. Esse instrumento é capaz de digitalizar os espectros existentes na literatura ou registrados em papel pelo espectrofotômetro, e 2) diretamente de um espectrofotômetro de I.V., por tradução de arquivos de dados gerados em discos flexíveis pelo microcomputador do espectrofotômetro.

III.3.1 - AQUISIÇÃO DE DADOS COM A MESA DIGITALIZADORA

Inicialmente, seleciona-se no menu principal a opção número 1 (atributos da escala do espectro a ser digitalizado), para entrar com os valores máximos e mínimos das transmitâncias e números de ondas do espectro a ser digitalizado. A figura III-4 mostra o diálogo originado dessa opção, para um espectro da biblioteca SADTLER.

Também, o programa oferece condições para que sejam digitalizados espectros com escalas horizontais descontínuas. Na maior parte dos casos, os espectros de I.V. tem sua escala horizontal dividida em duas partes, ou seja, de 4000 a 2000 e de 2000 a 600 ou 400 cm^{-1} . Embora mais raros, espectros com duas descontinuidades, também podem ser digitalizados. Através do emprego das sub-rotinas "SCALE1", "SCALE2", "SCALE3", "CONVERTHOMEMADE" e "CONVERTCOMMERCIAL", que consideram estas descontinuidades, convertendo corretamente os valores de tensão em porcentagem de transmitância e número de onda, expressos em porcentagem e cm^{-1} , respectivamente.

```

<< ATRIBUTOS DE ESCALA PARA O ESPECTRO DE I.V. >>

QUANTAS ESCALAS TEM O ESPECTRO PARA DIGITALIZAÇÃO ?      2

<< NÚMEROS DE ONDAS >>

ENTRE COM O VALOR MÁXIMO E PRESSIONE ENTER: ?            4000
ENTRE COM O VALOR INTERMEDIÁRIO E PRESSIONE ENTER: ?      2000
ENTRE COM O VALOR MÍNIMO E PRESSIONE ENTER: ?             400

<< TRANSMITÂNCIAS >>

ENTRE COM O VALOR MÁXIMO E PRESSIONE ENTER: ?            100
ENTRE COM O VALOR MÍNIMO E PRESSIONE ENTER: ?             0

PRESSIONE < ESC > PARA RETORNAR AO MENU PRINCIPAL

```

FIGURA III-4 - Diálogo originado da opção número 1 do menu principal (Atributos da Escala para o Espectro de I.V.), para um espectro da biblioteca SADTLER, com duas escalas (de 4000 a 2000 e de 2000 a 400 cm^{-1}).

Depois, para iniciar a digitalização do espectro deve-se selecionar a opção número 2 do menu principal, que indaga o usuário com relação ao tipo de mesa digitalizadora a ser utilizada (HM ou tipo comercial).

Antes de iniciar a digitalização, coloca-se o cursor da mesa digitalizadora na posições de máximos e mínimos, de acordo com a escala inserida na opção número 1. Este procedimento permite que o microcomputador defina os coeficientes, que serão utilizados na conversão dos valores lidos em unidades de tensão (mV), para as coordenadas usuais

dos espectros de I.V. (transmitância e número de onda). A figura III-5 mostra o diálogo originado dessa opção até essa etapa do programa.

Quando inicia-se a digitalização propriamente dita, as mensagens anteriores são retiradas da tela do monitor e, em seguida, aparecem os valores decimais dos sinais (em mV, por exemplo caso a mesa HM esteja sendo empregada) que estão sendo coletados, quando da movimentação do cursor da mesa digitalizadora. Além disso, são mostradas abaixo mensagens relacionadas com as teclas de sinais (+) e (-), que alteram a diferença de tensão mínima no sentido vertical (sinal analógico), que é utilizada entre a aquisição dos pontos para que não se perca informações do espectro principalmente, quando se está aproximando da intensidade máxima de uma banda de absorção. Essas diferenças foram estipuladas em 12 ou 6 mV para a coordenada vertical e 6 ou 4 mV para a coordenada horizontal, com a mesa HM, respectivamente.

```


<< MESA DIGITALIZADORA >>



1. HM



2. Tipo Comercial (DaVinci) - RS-232



Entre com (1 ou 2) para escolher o tipo de mesa digitalizadora ?



<< DADOS PARA A DIGITALIZAÇÃO >>



NOME DO COMPOSTO: ? ISOBUTRALDEIDO



MOVA O CURSOR PARA O PONTO DA BASE ESQUERDA E PRESSIONE ENTER



MOVA O CURSOR PARA O PONTO DA BASE MÉDIA E PRESSIONE ENTER



MOVA O CURSOR PARA O TOPO DA BASE DIREITA E PRESSIONE ENTER



POSICIONE O CURSOR PARA INICIAR A DIGITALIZAÇÃO E PRESSIONE ENTER


```

FIGURA III-5 - Parte do diálogo originado pela opção número 2 do menu principal (Inicia a Digitalização).

Os valores de potencial relacionados as coordenadas horizontais e verticais do cursor da mesa digitalizadora (expressos em milivolts) são coletados, através da sub-rotina "ALTREAD", que seleciona o canal de leitura e fornece sempre em primeiro lugar, o valor da posição vertical, porque esta, em um espectro, está sujeita a variações mais bruscas por acompanhar as bandas de absorção. Os valores retornados são médias de 10 leituras do conversor A/D dinâmico. A interação com o usuário é facilitada pelo programa, que através das etapas necessárias, direciona-o com mensagens objetivas. As sub-rotinas encarregadas para a aquisição de dados espectrais, com o uso da digitalização seguida de arquivamento, podem ser vistas no diagrama de blocos da figura III-6.

Um contador automático, que mantém atualizado o número de espectros existentes no sub-diretório "sistspec\system" é salvo no arquivo de controle denominado "spectr.ctr". Os espectros coletados são armazenados através de nomes DOS, da seguinte forma: "spexxx.spt", onde xxx é um número entre 000 e 999, sendo, então, igual a 1000 o número máximo de espectros no mesmo diretório. Os arquivos dos espectros contêm o nome do composto e o número total de pares de dados amostrados (transmitância e número de onda). Também, para o último espectro digitalizado, é feita e arquivada uma cópia sob o nome DOS "spec.spt", que então passa a ser o espectro corrente da biblioteca. Como exemplo, uma parte desse arquivo de um composto orgânico é mostrada na figura III-7. Todos os arquivos dos espectros digitalizados são armazenados no sub-diretório "sistspec\spectrum".

Um outro arquivo sob o nome DOS "setspec.se" colocado no sub-diretório "sistspec\system" relaciona o número do espectro e o nome do composto referente ao espectro contido na biblioteca, por exemplo "spectrum\SPE006.SPT", "2-BROMO-3-METHYLBUTYRIC ACID (C₅H₉BrO₂)".

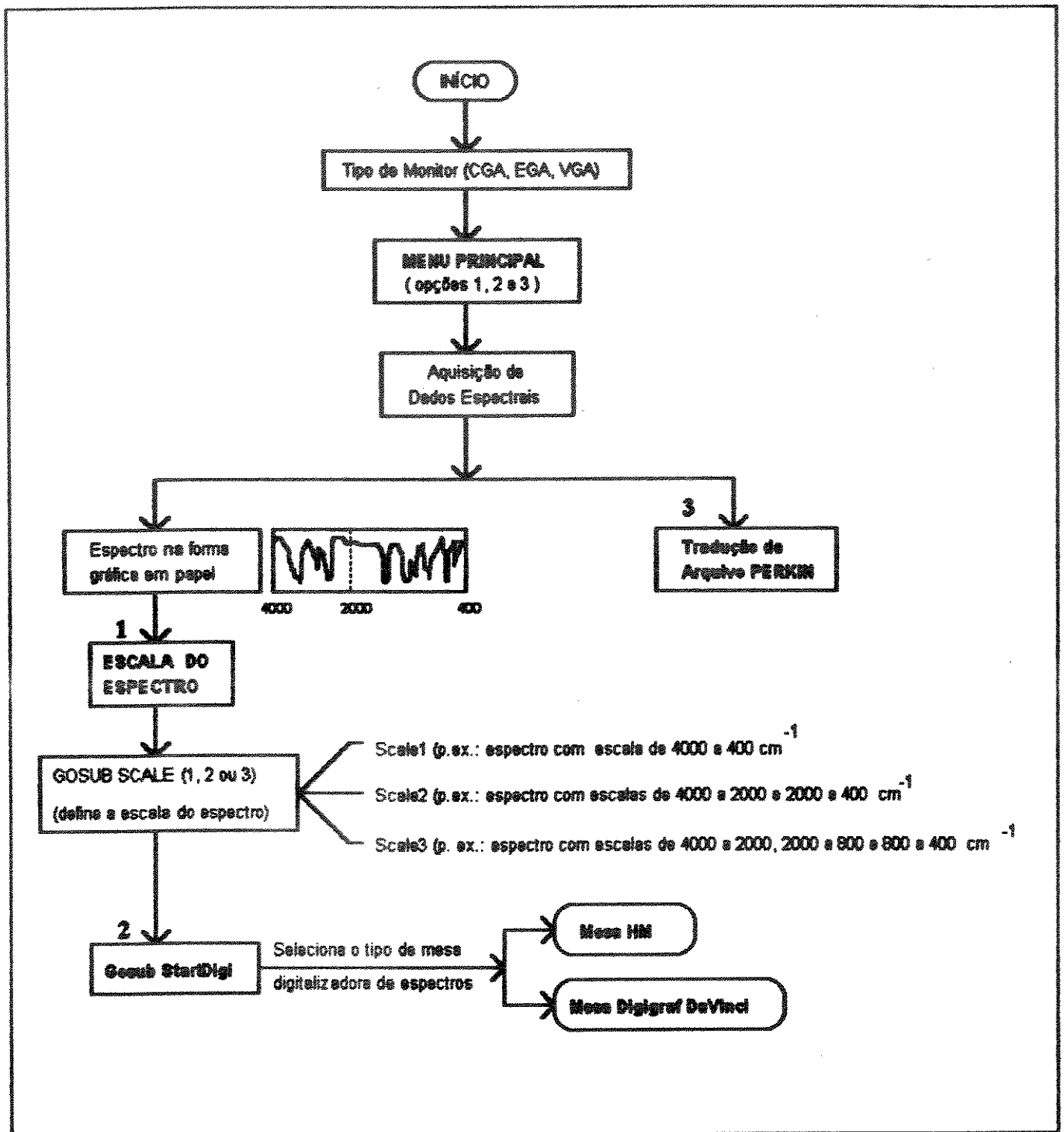


FIGURA III-6 - Diagrama de blocos das sub-rotinas encarregadas para aquisição de dados espectrais. Os números representam as opções do menu principal.

"3-BROMOPROPIONIC (Br-CH₂-CH₂-COOH)"

75,4000

75,3991

75,3978

75,3965

75,3950

75,3935

75,3921

75,3908

75,3893

76,3878

76,3865

76,3852

76,3838

76,3823

76,3810

76,3798

76,3783

76,3769

76,3754

76,3739

76,3724

76,3711

76,3696

76,3680

76,3667

76,3653

76,3639

76,3626

75,3611

75,3599

75,3584

75,3569

75,3555

75,3539

75,3524

74,3511

74,3497

74,3485

.....

.....

94,408

94,401

FIGURA III-7 - Parte do arquivo "spe001.SPT" referente ao espectro de I.V. do ácido 3-bromopropiônico, que é por exemplo, o espectro corrente denominado "spec.spt" existente na biblioteca. A primeira linha indica o nome do composto e as demais, os pares de pontos coletados (transmitância e número de onda).

Todos os arquivos dos espectros foram salvos pelo Quick-Basic (codificados em ASCII) e verificou-se que os mesmos podem ser acessados sem dificuldade, no formato de lista, através do S.E. desenvolvido e escrito em linguagem Turbo Prolog. Este fato é importante para a implementação de um programa especialista, que empregue essa linguagem.

Ao final da digitalização, isto é, quando o usuário ultrapassa os valores limites das coordenadas, inicialmente fornecidos (veja figura III-5) ocorre um sinal sonoro na forma de "BEEP" e imediatamente o espectro recém digitalizado é apresentado na tela do monitor. Abaixo, na tela do monitor, é apresentada uma mensagem perguntando se o usuário deseja ou não salvar esse espectro. Após uma confirmação positiva, o espectro será armazenado como `spexxx.spt`, e o nome do composto é arquivado em `"setspec.se"`

III.3.2 - AQUISIÇÃO DE DADOS COM UM ESPECTROFOTÔMETRO DE I.V.

A sub-rotina "PERKIN" foi implementada no PROGEIV para coletar os dados espectrais por meio da tradução de arquivos de dados gerados e existentes em discos flexíveis, através do microcomputador do espectrofotômetro.

Os sub-diretórios e arquivos DOS dessa sub-rotina, para controle das informações espectrais, estão assim distribuídos:

- a) `"sistspec\system"` - `"setperk.se"` arquivo de entrada, que relaciona os nomes dos espectros dos compostos orgânicos, quando transferidos para a biblioteca, por exemplo `"perkin\CETONA2.SP"`,
- b) `"sistspec\perkin"` - `"*.sp"` arquivo do espectro com um nome DOS qualquer e extensão `".SP"`, antes de ser traduzido (está na forma binária), por exemplo `"perkin\AMIDA4.SP"`,

c) "sistspec\realspec" - "*.spt" arquivo do espectro com um nome DOS qualquer e extensão ".spt", o qual passou para a forma ASCII decimal, logo após a sua expansão, por exemplo "AMIDA4.SPT",

d) "sistspec" - "perkin1.nam", arquivo de entrada, quando é feita a cópia no drive selecionado ("A", "B" ou "C"), dos arquivos dos espectros oriundos do espectrofotômetro.

As informações contidas num espectro de I.V., quando geradas em discos flexíveis, pelo microcomputador do espectrofotômetro estão codificadas na forma binária. Além disso, essa codificação é restrita ao instrumento, isto é, o químico apenas utilizará novamente estes dados, caso precisar do registro do espectro na forma gráfica em papel. Também, a codificação na forma binária dificulta o entendimento sobre os dados implicando, assim, num bloqueio para processar estes dados com outros programas. Portanto, o objetivo da implementação dessa sub-rotina no PROGEIV restringe-se principalmente, ao fato da sua capacidade em traduzir os dados oriundos do instrumento, da forma binária para a forma decimal, quando desejados. Conseqüentemente, os dados podem ser processados pelo usuário em qualquer programa. Um fluxograma completo da sub-rotina "PERKIN", pode ser observado na figura III-8, e as suas etapas durante a execução estão apresentadas abaixo.

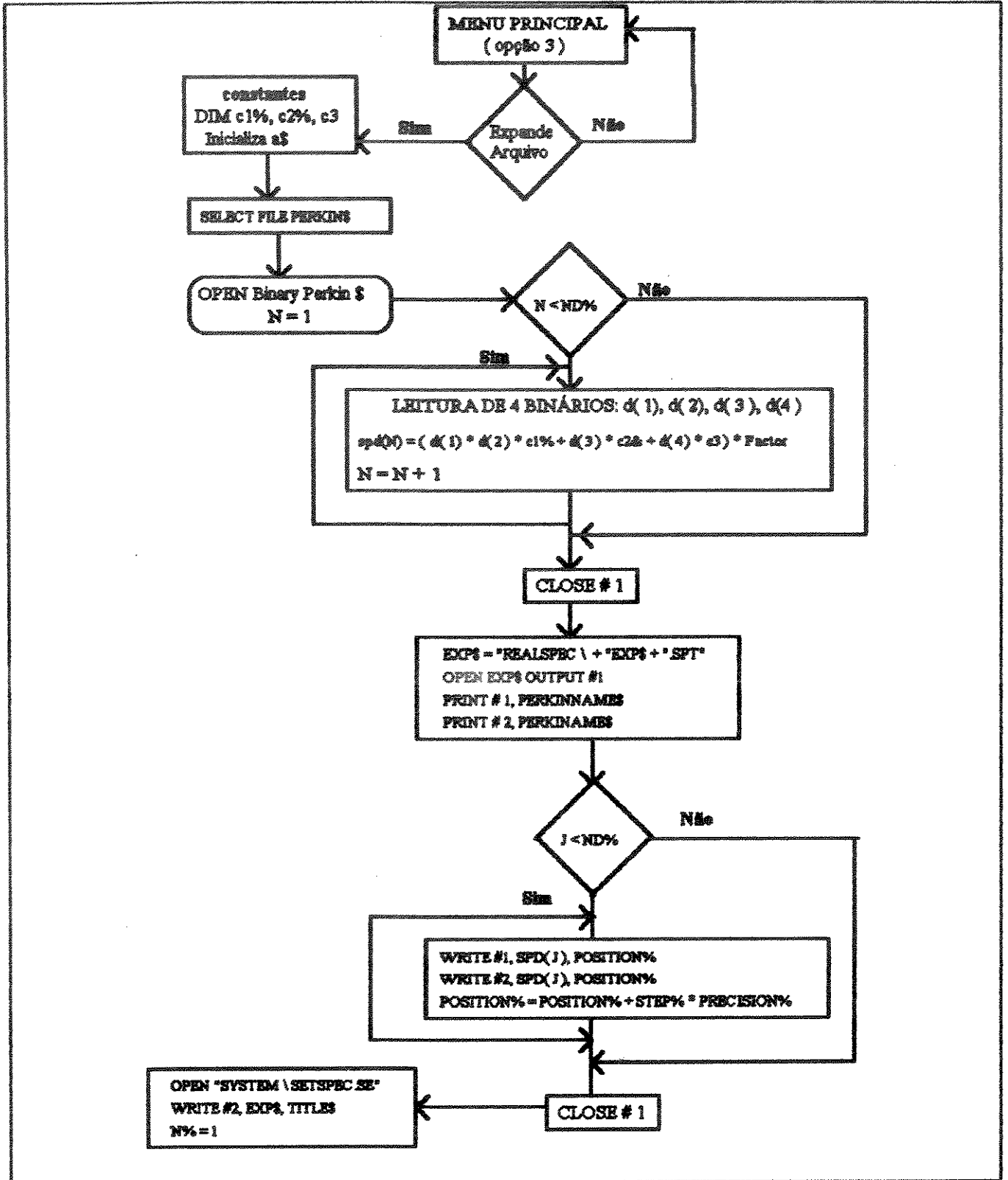


FIGURA III-8 - Fluxograma da sub-rotina "PERKIN".

Através do menu principal, seleciona-se a opção número 3 (Tradutor de Arquivos Perkin Elmer). Inicialmente, aparece na tela do monitor uma pergunta com relação a escolha do "drive" ("A", "B" ou "C"), onde será efetuada a leitura dos dados existente no arquivo acima exemplificado. Após a cópia do arquivo no "drive" selecionado, é criado o arquivo "perkinel.nam", colocado no diretório "sistspec", que contém o nome do composto do espectro obtido desse "drive".

Numa segunda etapa, após a seleção do drive, a última pergunta é retirada da tela do monitor, e retorna uma outra, sobre a expansão ou não, de um arquivo de um disco flexível. Em caso positivo, deve-se selecionar o arquivo "AMOSTRA1.SP", de modo que as informações codificadas na forma binária, sejam expandidas para a forma ASCII. Em seguida, o arquivo "AMOSTRA1.SP" é armazenado no sub-diretório "sistspec\perkin", e seu arquivo é expandido no sub-diretório "sistspec\realspec", com uma extensão ".spt", nesse caso "AMOSTRA1.SPT". Com esse procedimento, o usuário pode obter, quantas vezes desejar, um espectro na forma gráfica em papel pelo espectrofotômetro, chamando o arquivo "AMOSTRA1.SP", ou através do programa PROGEIV, conforme a implementação da sub-rotina "PRINTSPEC", que transfere os dados do arquivo "AMOSTRA1.SPT", apresentados na tela do monitor, para uma impressora. A tradução do último espectro também será salva no sub-diretório "sistspec\spectrum" pelo nome de "spec.spt", tornando-se um espectro corrente da biblioteca. A partir dessa etapa, o espectro pertence à biblioteca, e pode ser manipulado através das opções existentes no programa PROGEIV (mostrar o espectro na tela do monitor, comparar, remover, imprimir, etc.). Além disso, após a execução da sub-rotina "PERKIN", os números dos espectros e os nomes dos compostos desses espectros que foram expandidos, são arquivados no "setperk.se". Esse arquivo é colocado no diretório "sistspec\system".

A tradução de um arquivo de dados gerado pelo espectrofotômetro de I.V. Perkin Elmer, modelo 1600 e posteriormente arquivado em um disco flexível, através do seu microcomputador, está abaixo descrita.

A primeira parte dessa sub-rotina é para a identificação do arquivo contendo os dados a serem expandidos. Este arquivo denominado de PerkFile\$(N%).SP é fornecido pelo usuário, através de uma seleção no arquivo "perkinel.nam". O nome do arquivo do espectro não precisa ser entrado com a terminação ".sp" (a terminação padrão do Perkin Elmer), porque ela é automaticamente adicionada pela sub-rotina.

O arquivo é aberto para a entrada e o comprimento do arquivo (length%, em bytes) é retornado pela função Basic LOF(N%), onde N% é o número do arquivo aberto, como por exemplo PerkFile\$(4). A maior parte das linhas presentes no cabeçário do arquivo Perkin Elmer é omitida, até a identificação da "string" "#GR". Essa "string" marca o início do Cabeçário Gráfico no arquivo Perkin Elmer. As próximas dez "strings" contêm informação relevante para o processo de expansão e são entradas pela sub-rotina através da variável "string" indexada a\$(N%) e o arquivo é fechado. Esta etapa está implementada da seguinte maneira:

```

OPEN PERKIN$ FOR INPUT AS #1
length% = LOF(1)
WHILE S$ <> "#GR"
  INPUT #1, S$
  IF N% = 9 THEN
    PerkinName$ = S$
    N% = N% + 1
  ELSE
    N% = N% + 1
  END IF
WEND
N% = 1
FOR N% = 1 TO 10
  INPUT #1, a$(N%)
NEXT
CLOSE #1

```

Em seguida, a sub-rotina verifica a técnica de compressão empregada. A string a\$(8) deve ser igual a "8", ao contrário, a sub-rotina não pode ser usada. Se a técnica de compressão está correta, o número de pontos presentes no arquivo (nd%) é lido da string

a\$(7). O fator que converte os números inteiros para a escala das ordenadas é lido da "string" a\$(3). Então, a sub-rotina calcula a posição do primeiro byte de dados no arquivo usando os valores de length% e o nd%, como mostra abaixo:

```
LOCATE 17, 20: PRINT "WAIT FOR FILE EXPANDING"  
precision% = 3 ' precision por enquanto deve ser constante
```

```
IF Title$ = "" THEN Title$ = PerkinName$  
IF a$(8) <> "8" THEN GOTO ErrTrat
```

```
POSITION% = VAL(a$(5))  
STEPS% = VAL(a$(6))  
nd% = VAL(a$(7))  
Factor = VAL(a$(3))  
firstbyte% = length% - 4 * nd% + 1
```

Agora, conhecendo a posição do primeiro byte de dados no arquivo, ela pode ser aberta como binária. A declaração Basic "SEEK" é usada para mover o apontador de entrada para a posição do primeiro byte (firstbyte%). A entrada sequencial de grupos de quatro bytes é executada e os dados expandidos (expressos em unidades da ordenada) são gerados e armazenados na variável indexada spd(). O arquivo é fechado e um novo arquivo de dados expandido é criado sob o nome de "spec.spt". O arquivo contém o cabeçário padrão Perkin Elmer, seguido pelo dado da ordenada na forma ASCII. Se uma escala de transmitância (a mais comum) é usada, deve ser útil limitar os dados em 4 dígitos, a fim de salvar espaço no disco flexível. Esta etapa está implementada do seguinte modo:

```
OPEN PERKINS$ FOR BINARY AS #1
SEEK #1, firstbyte%
FOR N% = 1 TO nd%
    FOR K% = 1 TO 4
        d$ = INPUT$(1, #1)
        d(K%) = ASC(d$)
    NEXT
    spd%(N%) = (d(1) + d(2) * c1% + d(3) * c2% + d(4) * c3) * Factor
NEXT
CLOSE #1
```

A figura III-9 ilustra um espectro de I.V. do hexanaldeído, obtido da biblioteca do espectrofotômetro Perkin Elmer, modelo 1600.

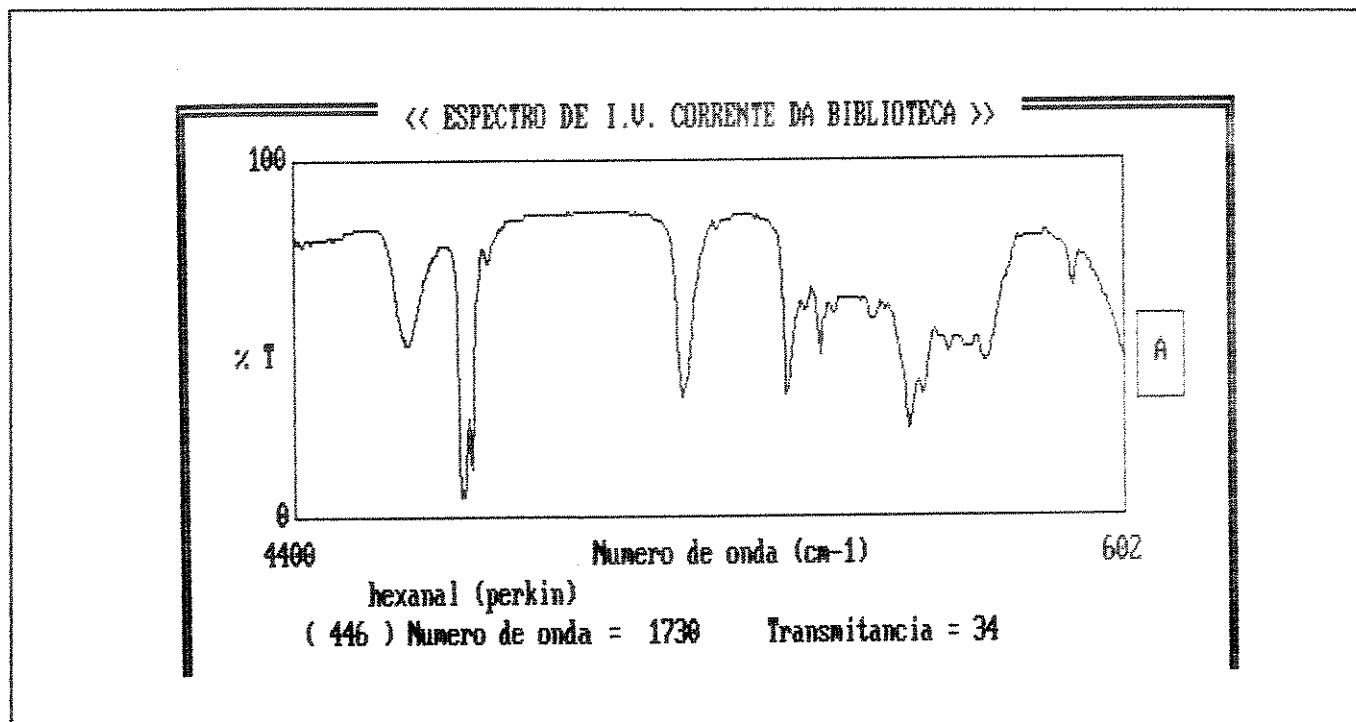


FIGURA III-9 - Espectro de I.V. do hexanaldeído obtido da biblioteca do espectrofotômetro Perkin Elmer, modelo 1600.

III.4 - MANIPULAÇÕES DOS ESPECTROS DA BIBLIOTECA

No menu principal do programa PROGEIV (veja figura III-3) pode-se selecionar as opções números 4, 5, 6 e 9, citadas anteriormente, que permitem o direcionamento das manipulações dos espectros da biblioteca, de acordo com as necessidades do usuário. Essas manipulações sempre são acompanhadas de mensagens claras e objetivas na tela do monitor, que direcionam o usuário para selecionar, imprimir, percorrer o cursor gráfico no espectro, retornar ao menu principal, etc.. As principais sub-rotinas para as manipulações dos espectros na biblioteca são: a) "ShowCurrentSpectrum", que se encarrega de apresentar na tela do monitor o espectro corrente, b) "CompareCurrentSpectrum" implementada para comparar o espectro corrente com um outro espectro selecionado na biblioteca, c) "ShowTwoSpectra" que mostra dois espectros escolhidos aleatoriamente para comparação, e d) "HouseKeeping" utilizada para remover um espectro não mais desejado na biblioteca. Um diagrama de blocos mostrando essas sub-rotinas de manipulações dos espectros, está apresentado na figura III-10.

As quatro principais sub-rotinas de manipulações acima mencionadas, de espectros da biblioteca, oferecem um recurso para melhorar a observação visual do espectro, através de um cursor gráfico (teclas de setas para a direita e/ou para a esquerda). Quando uma dessas teclas é pressionada continuamente, o cursor percorre um ponto na linha gráfica do espectro ("pixel"), e mais abaixo do gráfico aparecem o nome do composto, os números dos pontos amostrados, e os valores dos números de ondas e transmitâncias correspondentes. Pode-se também pressionar as teclas de sinais (+) ou (-) para o cursor gráfico avançar de uma só vez ou não 30 pontos amostrados. Por exemplo, se o ponto está no número 10 e pressiona-se a tecla de sinal (+), o cursor gráfico avança para o ponto de número 40 mostrando abaixo, os seus respectivos valores. Através do cursor gráfico é possível observar as ocorrências espectroscópicas como a posição e a intensidade da banda de absorção de um determinado composto orgânico. Este cursor gráfico pode ser

manuseado pelas sub-rotinas "ShowCurrentSpectrum", "CompareCurrent Spectrum" e "ShowTwoSpectrum".

Um outro recurso permite transferir o espectro corrente apresentado na tela do monitor, para uma impressora pressionando-se simplesmente a tecla "P". A sub-rotina "PRINTSPECTRUM" se encarrega dessa tarefa.

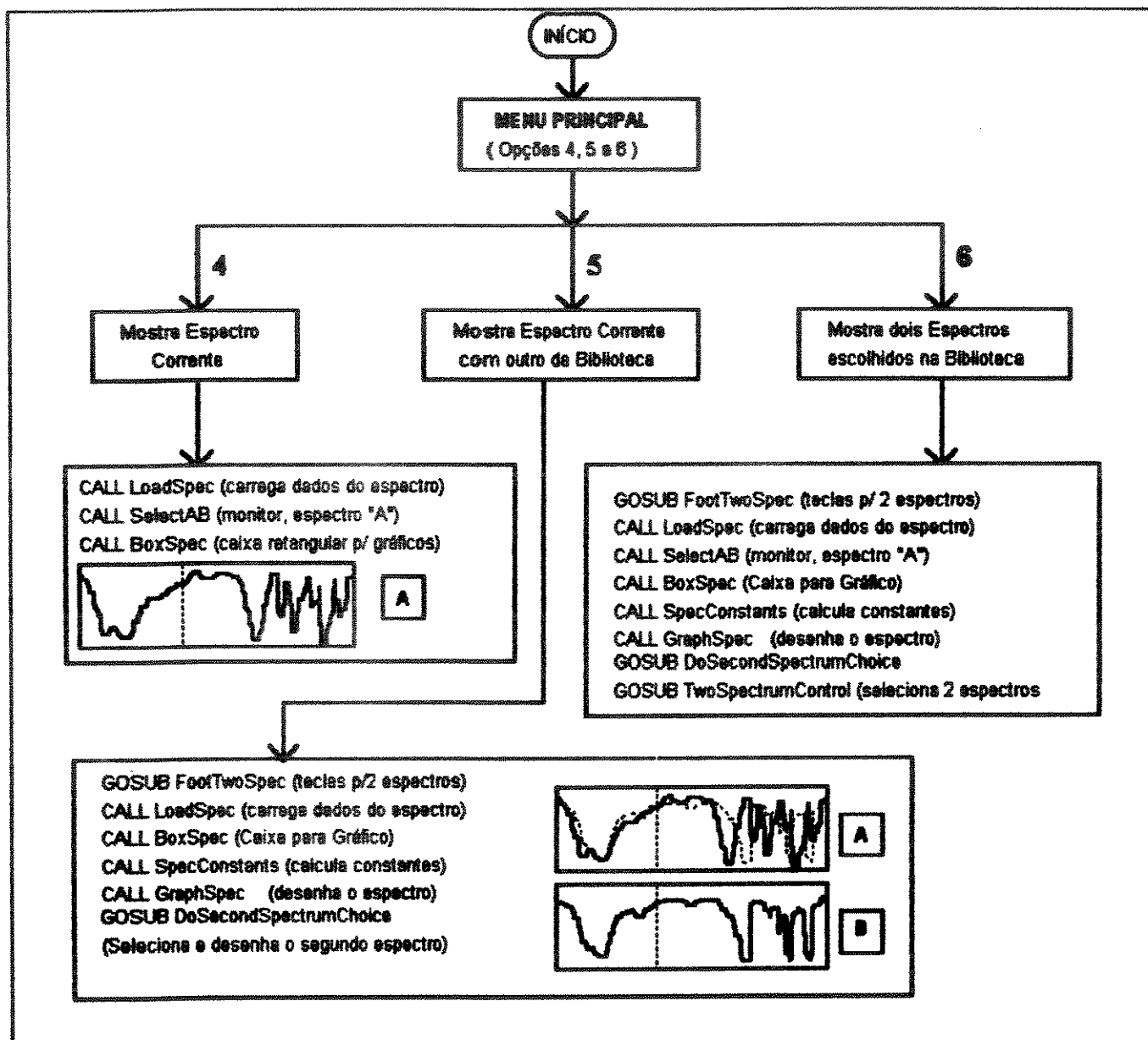


FIGURA III-10 - Diagrama de blocos das sub-rotinas de manipulações dos espectros da biblioteca. Os números representam as opções do menu principal.

A sub-rotina "SuperPosition" é um outro recurso muito importante para as opções 5 e 6 do menu principal ("CompareCurrentSpectrum" e "ShowTwoSpectrum"), porém, requer a apresentação de dois espectros na tela do monitor. A "Superposition", é utilizada para fazer a comparação de dois espectros referidos como "A" e "B" existentes na biblioteca. Para tal comparação deve-se apenas pressionar a tecla "S", que o espectro "B" sobrepõe-se ao espectro "A". No entanto, o espectro "B" original, isto é, aquele inicialmente mostrado na tela do monitor, quando selecionado, permanece abaixo do espectro "A". A principal vantagem em se utilizar a "SuperPosition" em conjunto com o cursor gráfico está na sua habilidade em comparar, automaticamente dois espectros, mostrando simultaneamente as suas ocorrências espectroscópicas.

As etapas das quatro principais sub-rotinas de manipulações de espectros da biblioteca são descritas abaixo.

III.4.1 - SUB-ROTINAS PARA MANIPULAÇÕES DE ESPECTROS

A - SHOWCURRENTSPECTRUM

Essa sub-rotina é acionada quando seleciona-se a opção número 4 do menu principal. Ela foi implementada para mostrar o espectro corrente na tela do monitor, selecionado pelo usuário, através da opção número 8 do menu principal (Tomar um Espectro Corrente), ou então o último espectro coletado pelo programa PROGEIV, como descrito anteriormente em aquisição de dados espectrais.

Inicialmente, a "ShowCurrentSpectrum" chama a sub-rotina "LOADSPEC" para carregar os dados do arquivo do espectro. Em seguida, configura-se na tela do monitor uma caixa em forma de retângulo, através da sub-rotina "BOXSPEC", conforme o tipo de monitor escolhido no início da execução do programa. Finalmente, a sub-rotina

"GRAPHSPEC" coloca dentro dessa caixa o espectro, a partir dos valores dos pares de dados (transmitâncias e números de ondas).

B - COMPARECURRENTSPECTRUM

Essa sub-rotina é executada quando seleciona-se a opção número 5 do menu principal. Ela tem como meta comparar o espectro corrente com um outro espectro existente na biblioteca. Portanto, essa sub-rotina procede como a sub-rotina "ShowCurrentSpectrum", até o momento que o usuário seleciona um outro espectro da biblioteca, para tal comparação. Em seguida, a "CompareCurrentSpectrum" chama a sub-rotina "DoSecondSpectrumChoice". Após a seleção do espectro, esta encarrega-se de colocar abaixo do primeiro espectro, referido como espectro "A", um segundo espectro, agora referido como espectro "B".

C - SHOWTWOPECTRUM

Essa sub-rotina, quando acionada, através da opção número 6 do menu principal, encarrega-se de apresentar na tela do monitor para uma comparação, dois espectros existentes na biblioteca, escolhidos pelo usuário. A apresentação dos espectros na tela do monitor é praticamente similar ao da sub-rotina citada anteriormente, nesse caso; a única diferença é que o usuário deve selecionar os espectros "A" e "B". Também, nessa sub-rotina, quando os dois espectros estão apresentados na tela do monitor, a sub-rotina "TwoControlSpectrum" é chamada para controlá-los quanto ao uso do cursor gráfico. Deste modo, escolhe-se a tecla "A" ou a tecla "B", e pressiona-se continuamente uma das teclas do cursor gráfico (←, →, - ou +), e um ponto percorrerá o espectro. Mais abaixo das

respectivas áreas de tela originais desses espectros, são mostrados os nomes dos compostos, e os números dos pontos coletados, com seus valores dos números de ondas e transmitâncias correspondentes.

D - HOUSEKEEPING

Essa sub-rotina é utilizada quando seleciona-se a opção de número 9 do menu principal. A sua implementação tem como objetivo remover um determinado espectro da biblioteca. Inicialmente, aparece na tela do monitor uma pergunta com relação a retirada ou não do espectro. No caso de confirmação, o usuário necessita apenas selecionar o espectro não mais desejado como componente da biblioteca. Após isso, a pergunta retorna novamente, até o usuário pressionar a tecla "N" para retornar ao menu principal. Cada vez que um espectro é removido da biblioteca, o arquivo "setspec.se" que contém os números e os nomes dos compostos existentes na biblioteca, é reformulado automaticamente, isto é, o número do espectro subsequente é decrementado em uma unidade. Por exemplo, em uma biblioteca contendo 100 espectros de I.V., o arquivo "setspec.se" armazena estes espectros da seguinte maneira:

```
"spectrum\SPE001.SPT", "3-BROMOPROPIONIC (Br-CH2-CH2-COOH)"
```

```
.....
```

```
"spectrum\SPE062.SPT", "PIVALIC ACID (PROPIONIC ACID_2-2-DIMETHYL)"
```

```
.....
```

```
"spectrum\SPE100.SPT", "ISOBUTYRIC ACID ((CH3)2-CH-COOH)"
```

Se o espectro "SPE062.SPT" foi escolhido para ser removido, após a execução, o arquivo "setspec.se" é reformulado permanecendo da seguinte maneira:

```
"spectrum\SPE001.SPT", "3-BROMOPROPIONIC (Br-CH2-CH2-COOH)"
```

```
.....
```

```
"spectrum\SPE099.SPT", "ISOBUTYRIC ACID ((CH3)2-CH-COOH)"
```

III.5 - RESULTADOS E DISCUSSÕES

O programa PROGEIV foi desenvolvido para a elaboração, manipulação e manutenção de uma biblioteca de espectros de I.V.. Para a aquisição de dados espectrais, o programa permite a utilização de uma mesa digitalizadora, como a HM construída no laboratório, ou uma mesa tipo comercial, como a modelo Digigraf DaVinci. Uma outra alternativa é traduzir arquivos de discos flexíveis gerados pelo microcomputador do espectrofotômetro. Estes arquivos são obtidos na forma binária e, posteriormente, são expandidos para a forma decimal em ASCII, pela sub-rotina "PERKIN".

O acesso para a seleção das diferentes rotinas do programa é feita através de menus. Estes menus foram elaborados de maneira clara e objetiva, direcionando o usuário quanto ao controle do programa. Uma biblioteca de até mil espectros pode ser controlada pelo programa PROGEIV. Os arquivos dos espectros contêm dados como o nome do composto e pares de pontos amostrados (transmitância e números de ondas).

O programa PROGEIV permite o manuseio de diversas operações do arquivo do espectro. Por exemplo, o usuário pode observar os espectros na forma gráfica na tela do monitor. Na figura III-11 estão ilustrados separadamente, os espectros do ácido dibromoacético. Na parte superior da figura, consta o espectro digitalizado pela mesa HM, apresentado na tela do monitor, através da seleção da sub-rotina "ShowCurrentSpectrum". A parte inferior da figura exibe o espectro original da biblioteca SADTLER do mesmo composto. Uma comparação pode ser feita, com relação às informações coletadas pela mesa digitalizadora e o espectro original. Através do cursor gráfico observa-se que a posição e a intensidade da banda de absorção da carbonila, para o ácido carboxílico são praticamente idênticas às do espectro original. Portanto, pode-se concluir que a mesa digitalizadora preserva, essencialmente, as informações contidas no espectro original.

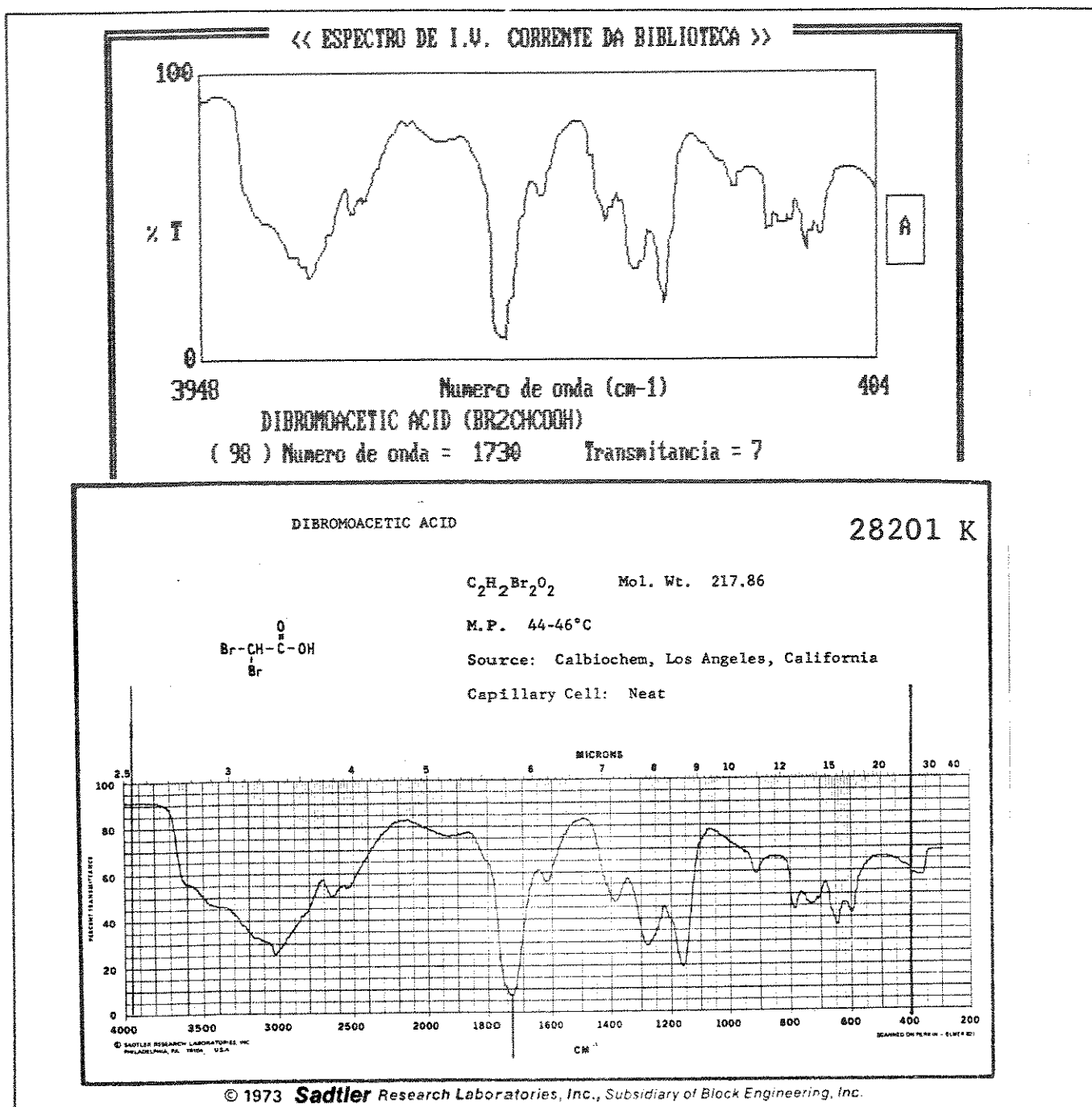


FIGURA III-11 - Espectros de I.V. do ácido dibromoacético. A parte superior da figura mostra o espectro digitalizado pela mesa HM e apresentado na tela do monitor pela sub-rotina "ShowCurrentSpectrum". Mais abaixo, são mostrados o nome do composto, e o número do ponto coletado, com seus valores de número de onda e transmitância, para a banda da carbonila (C=O). A parte inferior, ilustra o espectro original da biblioteca SADTLER do mesmo composto.

A mesa digitalizadora HM foi construída no laboratório e utilizada neste trabalho como uma ferramenta de base que permitiu a inicialização deste trabalho, uma vez que uma base de dados espectroscópica de I.V., se fazia necessária ao desenvolvimento do S.E.. Sua construção facilitou o entendimento da parte eletrônica, quanto à conversão de dados analógicos para digitais e a sua coleta automática para um microcomputador, similar aos da instrumentação moderna. Também, no início do projeto não existia uma mesa digitalizadora tipo comercial, disponível para tais objetivos. No entanto, durante o desenvolvimento do programa PROGEIV constatou-se que uma mesa digitalizadora, modelo DaVinci, estava disponível num laboratório do IQ/UNICAMP. Através de um interesse comum entre o grupo de pesquisa deste trabalho e a professora responsável pela aquisição dessa mesa digitalizadora, foi possível digitalizar vários espectros de I.V. de compostos orgânicos, oriundos da biblioteca SADTLER. Para isso, foi necessário implementar uma sub-rotina para a coleta de dados provenientes desse tipo de mesa, o que tornou o programa PROGEIV de uso mais amplo. A mesa DaVinci é melhor confeccionada do que a da mesa HM, o que facilita em muito a digitalização dos espectros. Esses fatos justificam porque a maior parte dos espectros utilizados neste trabalho, foram digitalizados através da mesa DaVinci.

Um outro recurso oferecido pelo programa PROGEIV é a possibilidade de apresentar simultaneamente, dois espectros na tela do monitor, para uma comparação, através das sub-rotinas "CompareCurrentSpectrum" ou "ShowTwoSpectrum". Um cursor gráfico selecionado para cada espectro ("A" ou "B") pode ser posicionado ao longo do eixo dos números de ondas (posição horizontal). Portanto, quando uma das teclas que direcionam o cursor é pressionada aparecem em tempo real, abaixo dos gráficos originais, o nome do composto, o número do ponto amostrado e os valores do número de onda com a transmitância correspondente.

A sub-rotina "SuperPosition" é utilizada para sobrepor um espectro "B" num espectro "A", com linhas portilhadas, sem retirar o espectro "B" original, da tela do

monitor. As linhas pontilhadas são importantes não apenas para a observação durante a comparação de espectros, mas para distingui-los entre si, quando da transferência dos espectros para uma impressora.

A figura III-12 ilustra a maneira como é apresentada na tela do monitor uma comparação entre um espectro corrente e um outro, selecionado pelo usuário, através da sub-rotina "CompareCurrentSpectrum". Essa visualização na tela do monitor facilita muito o pesquisador quanto à procura de um fato espectroscópico, como por exemplo, na comparação das posição e intensidade de uma banda de absorção. Na figura III-13 estão apresentados dois espectros existentes na biblioteca e escolhidos aleatoriamente pelo usuário, através da sub-rotina "ShowTwoSpectrum".

Finalmente, no programa PROGEIV, quando se seleciona a opção 7 do menu principal, ativa-se a sub-rotina "CALLPROLOG", que transforma o arquivo Quick-Basic "spec.spt" em um outro arquivo formatado como lista Turbo-Prolog. Desta forma, os dados são preparados para serem acessados pelo S.E. desenvolvido (EXPERTIV). A lista Prolog é armazenada através de nomes DOS "spexxx.ltd", colocados no diretório "SISTSPEC", onde xxx é um número entre 000 e 999, . A figura III-14 mostra parte um de arquivo de um espectro da biblioteca, referido como "spec.spt" e, seu arquivo formatado como lista "spec.ltd".

Também, deve-se ressaltar que embora o programa PROGEIV foi inicialmente desenvolvido com o objetivo de criar um banco de dados espectroscópico de I.V., para ser utilizado pelo programa especialista, resultou em um ferramenta muito importante para o químico orgânico, que poderá usufruir de todos os seus recursos descritos anteriormente.

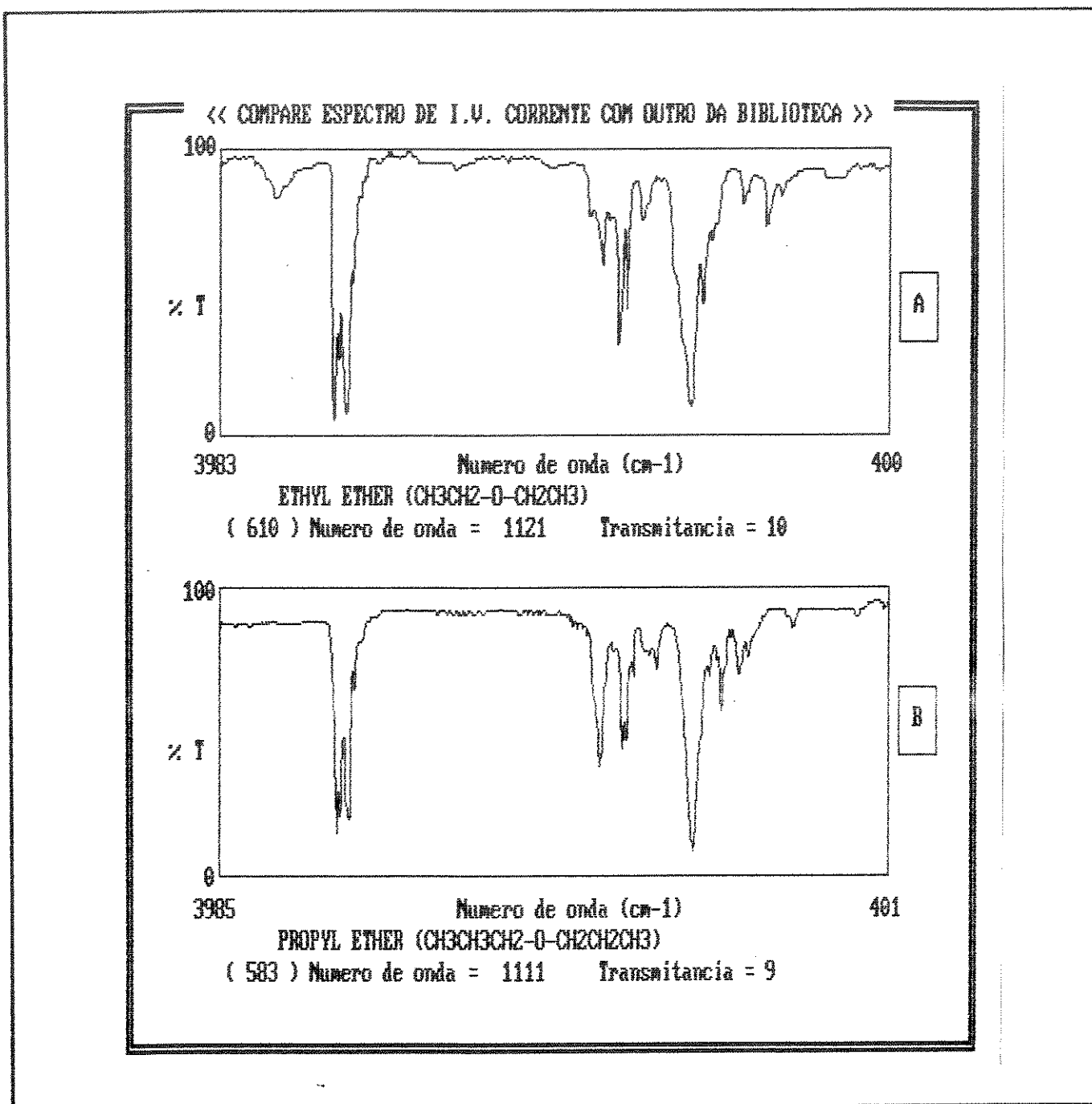


FIGURA III-12 - Tela do monitor do microcomputador mostrando o espectro de I.V. corrente e um outro espectro da biblioteca selecionado pelo usuário, para comparação. Abaixo dos espectros constam os nomes dos compostos, os pontos coletados e os seus números de ondas, referentes a posição de um cursor gráfico..

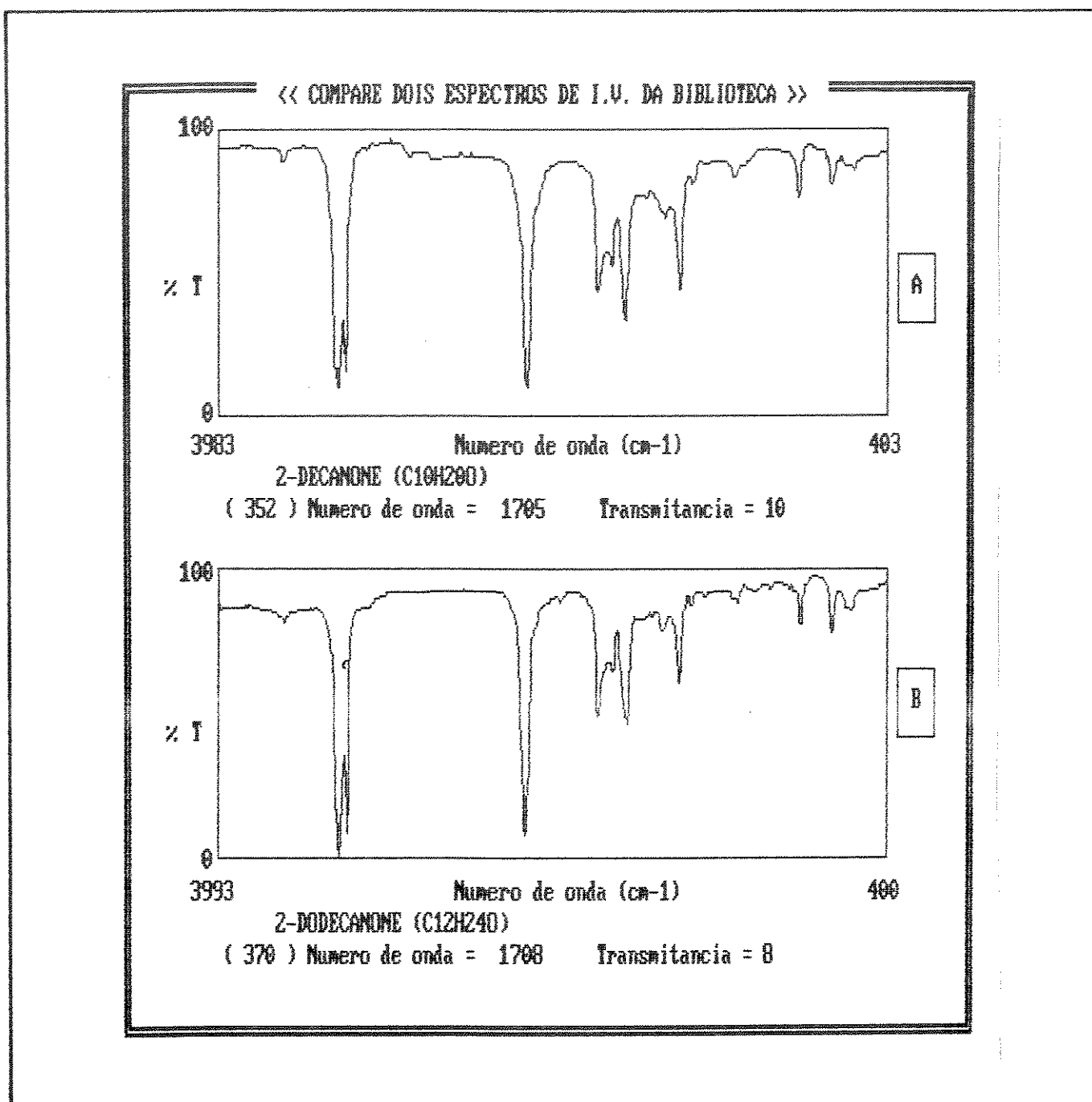


FIGURA III-13 - Tela do monitor do microcomputador mostrando dois espectros de I.V. presentes na biblioteca, selecionados aleatoriamente pelo usuário, para comparação. Abaixo dos espectros constam os nomes dos compostos, os pontos coletados e os seus números de ondas, referentes a posição de um cursor gráfico.

"2-PENTANOL (CH₃CH₂CH₂-CH(OH)CH₃)"

84,3886
 84,3833
 85,3772
 85,3719
 85,3658
 83,3612
 80,3567
 76,3559
 72,3529
 68,3529
 64,3521
 61,3506
 57,3506
 53,3506
 50,3460
 46,3452
 42,3452
 39,3437
 35,3430
 34,3384
 32,3384
 30,3376
 28,3376
 26,3376
 25,3323
 26,3285
 27,3255
 29,3240
 33,3240

 85,404

2-PENTANOL (CH₃CH₂CH₂-CH(OH)CH₃)

[84,3886,84,3833,85,3772,85,3719,85,3658,83,3612,80,3567,76,3559,72,3529,68,3529,64,3521,61,3506,57,3506,
 53,3506,50,3460,46,3452,42,3452,39,3437,35,3430,34,3384,32,3384,30,3376,28,3376,26,3376,25,3323,26,3285,
 27,3255,29,3240,33,3240,36,3217,40,3217,43,3179,47,3156,51,3148,55,3141,58,3110,62,3103,65,3087,68,3034,
 70,3019,73,3019,72,2981,70,2966,66,2943,62,2943,59,2943,54,2943,50,2943,46,2943,42,2943,39,2943,35,2943,
 31,2943,28,2943,21,2943,19,2943,17,2943,15,2943,13,2943,11,2943,13,2943,14,2943,16,2943,18,2928,19,2890,
 21,2890,23,2890,24,2882,28,2875,24,2875,22,2875,24,2867,26,2867,28,2867,32,2867,35,2852,39,2852,
82,450,83,427,85,404]

FIGURA III-14 - Partes dos arquivos DOS "spec.spt" (dados de entrada para o programa PROGEIV manipular o espectro) e "spec.ltd" que é a sua saída em formato de lista Turbo Prolog, para ser utilizada pelo S.E. desenvolvido, referente ao espectro de I.V. do álcool 2-pentílico presente na biblioteca.

CAPÍTULO IV

PROGRAMA ESPECIALISTA PARA IDENTIFICAÇÃO DE GRUPOS FUNCIONAIS DE COMPOSTOS ORGÂNICOS - EXPERTIV

Atualmente, o interesse no desenvolvimento de sistemas especialistas (S.E.) na área da química vem aumentando consideravelmente. Esta atenção tem mostrado ao químico em geral e, particularmente, ao químico analítico, a necessidade de ferramentas práticas para a resolução de problemas, baseadas no conhecimento de um especialista. A disponibilidade de uma linguagem declarativa de alto nível, do tipo Turbo Prolog, capaz de manusear informações simbólicas em microcomputadores, tende a aumentar o número de S.E. de menor porte, que são construídos para a resolução de problemas mais específicos.

Como o nome sugere, um S.E. ou sistema baseado em conhecimento é capaz de simular os procedimentos de um especialista humano, para a resolução de um problema de sua área. Em geral, um programa especialista, orientado para interpretação de espectros de I.V. incorpora uma base de conhecimento (que armazena regras de produção ou de inferência), construída através de tabelas de correlações da literatura, e também com a ajuda de especialistas na área de Química Orgânica. As regras de produção são amplamente utilizadas na geração de sistemas baseados em conhecimento. Uma regra de produção consiste de um conjunto de condições e um conjunto de ações que descrevem as conclusões a serem obtidas quando estas condições são encontradas. O fator determinante do desempenho de um S.E. está na qualidade do seu banco de conhecimento que permite, a um mecanismo de inferência, a obtenção de conclusões.

Como conseqüências da implementação de um S.E. para interpretação de espectros de I.V. podem distinguir-se como mais importantes; a) a possibilidade que o conhecimento dessa área estar presente em qualquer laboratório que utilize espectros de I.V. nos trabalhos de rotina, b) a possibilidade de empregá-lo como delimitador na

localização dos espectros em bibliotecas (sistemas de busca), ou seja, uma vez determinada a presença de certos grupos funcionais, a procura se limitaria aos espectros que contêm tais grupos funcionais. Neste caso, o S.E. seria utilizado como um filtro diminuindo assim, o tempo, a memória e o espaço de armazenamento necessários ao programa e, c) o programa, com pequenas modificações na sua parte interativa com o usuário pode ser utilizado no ensino da interpretação de espectros de I.V..

Deste modo, o programa especialista desenvolvido neste trabalho, denominado de EXPERTIV, aplicado na interpretação do espectro de I.V., busca num primeiro instante, sem a interferência do usuário, a identificação da presença de grupos funcionais existentes no composto orgânico que originou o espectro. Além da rapidez e da mínima intervenção do operador, buscou-se aproveitar uma das vantagens do sistema baseado em conhecimento, que é a de fornecer uma explicação plausível, ao se chegar a uma determinada conclusão no decorrer da interpretação.

Além disso, o desenvolvimento do programa EXPERTIV é útil, como uma primeira etapa, quando se visa um S.E. mais complexo, que contenha um banco de conhecimento referente à interpretação de espectros de massas e de RMN. Desta forma, poderia se utilizar dados dessas técnicas instrumentais, com objetivo de elucidar a estrutura de um composto orgânico. Em particular, convém lembrar que não é necessário escrever novamente o programa EXPERTIV, para acomodar novos dados.

IV.1 - DESCRIÇÃO DO PROGRAMA EXPERTIV

O programa EXPERTIV foi desenvolvido com o objetivo de identificar automaticamente grupos funcionais de compostos orgânicos e outras características específicas da molécula, através dos dados de espectros de I.V.. Esse programa foi escrito em linguagem Turbo Prolog versão 2.0 da Borland[87-91], executável num microcomputador tipo IBM-PC-AT/386. A Prolog foi escolhida como a linguagem de programação devido à sua facilidade de manipulação da informação simbólica para os dados espectrais. Uma vez que o espectro digitalizado é colocado na forma de uma lista (formato requerido pela Prolog), o programa pode ser utilizado para interpretação do espectro, através de regras baseadas em conhecimentos. Esta interpretação possibilita a identificação dos grupos funcionais presentes no composto que originou o espectro.

A interface usuário-microcomputador é efetuada, através da seleção da opção número 7 do menu principal do programa PROGEIV, o qual permite a conexão das linguagens Quick-Basic (de procedimentos) e Turbo Prolog (declarativa). Inicialmente, o programa PROGEIV chama a sub-rotina "CALLPROLOG", que por sua vez transforma os dados do espectro corrente da biblioteca, armazenado no arquivo DOS "spec.spt", em um outro arquivo DOS formatado como lista Turbo-Prolog, denominado "spec.ltd". Desta forma, os dados são preparados para serem acessados pelo programa EXPERTIV.

Uma das características de um sistema baseado em conhecimento é a separação entre a base de conhecimento e o mecanismo de inferência. A base de conhecimento, como o nome sugere, contém os conhecimentos (fatos) necessários para resolver os problemas do domínio especialista, enquanto que o mecanismo de inferência utiliza os algoritmos para manusear estes conhecimentos. Esta separação permite um desenvolvimento modular oferecendo melhor entendimento e fácil adaptação ao usuário.

A base de conhecimento consiste nas regras de produção implementadas à partir de dados dos espectros de I.V.. A base de conhecimento foi elaborada, inserindo as informações das tabelas de correlações para um arquivo DOS, obtidas de alguns livros de espectroscopia no I.V.[104,105] e principalmente do livro do "*Infrared Absorption Spectroscopy*"[106]. Algumas consultas com especialistas da área sobre seus métodos de resolução do problema (algumas vezes subjetivas), também contribuíram para a obtenção de resultados satisfatórios.

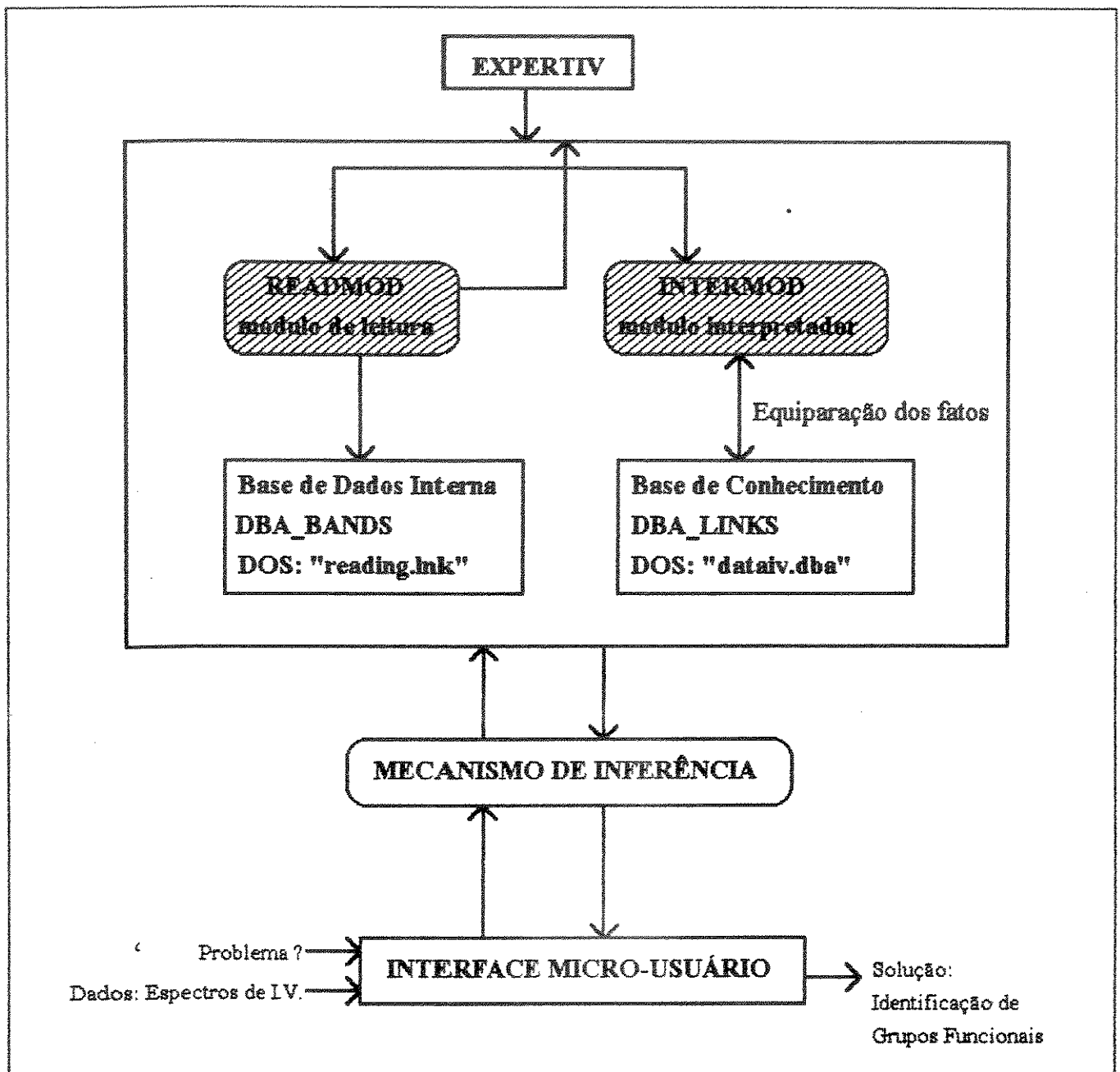
O programa atua primeiro sobre as regras impostas para efetuar uma leitura do espectro, extraíndo as bandas de absorção, e posteriormente nas regras de interpretação equiparando com a base de conhecimento, elaborada através de tabelas de correlações para a identificação de grupos funcionais de compostos orgânicos. Por exemplo, para se estabelecer que um composto é um ácido carboxílico devem ser extraídas do espectro, as bandas de absorção, correspondentes às vibrações de estiramento O-H, C=O e C-O. A partir desses fatos, as regras implementadas para interpretação do espectro os compara com a base de conhecimento, conferindo se essas vibrações estão dentro dos limites das faixas de absorção permitidas para o ácido carboxílico.

O programa EXPERTIV é constituído de dois módulos principais. O primeiro, denominado módulo de leitura, atua sobre a lista Turbo Prolog que contém os dados espectrais coletados pelo programa PROGEIV, produzindo, como resultado final, a extração de informações relevantes para a interpretação do espectro como a posição, a intensidade, a largura e o formato da banda de absorção. O segundo módulo é encarregado da interpretação do espectro propriamente dito, atuando sobre o conjunto de fatos amostrados no módulo de leitura (extração de ocorrências espectroscópicas) e utilizando inúmeras regras de inferência baseadas em conhecimento. Neste módulo, também foi implementada uma função matemática, similar às ferramentas de lógica difusa[107-111], com o objetivo de tornar o programa EXPERTIV mais completo e de direcionar algumas

conclusões de forma menos determinística. Normalmente, os especialistas da área utilizam termos relacionados às ocorrências espectrais de uma maneira subjetiva. A inclusão de ferramentas de lógica difusa, conduz a identificação de uma banda de absorção de forma mais flexível e portanto, mais próximas da linguagem dos especialistas. Por exemplo, eles se referem a "uma banda de absorção larga e de intensidade forte" ou "uma banda de absorção aguda de intensidade variável".

Na figura IV-1 está ilustrado o esquema de desenvolvimento dos componentes do programa EXPERTIV.

Os espectros de I.V. de compostos orgânicos puros foram obtidos da biblioteca SADTLER e de um espectrofotômetro de I.V., como descritos anteriormente. Até o momento, os espectros existentes na biblioteca do programa PROGEIV estão agrupados segundo a presença de suas funções orgânicas. No entanto, foram estudados pelo programa EXPERTIV, apenas quatorze grupos funcionais, sendo seis grupos de compostos carbonilados e oito não carbonilados.

**FIGURA IV-1** - Componentes do programa EXPERTIV.

IV.1.1 - MÓDULO DE LEITURA

O módulo de leitura denominado READMOD, tem por objetivo extrair da melhor forma possível e similarmente a um especialista da área, as informações relevantes do espectro de I.V. de um determinado composto orgânico, isto é, as suas ocorrências espectroscópicas como a posição, a intensidade, a largura e o formato da banda de absorção.

A base de conhecimento consiste de regras de produção implementadas para a extração dos dados espectrais, que são provenientes do programa PROGEIV. Em seguida, o mecanismo de inferência atua sobre estas regras armazenando os dados extraídos em um arquivo DOS, que é posteriormente requisitado pelo módulo de interpretação. No momento, o READMOD consta de 64 regras de produção e em sua forma executável ocupa 114 Kbytes de memória. A forma de operações desse módulo pode ser observada na figura IV-2.

O espectro de I.V. original do composto 2-butanona, obtido da biblioteca SADTLER, está ilustrado na figura IV-3. Os dados desse espectro são coletados pelo programa PROGEIV, armazenados no arquivo DOS "spec.spt" e, em seguida utilizados pelo READMOD. Durante a descrição desse módulo, serão apresentados os fatos obtidos na leitura do espectro do referido composto, através das figuras de seus arquivos DOS, quando suas metas são atingidas pelas regras implementadas.

No programa PROGEIV a sub-rotina "CALLPROLOG" prepara o arquivo DOS "spec.spt", referente ao espectro corrente da biblioteca, transformando-o no arquivo DOS "spec.ltd" em formato de lista Prolog, requisitado pelo programa EXPERTIV. A figura IV-4 mostra esse arquivo, referente ao espectro original do 2-butanona. Os elementos contidos na lista de informações espectrais são os pares de dados coletados, isto é, transmitâncias e números de ondas, em unidades de porcentagem e cm^{-1} , respectivamente.

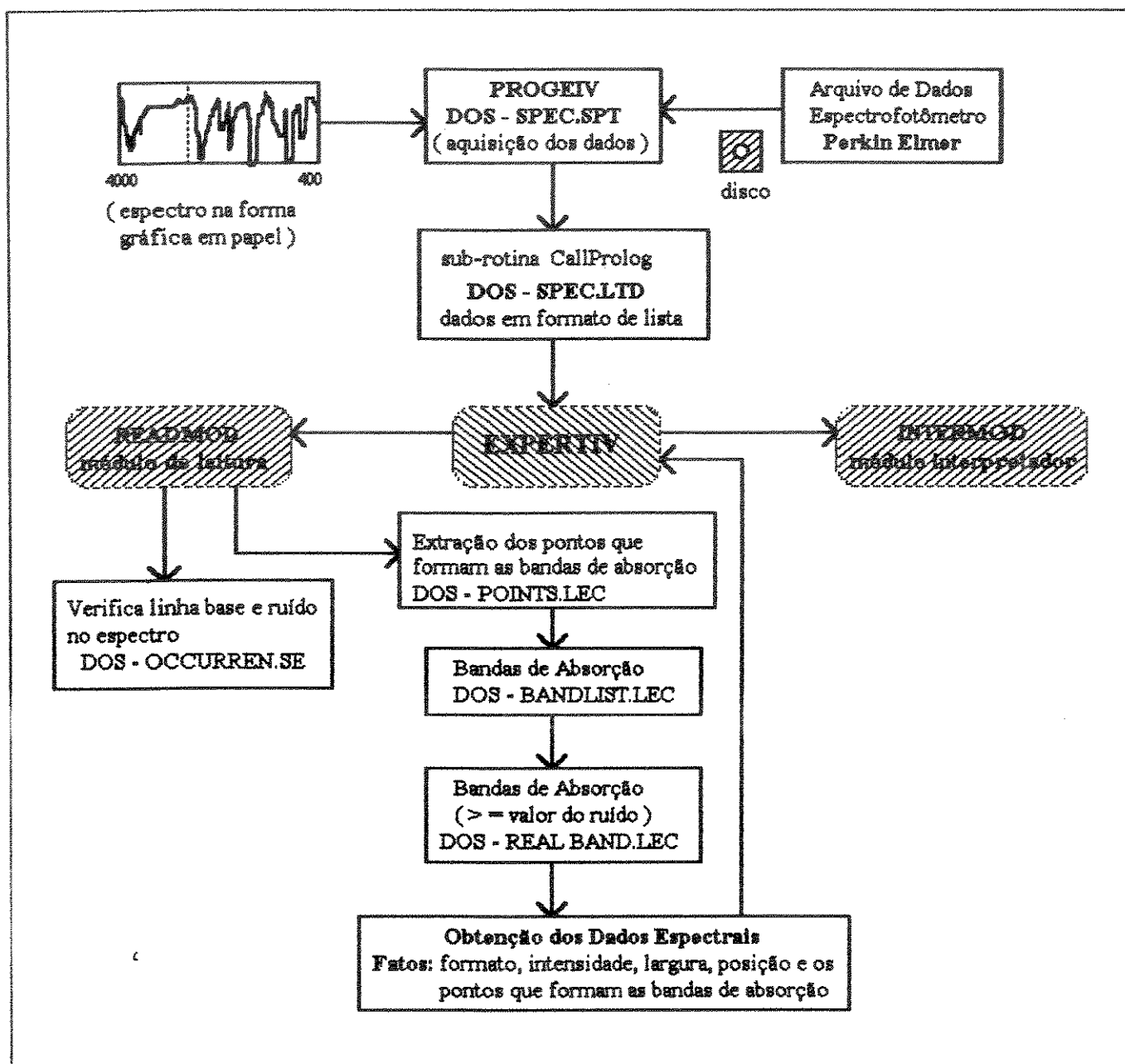


FIGURA IV-2 - Desenvolvimento do módulo de leitura READMOD para obtenção dos dados espectrais.

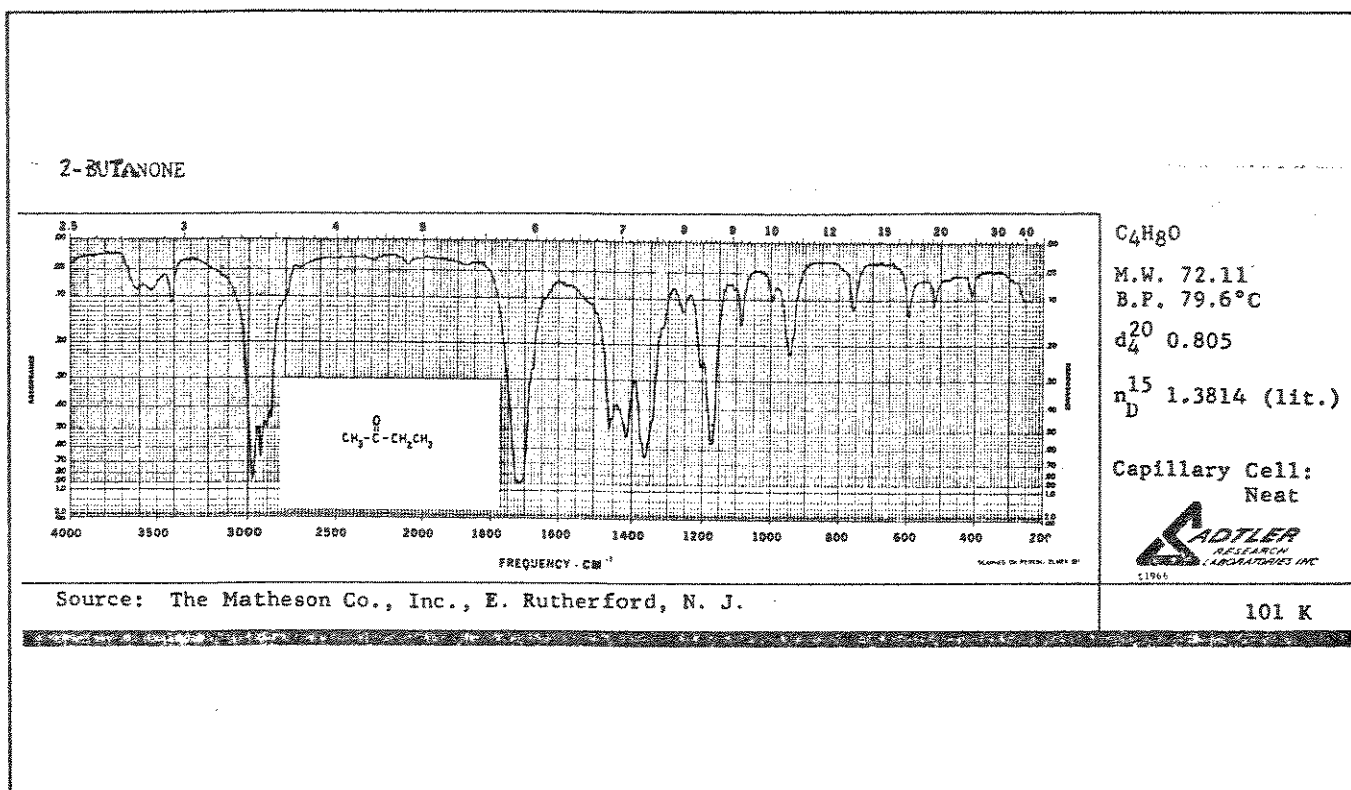


FIGURA IV-3 - Espectro de I.V. original da biblioteca SADTLER do 2-butanona.

2-BUTANONE (CH₃COCH₂CH₃)

93,3987,93,3975,94,3964,94,3951,94,3937,95,3925,95,3910,96,3897,95,3886,95,3871,95,3859,95,3847,96,3834,
 96,3820,96,3805,96,3791,96,3780,96,3766,96,3751,96,3732,96,3720,96,3708,95,3697,94,3688,93,3682,92,3671,
 90,3666,89,3661,87,3654,86,3645,85,3632,84,3619,84,3605,84,3594,85,3582,85,3570,84,3556,83,3544,83,3530,
 84,3517,85,3504,85,3493,87,3485,88,3473,89,3461,89,3450,87,3444,86,3434,84,3428,83,3427,81,3427,80,3423,
 79,3412,80,3405,82,3405,83,3403,85,3402,86,3398,88,3397,89,3387,91,3380,92,3373,94,3361,94,3349,94,3337,
 94,3326,94,3313,94,3299,94,3284,94,3269,94,3257,94,3246,94,3233,93,3219,93,3205,92,3194,91,3181,91,3169,
 90,3153,90,3142,89,3129,88,3116,88,3104,86,3093,85,3080,83,3074,82,3067,80,3054,79,3042,78,3034,76,3030,
 75,3027,73,3024,72,3016,70,3014,69,3014,67,3012,65,3009,64,3008,62,3007,60,3006,59,3006,57,3005,56,3003,
 54,3001,52,2999,51,2996,49,2994,48,2993,46,2991,45,2991,43,2991,41,2991,40,2990,38,2988,37,2986,35,2985,
 33,2983,32,2982,30,2981,28,2980,27,2978,25,2977,23,2976,22,2976,20,2975,19,2972,17,2971,15,2969,14,2967,
 15,2961,17,2961,18,2961,20,2961,909,34,2907,33,2895,35,2887,36,2886,38,2881,37,2869,39,2867,40,2867,42,
 2866,43,2866,45,2866,46,2866,48,2865,49,2864,52,2864,54,2862,56,2861,57,2859,59,2859,60,2859,62,2859,63,
 2858,64,2858,66,2856,68,2854,70,2850,71,2844,73,2839,74,2839,76,2837,77,2828,77,2817,79,2808,80,2796,82,
 2783,83,2779,85,2774,86,2769,88,2764,89,2760,91,2754,92,2742,92,2729,92,2716,92,2704,92,2693,92,2681,92,
 2668,93,2655,93,2643,93,2631,94,2620,94,2607,94,2591,94,2576,94,2563,94,2550,94,2538,94,2524,94,2512,95,
 2497,95,2483,95,2470,95,2456,95,2442,95,2427,95,2412,95,2396,95,2383,95,2371,95,2358,95,2346,94,2334,94,
 2323,94,2311,94,2296,94,2283,94,2270,94,2258,94,2242,95,2228,95,2214,95,2201,95,2186,95,2174,95,2162,95,
 2150,95,2136,94,2123,94,2111,93,2099,92,2087,92,2076,93,2064,94,2051,94,2039,94,2027,94,2016,94,2002,94,
 1995,95,1989,95,1983,95,1976,94,1968,94,1961,94,1951,94,1943,94,1936,94,1930,94,1924,94,1918,93,1912,93,
 1906,93,1900,93,1894,93,1888,92,1882,92,1875,92,1868,92,1862,93,1856,93,1850,93,1843,93,1837,93,1830,93,
 824,92,1818,91,1812,90,1806,89,1800,88,1796,86,1790,85,1787,83,1785,82,1782,80,1778,78,1776,77,1774,75,

continua na próxima página a lista Prolog

continuação

1772,74,1769,72,1767,70,1766,69,1765,67,1765,66,1762,64,1761,63,1759,61,1758,59,1756,58,1756,56,1755,55,
 1755,53,1752,52,1751,50,1750,48,1749,47,1747,45,1747,44,1746,42,1746,40,1745,39,1743,37,1741,36,1739,34,
 1738,32,1737,31,1737,29,1736,27,1736,26,1735,24,1732,23,1731,21,1731,19,1730,18,1729,16,1727,15,1725,13,
 1720,13,1713,13,1706,14,1700,15,1697,17,1696,18,1696,19,1696,21,1696,23,1694,24,1693,26,1692,27,1691,29,
 1691,31,1691,32,1691,34,1690,36,1689,37,1689,39,1689,41,1689,42,1689,44,1688,45,1688,47,1687,48,1687,50,
 1683,52,1682,53,1681,55,1676,56,1674,58,1673,60,1673,61,1673,63,1670,64,1670,66,1670,67,1669,69,1663,70,
 1661,72,1661,73,1659,75,1655,76,1652,77,1648,79,1646,79,1640,81,1634,82,1628,83,1622,85,1616,85,1610,86,
 1605,86,1599,86,1593,86,1587,85,1580,85,1574,85,1568,85,1562,85,1555,84,1548,82,1542,81,1536,80,1530,79,
 1524,79,1517,78,1512,78,1505,76,1501,75,1497,73,1492,72,1488,70,1484,69,1480,67,1479,66,1478,64,1476,63,
 1474,61,1474,60,1473,58,1471,57,1470,55,1470,53,1470,52,1469,50,1468,48,1467,47,1467,45,1465,44,1464,42,
 1463,41,1463,39,1463,38,1463,36,1462,34,1462,33,1462,35,1457,36,1455,36,1450,38,1448,40,1448,41,1442,40,
 1436,38,1430,37,1427,35,1424,34,1422,32,1420,30,1418,31,1413,32,1410,34,1410,35,1410,37,1409,38,1408,40,
 1407,41,1405,43,1403,45,1402,46,1401,48,1401,49,1398,48,1393,46,1393,45,1389,43,1389,42,1386,40,1385,39,
 1385,37,1384,36,1383,34,1382,32,1381,31,1381,29,1378,28,1374,26,1371,24,1370,23,1368,23,1363,24,1360,26,
 1359,28,1358,29,1356,31,1354,32,1353,34,1351,36,1348,37,1346,39,1344,40,1343,42,1343,43,1343,45,1342,47,
 1342,48,1341,50,1339,52,1338,53,1337,55,1337,57,1336,58,1335,60,1335,61,1334,63,1334,65,1330,67,1327,68,
 1326,70,1319,71,1312,72,1311,74,1308,75,1307,77,1305,78,1304,80,1302,81,1298,83,1294,84,1289,82,1283,82,
 1277,80,1275,78,1271,77,1266,76,1262,75,1257,77,1252,78,1250,80,1246,81,1242,81,1235,79,1232,78,1229,76,
 1225,75,1224,73,1223,71,1222,69,1221,68,1220,66,1219,65,1218,63,1216,62,1215,60,1214,59,1211,57,1211,56,
 1209,55,1206,56,1200,55,1194,53,1192,52,1191,50,1190,49,1190,47,1189,45,1189,44,1187,42,1187,41,1187,39,
 1186,37,1185,36,1184,34,1183,33,1182,31,1181,30,1181,28,1179,28,1173,29,1172,31,1171,32,1170,34,1169,35,
 1168,37,1167,39,1166,40,1166,42,1165,44,1165,45,1165,46,1165,48,1164,50,1164,52,1163,53,1163,55,1163,56,
 1160,58,1159,60,1157,61,1156,63,1156,65,1155,66,1154,68,1153,70,1153,71,1151,73,1147,75,1143,76,1143,78,
 1143,79,1143,81,1141,82,1137,84,1132,85,1125,85,1120,85,1113,83,1112,84,1106,84,1100,82,1097,81,1096,79,
 1094,78,1091,76,1090,75,1090,73,1090,72,1088,73,1084,75,1083,76,1082,78,1081,80,1080,81,1078,83,1074,85,
 1073,86,1070,88,1064,89,1060,90,1054,90,1047,90,1041,90,1035,90,1029,89,1024,89,1018,87,1012,86,1011,85,
 1005,83,1003,82,1003,80,1003,79,996,80,994,82,993,84,989,84,983,83,976,81,971,80,969,78,968,77,966,75,963,
 73,962,72,961,70,960,69,959,67,958,66,957,64,956,63,954,61,953,60,952,59,945,61,945,62,945,64,944,65,943,6
 7,939,69,937,70,936,72,935,73,934,75,934,77,933,78,931,80,928,81,926,83,923,85,921,86,918,88,913,89,910,90,
 906,92,899,93,892,93,886,93,878,93,863,93,845,93,838,93,831,93,824,93,817,93,811,93,805,91,799,91,792,90,7
 86,89,780,87,776,86,774,84,774,83,773,81,771,80,770,78,769,76,768,77,761,78,757,80,757,81,755,83,753,85,74
 9,86,746,88,744,89,742,90,736,92,732,93,726,93,720,93,713,93,707,93,701,93,695,93,689,93,682,93,677,92,671,
 92,663,93,658,93,651,92,644,92,638,91,632,91,625,89,622,89,617,87,614,85,611,84,609,82,609,80,608,79,607,7
 7,605,76,605,74,603,73,597,75,597,77,597,78,597,80,594,81,590,83,588,84,586,85,580,86,574,86,568,86,561,86,
 555,86,549,86,543,86,537,84,532,83,529,81,526,79,526,78,525,79,519,81,516,82,512,84,511,85,505,87,499,87,4
 94,87,488,87,481,87,474,87,467,88,460,88,454,88,447,88,441,88,435,88,429,86,422,84,417,83,414,81,412,82,
 406,83,405]c

FIGURA IV-4 - Arquivo DOS "spec.ltd" em formato de lista Prolog, referente ao espectro de I.V. original da biblioteca SADTLER do 2-butanona, utilizado como dados de entrada (fatos) no READMOD.

Inicialmente, observa-se através do READMOD as ocorrências de algumas características do espectro, como por exemplo se o primeiro ponto coletado é maior ou igual a 3850 cm^{-1} . A necessidade de que o espectro sob análise contenha dados acima do referido número de onda, deve-se ao fato de que o READMOD infere sobre a presença e o nível de ruído em valores de número de onda acima de 3850 cm^{-1} . Além disso, verifica a posição da linha base do espectro, de forma a analisar se está ou não adequada para a sua leitura e interpretação. As regras de produção implementadas para essas metas são determinísticas. Elas sempre devem ser verdadeiras para que se possa continuar com esse módulo. Quando falham na obtenção de alguma meta, impede-se o prosseguimento para o módulo de interpretação.

Para verificar o primeiro número de onda da lista Prolog (pares de dados amostrados), se é maior ou igual a 3850 cm^{-1} . A regra foi imposta da seguinte maneira:

% nota-se que na regra abaixo o predicado tem 2 argumentos (duas listas de n^{os}. inteiros)

`verifica_primeiro_numero_de_onda([_,Wavenum_],DeltaWave) :-`

`Wavenum >= 3850,`

`% Wavenum, primeiro número de onda da lista`

`DeltaWave = Wavenum - 100, !.`

`% DeltaWave, variável para usar em outra regra`

Assim, ao confirmar o corpo dessa regra (as duas sub-metas), diz-se que ela foi bem sucedida. Caso contrário, admite-se que o espectro não está adequado para o READMOD continuar com a sua leitura. A segunda observação está relacionada com a presença ou não de ruídos no espectro. Nesse caso, as informações espectrais ficam visivelmente prejudicadas, principalmente em se tratando das sobreposições de bandas e suas transmitâncias mínimas. A variável de saída DeltaWave, colocada na regra mencionada anteriormente, liga em uma outra regra para calcular o nível de ruído no espectro. Essa variável é uma lista de pares de dados compreendendo a faixa de absorção do primeiro número de onda coletado menos 100 (cem). Por exemplo, na lista `[_3987|Lista_de_Pontos]`, subtrai-se 100 do número de onda 3987, então, a faixa de

absorção para verificar o ruído do espectro é de 3987 a 3887 cm^{-1} , aplicada aos espectros dos compostos orgânicos utilizados neste trabalho. Este procedimento foi definido desta forma, porque nesse intervalo do espectro, normalmente, não há ocorrência de bandas de absorção. O ruído no espectro é calculado pela diferença dos valores de transmitâncias máxima e mínima dessa faixa de absorção, não podendo ultrapassar a 9 %T. Todavia, se essa diferença for menor ou igual a 1 %T, estipula-se o valor mínimo de variação da transmitância igual 3 %T, para considerá-la como uma possível banda de absorção, a ser extraída pelo módulo de leitura.

Quanto à linha base do espectro, foi estipulado que ela não pode ser menor do que 70 %T, compreendendo a faixa de absorção de 2500 a 2000 cm^{-1} . O principal motivo para essa escolha é a não ocorrência de muitas absorções nessa região, ao contrário da região de "impressão digital" e da região acima de 2500 cm^{-1} . Para atingir esta meta no programa, descarta-se as bandas de absorção que ocorrerem na referida região do espectro. Depois, é feita uma somatória, apenas com os outros pontos amostrados e, finalmente, calcula-se o valor médio das transmitâncias, obtendo-se assim o valor da transmitância da linha base do espectro.

Um arquivo DOS denominado "ocurren.se" armazena as ocorrências observadas durante a leitura do espectro, retornando na tela do monitor uma mensagem para o usuário, referente ao fato do espectro estar ou não adequado para uma interpretação.

Todas as regras implementadas para observar as possíveis anomalias de um espectro de I.V. devem ser bem sucedidas. Deste modo, considera-se que o espectro está devidamente adequado para continuar sua leitura, ou seja, com a extração dos pontos que formam as bandas de absorção.

Uma regra foi implementada de maneira recursiva para a extração dos pontos, constituindo-se das seguintes sub-metas: a) uma regra com uma constante mínima para uma derivada, onde são determinados os pontos amostrados, quando há uma inflexão na curva

do espectro e, b) uma regra para verificar onde inicia a próxima inflexão na curva do espectro. Essas regras foram inseridas no READMOD da seguinte forma:

```

/* Regra recursiva com 2 sub-metas. A segunda sub-meta é quando a inflexão da curva é para baixo */
extraindo_pontos([Transmit,Wavenum,TransmitB,WavenumB|PointsList],
                 [TransDown,WaveDown|JustPoints]) :-
derivada([Transmit,Wavenum,TransmitB,WavenumB],Derivada),
        Derivada > derMinDown,                % derMinDown é um valor constante = 0.08
pontos_descendo([Transmit,Wavenum,TransmitB,WavenumB|PointsList],
                [TransDown,WaveDown],DownList),
extraindo_pontos([TransDown,WaveDown|DownList],JustPoints), !.                % recursão

extraindo_pontos([],[]) :- !.                % condição de parada

derivada([Transmit,Wavenum,TransmitB,WavenumB],Derivada) :-
        Wavenum <> WavenumB,
        Derivada = (round (Transmit - TransmitB) / (Wavenum - WavenumB)), !.

```

Na cláusula acima citada, a primeira sub-meta derivada verifica se um número de onda é diferente do próximo número de onda da lista e, em seguida determina o valor da variável Derivada através da razão entre a diferença das transmitâncias e dos números de ondas. O valor encontrado é equiparado com o valor da constante mínima para a derivada (colocado na seção de constantes, derMinDown), dependendo da inflexão da curva do espectro, isto é, conforme os valores das transmitâncias diminuem, aumentam ou inclinam suavemente. A regra é recursiva, ou seja, vai retrocedendo até que a lista se torne vazia ([]) completando assim a extração dos pontos do arquivo "spec.ltd" (lista Prolog). A figura IV-5 ilustra o espectro original do 2-butanona, mas agora apenas com os pontos extraídos com as regras citadas, isto é, obtenção das informações úteis relacionadas às bandas de absorção. Como se pode observar nesta etapa do READMOD as regras atuam de forma a reduzir os dados do espectro original. Também, abaixo dessa figura está o arquivo DOS "points.lec", onde constam esses fatos.

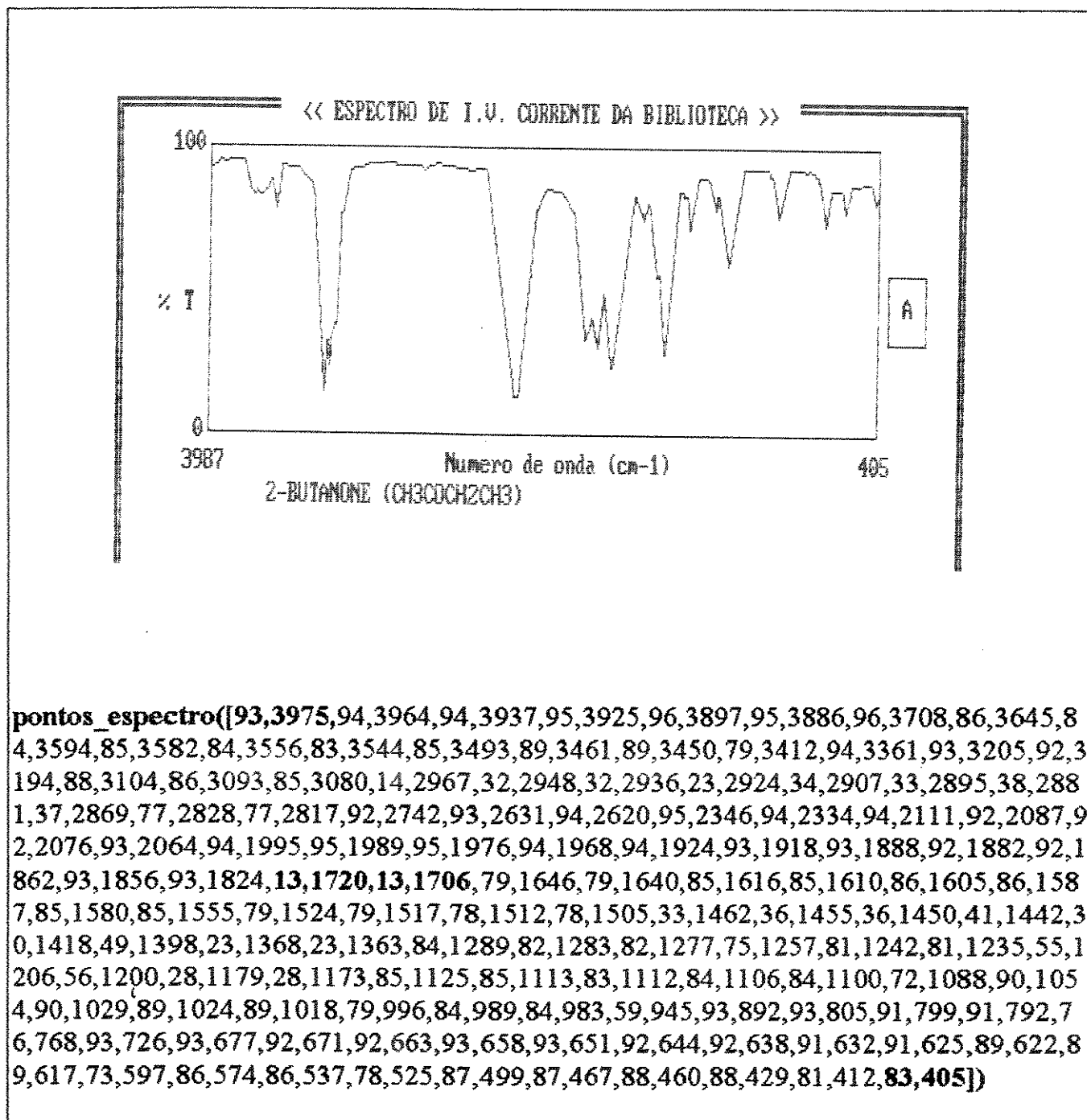


FIGURA IV-5 - A parte superior representa graficamente o espectro de I.V. original do 2-butanona, empregando-se somente os dados originados da atuação do módulo de leitura. Na parte inferior consta os fatos (pontos extraídos do espectro), que são armazenados no arquivo DOS "points.lec".

Uma vez que os pontos que formam as bandas de absorção do espectro foram extraídos, deve-se separá-los para obter o formato de cada banda de absorção. Desta maneira, pode-se calcular as suas larguras e transmitâncias mínimas. As regras inseridas para atingir esta meta são similares àquelas da extração de pontos, isto é, utilizam a regra da derivada para os pontos de inflexões da curva no espectro. O arquivo "points.lec" contém todos os pontos que formam uma banda de absorção, sendo necessário apenas acrescentar uma regra, que armazena os pontos que configuram o formato de cada banda. Os formatos das bandas mais comumente encontrados nos espectros de I.V. foram estabelecidos como simples, de base achatada, com sobreposição a esquerda e/ou a direita e, com sobreposição a esquerda e/ou a direita com bases achatadas, como mostra a figura IV-6. Por exemplo, para configurar o formato de uma banda simples é necessário extrair três pares de dados, como os elementos contidos na lista: [90,3045,10,2920,85,2765]. A figura IV-7 mostra a saída do seu arquivo DOS "bandlist.lec", referente ao espectro do 2-butanona, quando esta meta é alcançada com sucesso.

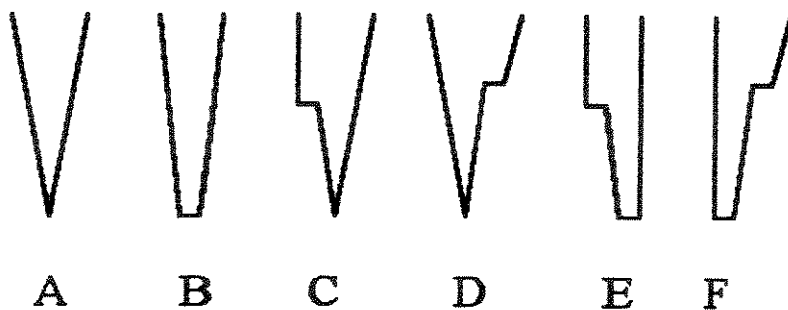


FIGURA IV-6 - Formatos das bandas de absorção mais comumente encontradas num espectro de I.V.. A) banda simples, B) base achatada, C) com sobreposição à esquerda, D) com

sobreposição à direita, E) com sobreposição à esquerda e base achatada e, F) com sobreposição à direita e base achatada.

```

lista_bandas([banda("sobreposição_esquerda_base_achatada",[96,3897,95,3886,96,3708,
86,3645,84,3594,85,3582]),banda("base_achatada",[84,3556,83,3544,85,3493,89,3461]),
banda("banda_normal",[89,3450,79,3412,94,3361]),
banda("sobreposição_esquerda",[88,3104,86,3093,85,3080,14,2967,32,2948]),
banda("banda_normal",[32,2936,23,2924,34,2907]),
banda("banda_normal",[34,2907,33,2895,38,2881]),
banda("banda_normal",[38,2881,37,2869,77,2828]),
banda("sobreposição_esquerda_base_achatada",[95,2346,94,2334,94,2111,92,2087,92,20
76,93,2064]),banda("sobreposição_esquerda_base_achatada",[94,1924,93,1918,93,1888,9
2,1882,92,1862,93,1856]),
banda("base_achatada",[93,1824,13,1720,13,1706,79,1646]),
banda("sobreposição_esquerda",[79,1517,78,1512,78,1505,33,1462,36,1455]),
banda("banda_normal",[41,1442,30,1418,49,1398]),
banda("base_achatada",[49,1398,23,1368,23,1363,84,1289]),
banda("sobreposição_esquerda",[84,1289,82,1283,82,1277,75,1257,81,1242]),
banda("banda_normal",[81,1235,55,1206,56,1200]),
banda("base_achatada",[56,1200,28,1179,28,1173,85,1125]),
banda("banda_normal",[85,1113,83,1112,84,1106]),
banda("banda_normal",[84,1100,72,1088,90,1054]),
banda("sobreposição_esquerda",[90,1029,89,1024,89,1018,79,996,84,989]),
banda("banda_normal",[84,983,59,945,93,892]),
banda("sobreposição_esquerda",[93,805,91,799,91,792,76,768,93,726]),
banda("base_achatada",[93,677,92,671,92,663,93,658]),
banda("sobreposição_esquerda",[91,625,89,622,89,617,73,597,86,574]),
banda("banda_normal",[86,537,78,525,87,499]),
banda("banda_normal",[88,429,81,412,83,405]))

```

FIGURA IV-7 - Arquivo DOS "bandlist.lec" para o espectro de I.V. do 2-butanona, mostrando separadamente as bandas de absorção extraídas (fatos).

Após a devida separação das bandas de absorção, ainda é necessário descartar as que tem valores de transmitâncias menores do que a calculada na faixa de absorção, já conferidas, com a presença ou não de ruído no espectro. Desta forma, os fatos do arquivo DOS "bandlist.lec" são ligados em uma outra regra que os filtra, armazenando somente as bandas que fornecem dados espectrais considerados suficientemente seguros para uma

interpretação. A figura IV-8 apresenta o arquivo DOS "realband.lec", onde constam as bandas que devem ser posteriormente analisadas, para calcular a largura, a intensidade e a posição da banda de absorção.

```
real_bandas([banda("sobreposição_esquerda_base_achatada",
[96,3897,95,3886,96,3708,86,3645,84,3594,85,3582]),
banda("banda_normal",[89,3450,79,3412,94,3361]),
banda("sobreposição_esquerda",[88,3104,86,3093,85,3080,14,2967,32,2948]),
banda("banda_normal",[32,2936,23,2924,34,2907]),
banda("banda_normal",[38,2881,37,2869,77,2828]),
banda("base_achatada",[93,1824,13,1720,13,1706,79,1646]),
banda("sobreposição_esquerda",[79,1517,78,1512,78,1505,33,1462,36,1455]),
banda("banda_normal",[41,1442,30,1418,49,1398]),
banda("base_achatada",[49,1398,23,1368,23,1363,84,1289]),
banda("sobreposição_esquerda",[84,1289,82,1283,82,1277,75,1257,81,1242]),
banda("banda_normal",[81,1235,55,1206,56,1200]),
banda("base_achatada",[56,1200,28,1179,28,1173,85,1125]),
banda("banda_normal",[84,1100,72,1088,90,1054]),
banda("sobreposição_esquerda",[90,1029,89,1024,89,1018,79,996,84,989]),
banda("banda_normal",[84,983,59,945,93,892]),
banda("sobreposição_esquerda",[93,805,91,799,91,792,76,768,93,726]),
banda("sobreposição_esquerda",[91,625,89,622,89,617,73,597,86,574]),
banda("banda_normal",[86,537,78,525,87,499])])
```

FIGURA IV-8 - Arquivo DOS "realband.lec" para o espectro de I.V. do 2-butanona, mostrando as bandas de absorção maiores ou iguais ao valor calculado para o ruído do espectro.

Em seguida, para cada banda de absorção, independente do seu formato, calcula-se a largura e a transmitância mínima. A posição da banda de absorção é considerada como o número de onda correspondente com a sua transmitância mínima.

A largura das bandas de absorção são calculadas com base na região do espectro em que estão localizadas. Por exemplo, a largura das bandas de absorção acima

de 2500 cm^{-1} , é calculada a três quartos da altura (largura a $\frac{3}{4}$ da altura) e das bandas de absorção abaixo dessa região, é calculada a um quarto da altura (largura a $\frac{1}{4}$ da altura). Estes critérios foram estabelecidos conforme consultas feitas com alguns especialistas da área de Química Orgânica.

As regras implementadas para calcular a largura de uma banda de absorção foram inseridas com a primeira sub-meta encarregada de encontrar a transmitância média e a segunda sub-meta para encontrar o número de onda mais próximo desta transmitância média. Essas regras atuam para os dois lados (esquerdo e direito) de uma banda de absorção, como ilustrada na figura IV-9. Por exemplo, para calcular a largura da banda de absorção, referente à vibração de estiramento C=O do espectro de I.V. do 2-butanona, uma regra foi imposta da seguinte maneira:

```

encontre_transmitancia_media("base_achatada",
  [Transmit,_,TransmitB,_,TransmitC,WaveC,TransmitD,WaveD],           % dados de entrada
  [TransmitMed1,TransmitC,WaveC,TransmitMed2, TransmitD,WaveD]) :-   % dados de saída
    TransMed1 = ( trunc ( Transmit + TransmitB ) / 2 ),
    TransmitMed1 = ( trunc ( TransmitB + TransMed1 ) / 2 ),
    TransMed2 = ( trunc ( TransmitC + TransmitD ) / 2 ),
    TransmitMed2 = ( trunc ( TransmitB + TransMed2 ) / 2 ), !.

```

No caso da banda de absorção da carbonila extraída do arquivo "realband.lec" (veja figura IV-8), o cálculo é efetuado da seguinte maneira:

% banda("base_achatada",[93,1824,13,1720,13,1706,79,1646] referente a banda de carbonila

```

encontre_transmitancia_media("base_achatada",[93,_,13,_,13,1706,79,1646],
  [TransmitMed1,13,1706,TransmitMed2,79,1646]) :-           % dados de saída
  TransMed1 = ( trunc ( 93 + 13 ) / 2 ),                     % ½ altura, lado esquerdo
  TransmitMed1 = ( trunc ( 13 + 53 ) / 2 ),                  % ¼ da altura, lado esquerdo
  TransMed2 = ( trunc ( 13 + 79 ) / 2 ),                     % ½ altura, lado direito
  TransmitMed2 = ( trunc ( 13 + 46 ) / 2 ), !.                % ¼ da altura, lado direito

```


Em seguida, uma regra é colocada para encontrar os números de ondas iguais ou mais próximos possíveis, dos valores das variáveis calculadas anteriormente ($\text{TransmitMed1} = 33 \% T$ e $\text{TransmitMed2} = 30 \% T$), correspondendo aos números de ondas encontrados 1737 e 1691 cm^{-1} para cada lado da banda de absorção. Portanto, a largura da banda de absorção da carbonila (largura a $\frac{1}{4}$ da altura), do 2-butanona é igual a 46 cm^{-1} , que é a diferença dos referidos números de ondas e, considerada no programa EXPERTIV como sendo de largura média.

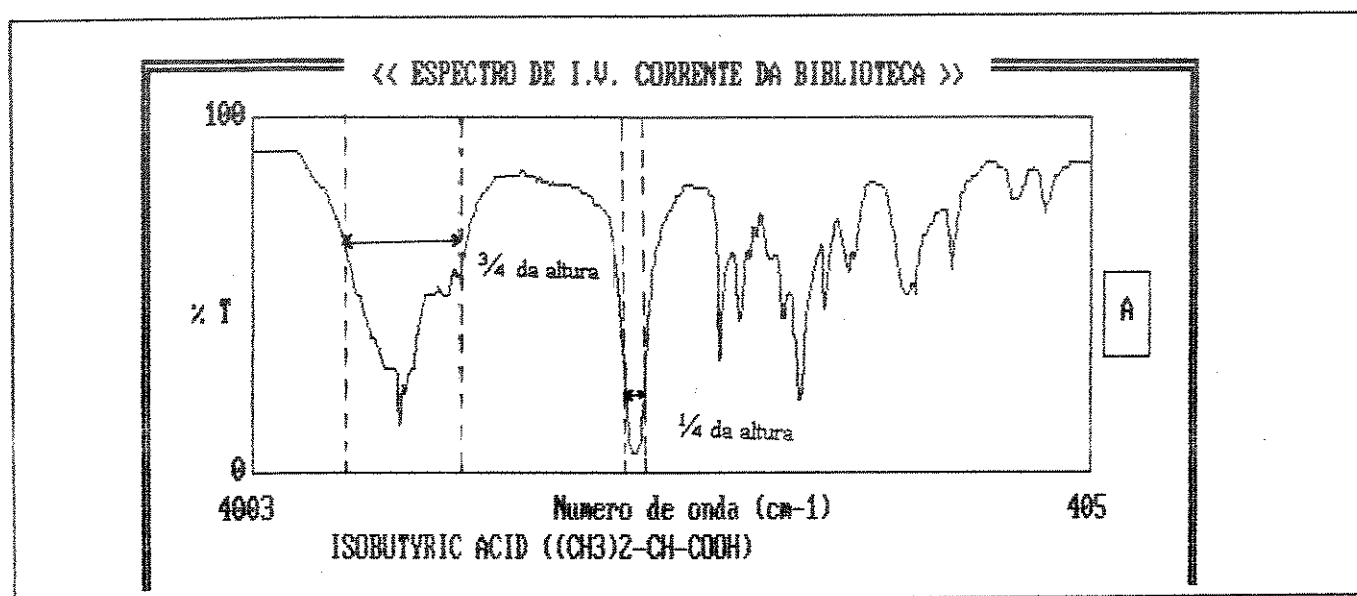


FIGURA IV-9 - Bandas de absorção de uma vibração de estiramento O-H e de uma carbonila, mostrando as larguras a $\frac{3}{4}$ e a $\frac{1}{4}$ da altura.

Através do formato de uma banda de absorção determina-se a transmitância mínima e o seu número de onda correspondentes à sua posição. Por exemplo, para a banda de absorção da carbonila do 2-butanona, as regras foram assim impostas:

% lista de pontos que configuram banda da carbonila: [93,1824,13,1720,13,1706,79,1646]

determine_transmitância("base_achatada", [_,_,13,1720,13,1706|_], MinTransmit,
posição(Posição)) :-

Posição = (trunc (1720 + 1706) / 2),

transmitância_mínima([13,13], MinTransmit), !.

A transmitância mínima e a posição encontradas para a banda de absorção da carbonila do 2-butanona são iguais a 13 %T e 1713 cm^{-1} , respectivamente.

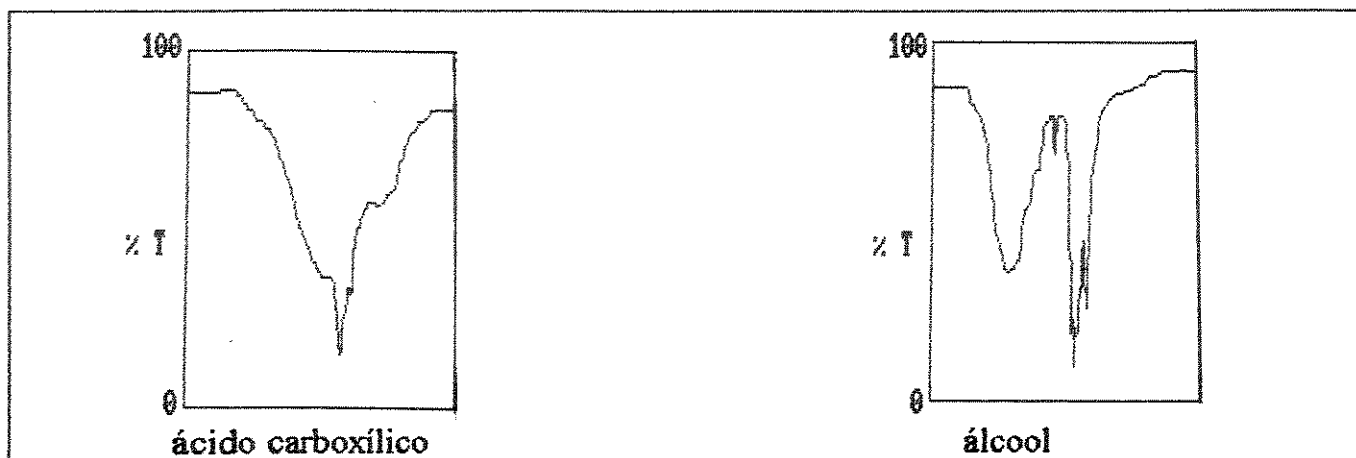
As metas do programa, até o momento descritas, devem ser bem sucedidas para formar um banco de dados espectrais com o formato, a largura, a intensidade, a posição e os pares de dados que configuram as bandas de absorção do espectro (transmitância e número de onda). Esses dados serão posteriormente requisitados pelo módulo interpretador, para identificação de grupos funcionais de um determinado composto orgânico.

Além das informações obtidas na primeira leitura do espectro, o READMOD verifica a existência de alguma banda de absorção nas características de uma vibração de estiramento C=O. Em caso positivo, é feita uma segunda leitura na região entre 3640 a 2500 cm^{-1} , onde os espectros de I.V. de álcoois, fenóis e ácidos carboxílicos podem apresentar uma ou mais bandas de absorção, de modo que a mais importante é a vibração de estiramento O-H, característica desses grupos funcionais. Caso contrário, os fatos da primeira leitura são suficientes para serem acessados ao módulo interpretador.

Baseada nas informações obtidas por especialistas da área de química orgânica, para distinguir um grupo funcional alcoólico ou fenólico de um ácido carboxílico, implementou-se uma regra para esta segunda leitura do espectro, que procura a ocorrência de uma banda de absorção com a sua intensidade maior do que a intensidade média da maior banda de absorção do espectro em análise. Também, foi utilizada novamente a regra da derivada, para determinar os pontos de inflexões da curva no espectro, porém o valor das constantes mínimas da derivada (derMinDown e derMinUp) foram alteradas para 0.15 e - 0.15, para a diminuição ou aumento dos valores de transmitâncias. A razão dessa mudança é tornar a extração dos pontos menos sensível, ou seja, evitar a obtenção dos pontos de uma provável sobreposição de bandas de absorção. Desta forma, uma regra foi implementada para buscar o primeiro e o último ponto de inflexão da curva na região

citada. Em seguida, é definida a transmitância mínima e calculada a largura a $\frac{3}{4}$ da altura, da mesma forma apresentada anteriormente (veja figura IV-9).

As bandas de absorção na região do espectro entre 3640 a 2500 cm^{-1} , para a banda de estiramento O-H de um ácido carboxílico e de um álcool, podem ser vistas na figura IV-10, onde abaixo da mesma estão os seus arquivos DOS "highreg.lnk" com as informações obtidas dessas bandas.



```
banda(tipo("regiao_alta"),largura(3401,2520,881,"cm-1"),intensidade(14,"%T"),
posicao(2952),pontos([87,3709,14,2952,60,2594])))
```

```
banda(tipo("regiao_alta"),largura(3534,3198,336,"cm-1"),intensidade(36,"%T"),
posicao(3359),pontos([88,3656,36,3359,72,3061])))
```

```
banda(tipo("regiao_alta"),largura(3061,3053,8,"cm-1"),intensidade(69,"%T"),
posicao(3053),pontos([71,3061,69,3053,72,2992])))
```

```
banda(tipo("regiao_alta"),largura(2969,2893,76,"cm-1"),intensidade(9,"%T"),
posicao(2924),pontos([72,2992,9,2924,45,2863])))
```

```
banda(tipo("regiao_alta"),largura(2855,2840,15,"cm-1"),intensidade(26,"%T"),
posicao(2847),pontos([32,2863,26,2847,81,2756])))
```

FIGURA IV-10 - Bandas de estiramento O-H de um ácido carboxílico e de um álcool. Abaixo, estão as saídas dos seus arquivos DOS "highreg.lnk", com as suas devidas informações espectrais da região acima de 2500 cm^{-1} .

Finalmente, após essas leituras, os fatos conhecidos sobre o espectro em análise (dados espectrais) são armazenados em tempo real na base de dados interna "*dba_bands*" e salvos no arquivo DOS "*reading.lnk*". No momento, pode-se considerar que estes fatos representam da melhor forma (simbólica), os dados extraídos de um espectro de I.V., os quais podem ser acessados ao módulo de interpretação. A figura IV-11 mostra este tipo de arquivo, referente aos fatos extraídos do 2-butanona.

```

banda(banda(formato("sobreposição_esquerda_base_achatada"),
largura(3661,3605,56,"cm-1"),intensidade(84,"%T"),posição(3620),
pontos([96,3897,95,3886,96,3708,86,3645,84,3594,85,3582])))
banda(banda(formato("banda_normal"),largura(3427,3403,24,"cm-1"),
intensidade(79,"%T"),posição(3412),pontos([89,3450,79,3412,94,3361])))
banda(banda(formato("sobreposição_esquerda"),largura(2983,2961,22,"cm-1"),
intensidade(14,"%T"),posição(2967),pontos([88,3104,86,3093,85,3080,14,2967,32,2948])))
banda(banda(formato("banda_normal"),largura(2925,2921,4,"cm-1"),intensidade(23,"%T"),
posição(2924),pontos([32,2936,23,2924,34,2907])))
banda(banda(formato("banda_normal"),largura(2881,2865,16,"cm-1"),
intensidade(37,"%T"),posição(2869),pontos([38,2881,37,2869,77,2828])))
banda(banda(formato("base_achatada"),largura(1737,1691,46,"cm-1"),
intensidade(13,"%T"),posição(1713),pontos([93,1824,13,1720,13,1706,79,1646])))
banda(banda(formato("sobreposição_esquerda"),largura(1465,1457,8,"cm-1"),
intensidade(33,"%T"),posição(1462),pontos([79,1517,78,1512,78,1505,33,1462,36,1455])))
banda(banda(formato("banda_normal"),largura(1420,1410,10,"cm-1"),intensidade(30,"%T"),
posição(1418),pontos([41,1442,30,1418,49,1398])))
banda(banda(formato("base_achatada"),largura(1378,1344,34,"cm-1"),
intensidade(23,"%T"),posição(1366),pontos([49,1398,23,1368,23,1363,84,1289])))
banda(banda(formato("sobreposição_esquerda"),largura(1266,1252,14,"cm-1"),
intensidade(75,"%T"),posição(1257),pontos([84,1289,82,1283,82,1277,75,1257,81,1242])))
banda(banda(formato("banda_normal"),largura(1215,1206,9,"cm-1"),intensidade(55,"%T"),
posição(1206),pontos([81,1235,55,1206,56,1200])))
banda(banda(formato("base_achatada"),largura(1183,1165,18,"cm-1"),
intensidade(28,"%T"),posição(1176),pontos([56,1200,28,1179,28,1173,85,1125])))
banda(banda(formato("banda_normal"),largura(1090,1081,9,"cm-1"),intensidade(72,"%T"),
posição(1088),pontos([84,1100,72,1088,90,1054])))
banda(banda(formato("sobreposição_esquerda"),largura(1003,994,9,"cm-1"),intensidade(79,"%T"),
posição(996),pontos([90,1029,89,1024,89,1018,79,996,84,989])))
banda(banda(formato("banda_normal"),largura(956,937,19,"cm-1"),intensidade(59,"%T"),
posição(945),pontos([84,983,59,945,93,892])))
banda(banda(formato("sobreposição_esquerda"),largura(770,757,13,"cm-1"),
intensidade(76,"%T"),posição(768),pontos([93,805,91,799,91,792,76,768,93,726])))
banda(banda(formato("sobreposição_esquerda"),largura(605,597,8,"cm-1"),
intensidade(73,"%T"),posição(597),pontos([91,625,89,622,89,617,73,597,86,574])))
banda(banda(formato("banda_normal"),largura(526,516,10,"cm-1"),intensidade(78,"%T"),
posição(525),pontos([86,537,78,525,87,499])))

```

FIGURA IV-11 - Fatos da base de dados interna "*dba-bands*", inseridos em tempo de execução através do predicado padrão "*assertz*" e salvos no arquivo DOS "*reading.lnk*", referente a leitura do espectro de I.V. do 2-butanona. Esses fatos estão prontos para serem acessados no módulo interpretador. Os dados da banda de absorção posicionada em 1713 cm^{-1} , referem-se aos fatos obtidos de uma vibração de estiramento $\text{C}=\text{O}$.

IV.1.2 - MÓDULO INTERPRETADOR

Uma vez que os dados espectrais são extraídos pelo READMOD, o módulo interpretador denominado INTERMOD pode requisitá-los, para proceder a identificação dos grupos funcionais presentes num determinado composto orgânico. A forma de atuação desse módulo pode ser observada na figura IV-12.

A base de conhecimento do INTERMOD é constituída da implementação das regras de produção e da base de dados interna "*dba_bands*", contendo os fatos provenientes do READMOD, que foram salvos no arquivo DOS "reading.lnk" (como o formato, a largura, a intensidade, a posição e os pontos que formam uma banda de absorção) e, os fatos da "*dba_links*", existentes no arquivo DOS "dataiv.dba". Essa base de conhecimento foi elaborada através das tabelas de correlações de alguns livros sobre espectroscopia no I.V.[104-106]. Ela foi implementada com predicados contendo os seus argumentos (ou objetos), que são fatos conhecidos, da seguinte maneira:

% Base de Conhecimento "*dba_links*"

faixa_absorcao(Modo_Vibracional,[Faixa_Absorcao]).

vibacoes_caracteristicas(Grupo_Funcional,[Vibacoes_Caracteristicas]).

posicao_tipo_vibracao(Grupo_Funcional,Vibracao_Caracteristica,[Faixa_Absorcao],
intensidade([" "]),largura([" "])).

posicao_tipo_especifico(Grupo_Funcional,Vibracao_Caracteristica,[Faixa_Absorcao],
Tipo_Specifico).

ausente_outro_grupo(Grupo_Funcional,[Lista_Outros_Grupos]),

vibacoes_caracteristicas_outro_grupo(Grupo_Funcional,[Vibacoes_Caracteristicas]).

tipo_vibracao_outro_grupo(Grupo_Funcional,Vibracao_Caracteristica,[Faixa_Absorcao],
intensidade([" "]),largura([" "])).

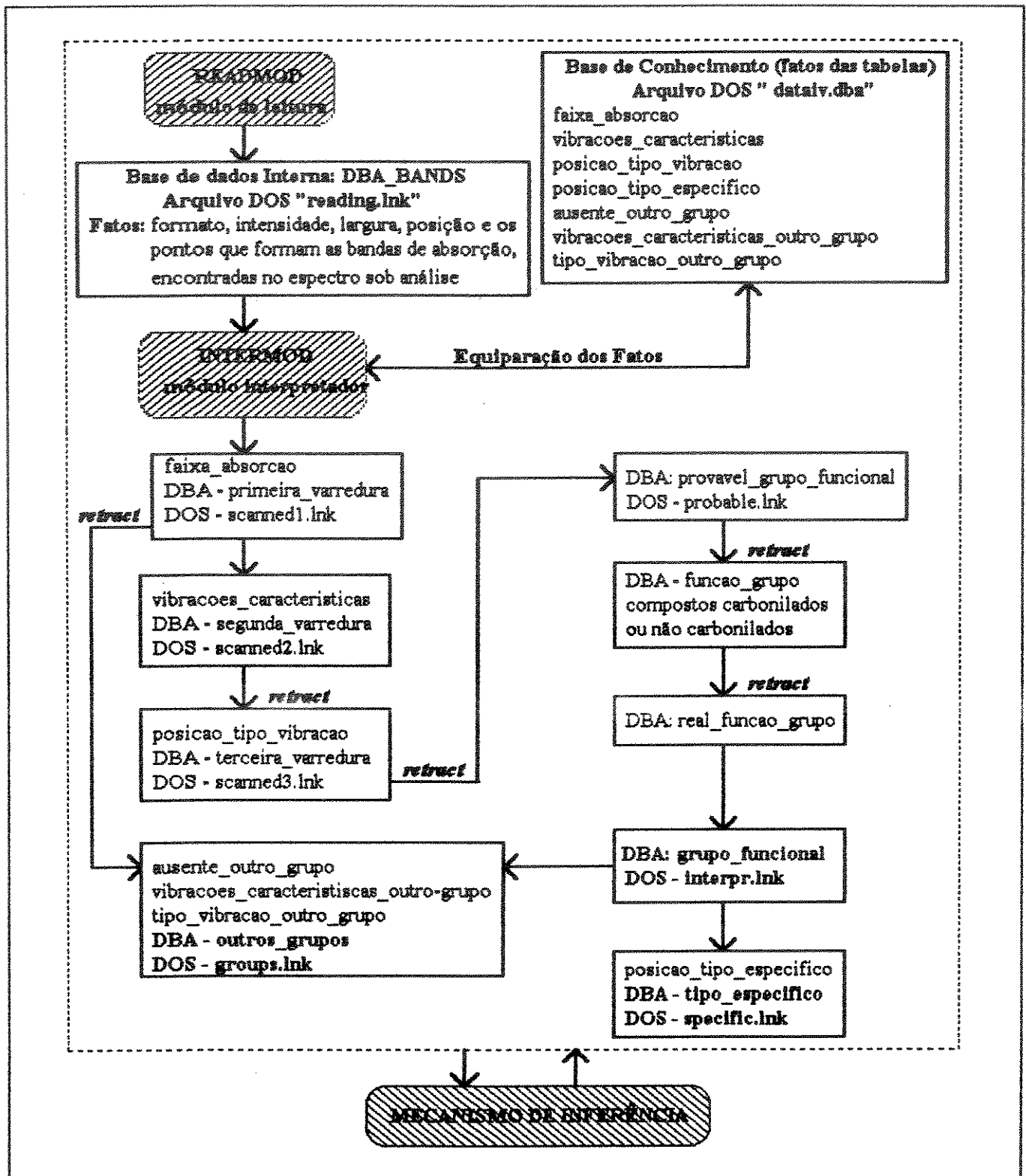


FIGURA IV-12 - Componentes do módulo interpretador INTERMOD, para a identificação dos grupos funcionais, tipos específicos e outros grupos mais prováveis de estarem presentes em um composto orgânico. O predicado padrão "retract" é utilizado para a remoção de fatos de uma base de dados interna.

A implementação desses predicados na base de conhecimento "*dba_links*" foi baseada nas informações dos livros sobre espectroscopia no I.V.[104,105], principalmente o livro de Nakanishi[106]. O predicado "*faixa_absorcao*" tem dois argumentos, o modo vibracional (uma string) e a sua faixa máxima de absorção (uma lista de números inteiros). Nesse caso, a faixa de absorção é independente do grupo funcional da molécula, por exemplo a vibração de estiramento O-H pode ser localizada no espectro de I.V. na região de 3640 a 2500 cm^{-1} para os grupos funcionais ácido carboxílico, álcool e fenol. O mesmo tipo de argumentos é utilizado pelo predicado "*vibracoes_caracteristicas*", mas relacionando o grupo funcional com as vibrações que o caracteriza.

No caso do predicado "*posicao_tipo_vibracao*" são utilizados cinco argumentos de modo que, os três últimos estão relacionados com a posição, a intensidade e a largura da banda de absorção de uma vibração característica de determinado grupo funcional. Conforme as informações da literatura citada a posição de uma banda de absorção pode ser encontrada com um desvio de cerca de $\pm 10 \text{ cm}^{-1}$. Dessa forma, estabeleceu-se para a maioria das posições o referido desvio, como por exemplo a banda da carbonila de uma cetona normal encontra-se em 1715 cm^{-1} portanto, a faixa de absorção foi imposta de 1725 a 1705 cm^{-1} . Nem sempre são fornecidas pelos livros dados referentes a largura e a intensidade de uma banda de absorção. Alguns critérios determinísticos foram estabelecidos inicialmente para classificar as bandas encontradas na base de dados interna "*dba_bands*", com relação a sua largura e intensidade. A tabela IV-1 relaciona os valores atribuídos para esta equiparação.

O mesmo procedimento do predicado acima mencionado, foi adotado para o predicado "*posicao_tipo_especifico*", contendo os fatos dos grupos funcionais de tipos específicos associados ao grupo funcional, como por exemplo a banda da carbonila para uma cetona α - β -insaturada posicionando-se em 1690 cm^{-1} , de modo que a sua faixa de absorção é inserida no seguinte formato de lista: [1700,1680].

O predicado "ausente_outro_grupo" é para descartar certos grupos ligados na molécula quando estes já foram identificados anteriormente como um dos grupos funcionais mais prováveis. Por exemplo, se um grupo aromático foi obtido nessas condições, isto é, identificado como um grupo funcional, ele não será analisado como um outro grupo da molécula.

Os predicados "vibrações_características_outro_grupo" e "tipo_vibração_outro_grupo" são requisitados no programa para a busca de outros grupos possíveis de estarem presentes ao composto, como um grupo alceno, nitro, tio, halogênio, amino, aromático, etc., ligados em um grupo cetônico. A figura IV-13 apresenta uma parte da base de conhecimento "dba_links", para o grupo funcional ácido carboxílico.

TABELA IV-1 - Relação dos valores adotados para a classificação das larguras e intensidades das bandas de absorção para os espectro de I.V..

LARGURAS (cm^{-1})		INTENSIDADES (% T)
Acima da região de 2000	Abaixo da região de 2000	muito fraca > 85 e <= 100
aguda ≥ 0 e <= 30	aguda <= 15	fraca > 60 e <= 85
média > 75 e <= 150	média > 15 e <= 75	média > 35 e <= 60
larga > 150	larga > 75	forte > 15 e <= 35
		muito forte ≥ 0 e <= 15

```
faixa_absorcao("estiramento_C=O",[1900,1600]).
faixa_absorcao("estiramento_C-O",[1440,1010]).
faixa_absorcao("estiramento_O-H",[3640,2500]).
vibracoes_caracteristicas("acido_carboxilico_(-COOH)","estiramento_C=O",
"estiramento_O-H","estiramento_C-O").
posicao_tipo_vibracao("acido_carboxilico_(-COOH)","estiramento_C=O",[1770,1700],
intensidades(["média","forte","muito_forte"]),largura([])).
posicao_tipo_vibracao("acido_carboxilico_(-COOH)","estiramento_O-H",[3550,2500],
intensidades(["média","forte","muito_forte"]),largura(["larga"])).
posicao_tipo_vibracao("acido_carboxilico_(-COOH)","estiramento_C-O",[1440,1200],
intensidades(["média","forte","muito_forte"]),largura([])).
posicao_tipo_especifico("acido_carboxilico_(-
COOH)","estiramento_C=O",[1820,1800],"(Alfa_F)_monomero").
posicao_tipo_especifico("acido_carboxilico_(-COOH)","estiramento_C=O",[1780,1760],
"(Alfa_Br)_or_(Alfa_Cl)_monomero").
posicao_tipo_especifico("acido_carboxilico_(-COOH)","estiramento_C=O",[1770,1750],
"(Alfa_F)_dimer").
posicao_tipo_especifico("acido_carboxilico_(-COOH)","estiramento_C=O",[1730,1700],
"(Alfa_Br)_or_(Alfa_Cl)_dimer").
posicao_tipo_especifico("acido_carboxilico_(-COOH)","estiramento_C=O",[1730,1710],
"(C=C-COOH)_monomero").
posicao_tipo_especifico("acido_carboxilico_(-COOH)","estiramento_C=O",[1700,1680],
"(Ar-COOH)_dimer").
posicao_tipo_especifico("acido_carboxilico_(-COOH)","estiramento_O-H",[3000,2500],
"(-COOH)_dimer").
posicao_tipo_especifico("acido_carboxilico_(-COOH)","estiramento_O-H",[3550,3000],
"(-COOH)_monomero").
posicao_tipo_especifico("acido_carboxilico_(-COOH)","estiramento_C-O",[1440,1200],
"(dimer)").
```

FIGURA IV-13 - Parte do arquivo DOS "dataiv.dba" onde estão inseridos os fatos referentes a base de conhecimento "dba_links", apenas para o grupo funcional ácido carboxílico.

O programa atua sobre a base de dados interna "*dba_bands*" contendo os fatos obtidos do espectro sob análise, que são equiparados com os fatos da base de conhecimento "*dba_links*", baseando-se sempre nos modos vibracionais que caracterizam um certo grupo funcional. Esta equiparação é efetuada através das regras implementadas, que descartam os grupos funcionais improváveis de estarem presentes na molécula, retendo aqueles que apresentam uma alta probabilidade de estarem presentes na molécula. Atualmente, o "INTERMOD" contém cerca de 100 regras de produção para alcançar os seus objetivos e em sua forma executável ocupa 110 Kbytes de memória.

Inicialmente, o "INTERMOD" faz as leituras através do predicado padrão "*consult*" nos arquivos DOS "*reading.lnk*" e "*dataiv.dba*" contendo os fatos declarados, inserindo-os na seção da base de dados apropriada nesse caso, "*dba_bands*" e "*dba_links*", respectivamente. Por exemplo, *consult("reading.lnk",dba_bands)*, onde *dba_bands* é o nome da base de dados interna contendo os fatos do arquivo DOS "*reading.lnk*".

Em seguida, para buscar as vibrações de um composto através dos seus dados espectrais, foi implementada uma regra com o predicado "*vibracoes_do_composto*" que liga os fatos do predicado "*faixa_absorcao*", existente na base de conhecimento "*dba_links*", retornando apenas as bandas de absorção que estão dentro da faixa máxima de absorção permitida para um certo modo vibracional (por exemplo, vibração de estiramento O-H, vibração de deformação C-H). Esta regra atua para todas as bandas de absorção existentes na base de dados interna "*dba_bands*". Ela foi implementada no programa da seguinte forma:

% Nessa regra consta: um predicado da "dba_links" com seus fatos, duas sub-metas e o predicado padrão "assertz".

analisa_vibracoes_do_espectro :-

```

    absorption_range(Tipo_Vibracao,FaixaAbsorcao),           % predicado da dba_links
    vibracoes_do_composto(Tipo_Vibracao,FaixaAbsorcao,max_wavenum,BandsList),
    confirma_vibracoes(BandsList,YesNo),
    assertz(primeira_varredura(vibracao_espectro([Tipo_Vibracao,YesNo]),
    faixa_abasorcao(FaixaAbsorcao),BandsList)), !.

```

% quando as condições da regra tornam-se verdadeiras os fatos da "dba_bands" ligam-se à regra como dados de saída

```

vibracoes_do_composto(Tipo_Vibracao,[WavenumInit,WavenumEnd],GreaterPosition,
    [banda(BandShape,Intensity,band_position(BandPosition))|BandList]) :- % dados de saída
    banda(banda(_BandShape,Intensity,band_position(BandPosition),_)), %fatos da "dba_bands"
    BandPosition < GreaterPosition, % GreaterPosition inicia com 4500 cm-1
    BandPosition <= WavenumInit,
    BandPosition >= WavenumEnd,
    vibracoes_do_composto(Tipo_Vibracao,[WavenumInit,WavenumEnd],BandPosition,BandList), !.

```

Os modos vibracionais inseridos na base de conhecimento "dba_links" são os que caracterizam a presença de determinados grupos funcionais do composto em análise. Nesse caso, a regra não faz nenhuma distinção quanto a largura e a intensidade da banda de absorção, requerendo apenas que esta se localize dentro das faixas limites de absorção originada de um modo vibracional. No entanto, algumas faixas de absorção podem sobrepor a outras e como consequência, uma mesma banda de absorção pode ser classificada como originada por diferentes ligações contendo diferentes elementos. Por exemplo, dependendo da localização da banda de vibração de estiramento O-H, esta pode ser confundida com as vibrações de estiramento das ligações N-H e C-H.

Os fatos obtidos com a meta acima citada são inseridos em tempo de execução do INTERMOD na base de dados interna "primeira_varredura", através do predicado padrão "assertz" e salvos no arquivo DOS "scanned1.lnk". A figura IV-14 mostra uma parte desse tipo de arquivo, referente ao espectro de I.V. do ácido 3-bromopropiônico.

Aqui, o que se pretende esclarecer é o fato da ocorrência ou não de certas vibrações no composto, como por exemplo, as vibrações de estiramento das ligações C=O, C=C de alcenos, C=N (nitrilas), C≡C (alcinos) e C-H de grupos aromáticos. Também, pode-se observar que a banda posicionada em 1708 cm^{-1} encontrada pode servir para as vibrações de estiramento C=O e C=C de alcenos. Entretanto, não foram encontradas no espectro nenhuma banda de absorção correspondente as vibrações de estiramento C=N, C≡C e C-H de grupo aromático. Isso significa que esses grupos funcionais estão ausentes do composto e serão excluídos dos próximos procedimentos da interpretação.

Se a primeira meta do INTERMOD foi bem sucedida, prossegue-se com outras regras, agora ligando a cada grupo funcional às vibrações que o caracteriza, utilizando na base de conhecimento "*dba_links*", os fatos do predicado "*vibracoes_caracteristicas*". Por exemplo, as vibrações de estiramento O-H, C=O e C-O caracterizam o grupo funcional ácido carboxílico. A regra foi implementada com três sub-metas da seguinte maneira:

```
% busca na "dba_links" as vibrações características para cada grupo funcional em análise
```

```
verificar_vibracoes_caracteristicas :-
```

```
vibracoes_caracteristicas(GrupoFuncional,ListaVibracoes),           % predicado da "dba_links"
```

```
verificar_estiramento_CX(Grupo_Funcional,["estiramento_C-F","estiramento_C-Cl",
```

```
  "estiramento_C-Br","estiramento_C-I"],Lista_Vibracoes,Vibracoes_Caracteristicas),
```

```
confirma_vibracoes(Grupo_Funcional,VibracoesCaracteristicas,ListaVibracoesCaracteristicas),
```

```
melhor_banda_se_anidrido(Case,GrupoFuncional,ListaVibracoesCaracteristicas,
```

```
  RealListaVibracoes),
```

```
  assertz(segunda_varredura(GrupoFuncional,RealListaVibracoes)).
```

```
fail, !.
```

```
% predicado fail, para forçar o retrocesso
```

As vibrações de estiramento C=O e C-X em -CO-X, onde X pode ser F, Cl, Br e I, caracterizam o grupo funcional haleto de acila. A primeira sub-meta da cláusula acima apresentada `verificar_estiramento_CX` foi imposta para verificar uma possível presença do haleto de acila, buscando os fatos existentes na base de dados interna "`primeira_varredura`", que equiparam com os fatos das vibrações de estiramento C-X, da base de conhecimento "`dba_links`", de forma que as vibrações que forem bem sucedidas, são colocadas na lista de vibrações características desse grupo funcional. Em seguida, procede-se igualmente com a segunda sub-meta `confirma_vibracoes`, mas para os outros grupos funcionais utilizados neste trabalho.

```
primeira_varredura(modo_vibracional(["estiramento_C=O","SIM"]),
faixa_absorcao([1900,1600]),[detalhes_banda(largura(1741,1676,65,"cm-1"),
intensidade(9,"%T"),posicao(1708))])
primeira_varredura(modo_vibracional(["estiramento_C=C_insaturado","SIM"]),
faixa_absorcao([1790,1566]),[detalhes_banda(largura(1741,1676,65,"cm-1"),
intensidade(9,"%T"),posicao(1708))])
primeira_varredura(modo_vibracional(["estiramento_C-H_aromático","NÃO"]),
faixa_absorcao([3100,3000]),[])
primeira_varredura(modo_vibracional(["estiramento_C=N","NÃO"]),
faixa_absorcao([2260,2050]),[])
primeira_varredura(modo_vibracional(["estiramento_C=C","NÃO"]),
faixa_absorcao([2260,2100]),[])
```

FIGURA IV-14 - Parte do arquivo DOS "scanned1.lnk" relacionando as bandas de absorção cujas análises foram bem sucedidas para as vibrações de estiramento C=O e C=C, mas não para as vibrações de estiramento C-H de aromático, C=N (nitrila) e C=C (alcino), referente a interpretação do espectro de I.V. do ácido 3-bromopropiônico.

A terceira sub-meta *melhor_banda_se_anidrido* refere-se ao grupo funcional anidrido de ácido, que tem como característica, a ocorrência de duas vibrações de estiramento C=O e uma de C-O. Essa regra fundamenta-se na busca das duas melhores bandas de absorção, relacionadas às vibrações de estiramento da carbonila na região do espectro entre 1900 a 1600 cm^{-1} . Aqui, determinou-se que as bandas mais intensas são as ocorrências mais prováveis dessas vibrações para o composto em análise.

Portanto, para se chegar a conclusão da segunda meta do INTERMOD deve-se confirmar a presença de todas as vibrações características de um determinado grupo funcional. Todavia, se nos fatos inseridos na base de dados interna "*primeira_varredura*", não são encontradas qualquer banda de absorção relacionada a estas vibrações, o grupo funcional em questão será descartado do resto da interpretação do espectro. Ao contrário, os grupos funcionais bem sucedidos são inseridos durante a execução, na base de dados interna "*segunda_varredura*" (fatos: grupo funcional com suas vibrações características) e salvos no arquivo DOS "scanned2.lnk". Até este momento da interpretação, não foram consideradas a intensidade (muito forte, forte, média e fraca) e a largura (larga, média ou aguda) de cada banda de absorção, mas apenas a posição da mesma, que deve estar dentro da sua faixa de frequência mais ampla permitida para o grupo funcional.

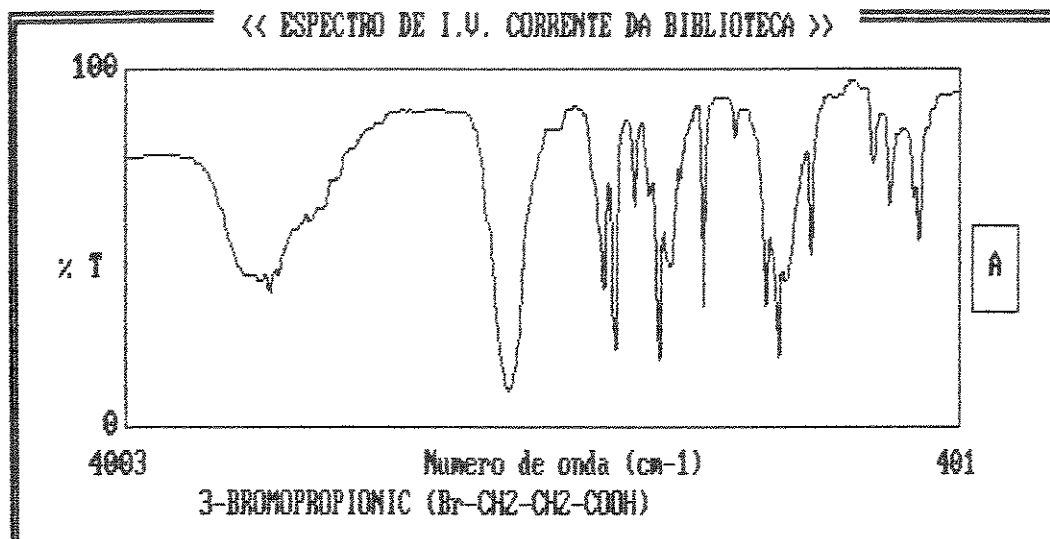
A tabela IV-2 relaciona os grupos funcionais utilizados neste trabalho com as suas respectivas vibrações características. A figura IV-15 ilustra o espectro de I.V. do ácido 3-bromopropiônico constando abaixo, uma parte do arquivo DOS "scanned2.lnk" apenas com os fatos para o grupo funcional ácido carboxílico.

Em geral, os fatos até aqui obtidos são os grupos funcionais com suas vibrações características, de modo que, dentro da faixa de absorção de uma certa vibração, podem constar uma ou mais bandas de absorção (veja figura IV-15, as bandas de absorção para as vibrações de estiramento O-H e C-O). Entretanto, necessita-se determinar qual ou quais bandas de absorção são as mais prováveis, para a vibração característica em análise, equiparando os fatos existentes na base de dados interna "*segunda_varredura*" com os da

base de conhecimento "*dba_links*", através do predicado "*posicao_tipo_vibracao*". Nesse caso, efetua-se uma análise completa, verificando a posição, a intensidade e a largura de cada banda de absorção. As regras implementadas quando sucedidas, inserem os fatos durante a execução na base de dados interna "*terceira_varredura*", agora, com um "*score*" para cada banda de absorção analisada e, conseqüentemente, para o grupo funcional em questão. O "*score*" é uma contagem de valores atribuídos, conforme as características observadas de cada banda de absorção no espectro.

TABELA IV-2 - Relação dos grupos funcionais com suas vibrações características, utilizados neste trabalho. ν = vibração de estiramento e δ = vibração de deformação.

GRUPOS FUNCIONAIS	VIBRAÇÕES CARACTERÍSTICAS
1. anidrido de ácido (-CO-O-CO)	ν C=O, ν C=O, ν C-O
2. ácido carboxílico (-COOH)	ν C=O, ν O-H, ν C-O
3. éster (-CO-O-)	ν C=O, ν C-O
4. haleto de acila (-CO-X)	ν C=O, ν C-X
5. cetona (-CO-)	ν C=O
6. aldeído (-CHO)	ν C=O, ν C-H alifático saturado, sobretudo δ C-H
7. álcool (-ROH)	ν O-H, ν C-O
8. fenol (-phOH)	ν O-H, ν C-O
9. aromático (ph)	ν C=C, ν C-H de grupo aromático
10. alceno (C_nH_{2n})	ν C=C, ν C-H de alifático insaturado
11. alcino (C_nH_{2n-2})	ν C≡C
12. nitrila (-C≡N)	ν C≡N
13. alceno (C_nH_{2n+2})	ν C-H assimétrico e ν C-H simétrico
14. éter (R-O-R)	ν C-O



```
segunda_varredura("acido_carboxilico_(-COOH)",
[vibracao(["estiramento_C=O","SIM"],
[banda(largura(1741,1676,65,"cm-1"),intensidade(9,"%T"),posicao(1708))]],
vibracao(["estiramento_O-H","SIM"],
[banda(largura(2997,2995,2,"cm-1"),intensidade(37,"%T"),posicao(2989)),
banda(largura(3314,2744,570,"cm-1"),intensidade(37,"%T"),posicao(2982)),
banda(largura(2958,2921,37,"cm-1"),intensidade(42,"%T"),posicao(2928))]],
vibracao(["estiramento_C-O","SIM"],
[banda(largura(1445,1429,16,"cm-1"),intensidade(39,"%T"),posicao(1435)),
banda(largura(1408,1395,13,"cm-1"),intensidade(21,"%T"),posicao(1400)),
banda(largura(1348,1341,7,"cm-1"),intensidade(62,"%T"),posicao(1345)),
banda(largura(1311,1301,10,"cm-1"),intensidade(65,"%T"),posicao(1306)),
banda(largura(1278,1268,10,"cm-1"),intensidade(18,"%T"),posicao(1273)),
banda(largura(1252,1232,20,"cm-1"),intensidade(45,"%T"),posicao(1241)),
banda(largura(1213,1209,4,"cm-1"),intensidade(69,"%T"),posicao(1213)),
banda(largura(1145,1141,4,"cm-1"),intensidade(34,"%T"),posicao(1143)),
banda(largura(1056,1048,8,"cm-1"),intensidade(81,"%T"),posicao(1052))]]))
```

FIGURA IV-15 - Espectro de I.V. do ácido 3-bromopropiônico e parte do arquivo DOS "scanned2.lnk" da base de dados interna "segunda_varredura", relacionando o grupo funcional ácido carboxílico com suas vibrações características. A busca das bandas de absorção foram efetuadas na base de dados interna "primeira_varredura".

As regras devem ser concluídas se forem bem sucedidas as seguintes sub-metas:

a) determinar a banda de absorção que se localiza na faixa mais provável do espectro, para aquela vibração característica, b) calcular o "score" de cada banda de absorção, c) buscar uma ou mais bandas com "scores" maiores, d) através do predicado padrão "assertz" inserir os fatos obtidos na base de dados interna "terceira_varredura" e salvar no arquivo DOS "scanned3.lnk" e, e) através do predicado padrão "retract" remover todos fatos existentes na base de dados interna "segunda_varredura". Durante a execução do INTERMOD os fatos inseridos em uma base de dados interna podem ser removidos, caso não haja mais necessidade de utilizá-los.

A determinação da melhor faixa de absorção para uma vibração característica de uma função orgânica, baseia-se no fato que ela deve normalmente posicionar-se em uma mesma região do espectro, ou às vezes, podendo situar-se fora dessa região, devido a outros grupos envolvidos na molécula. Por exemplo, a faixa de absorção para a banda da carbonila de um ácido carboxílico, posiciona-se na maioria dos espectros entre 1760 a 1710 cm^{-1} . No entanto, para um ácido carboxílico com um grupo α -Br ou um α -Cl, a mesma encontra-se entre 1780 e 1760 cm^{-1} .

Quando uma banda de absorção está localizada dentro da sua faixa normal de ocorrência no espectro, atribui-se um "score" igual a 100. Caso a regra aplicada não encontre uma solução para o problema, a mesma retrocede (backtracking), ou seja, procura na outra cláusula do mesmo predicado, uma alternativa para tal resolução, dentro das faixas de absorção específicas para o grupo funcional em análise. Se esta meta é bem sucedida, atribui-se também o mesmo score para a respectiva banda de absorção. Porém, há casos em que a busca não sucede nem para as faixas de absorção específicas. Desta forma, aplica-se uma regra com uma função matemática, que descreve o comportamento da posição da banda de absorção. Esta regra foi implementada para que a tomada de decisão do INTERMOD ao resolver este problema, não seja puramente determinística, isto é, se a banda de absorção não está posicionada na região espectral esperada, calcula-se o "score"

proporcional ao seu deslocamento. Este procedimento nos condiciona em não descartar a hipótese da banda de absorção pertencer ao grupo funcional em análise, mas, sim diminuir gradativamente o seu "score" a medida que a banda de absorção se distancie da posição ideal. As regras para buscar as informações da posição de uma banda de absorção, foram impostas da seguinte maneira:

```

posicao_difusa(Tipo_Vibracao,Posicao,[WavenumInit,WavenumEnd],Score,Desvio,DeltaMenor) :-
    DeltaEsquerda = (abs ( WavenumInit - Posicao )),
    DeltaDireita  = (abs ( WavenumEnd - Posicao )),
    encontre_menor_lado_desvio([DeltaEsquerda,DeltaDireita],Desvio,DeltaMenor),
    funcao_difusa_posicao(Tipo_Vibracao,DeltaMenor,Score), !.           % o cut (!) evita retrocesso

% A função matemática aplicada é { 1 / [ exp ( alfa * (X) ) ] }
funcao_difusa_posicao(Tipo_Vibracao,DeltaMenor,Score) :-
    verifica_tipo_vibracao(Tipo_Vibracao,X),
    AlfaX = ( X * DeltaMenor ), Exponencial = (exp ( AlfaX ) ),
    Valor_Exponencial = ( ( 1 / Exponencial ) * 100 ),
    Score = ( round ( Valor_Exponencial ) ).

```

Na cláusula acima descrita, após o cálculo dos deslocamentos dos lados direito e esquerdo, relacionados com a posição da banda de absorção, a primeira sub-meta `encontre_menor_lado_desvio` busca qual o lado de menor deslocamento (em cm^{-1}) e, em seguida começa atuar a segunda sub-meta `função_posicao_difusa`, que verifica o tipo de vibração que está sendo analisada e calcula o "score" da banda de absorção, através da aplicação da função matemática: $[1 / (e^{\alpha \cdot x}) * 100]$ onde, e = exponencial, α = desvio da posição e x = constante (depende do tipo de vibração)

A variável "X" é um valor constante, dependente do tipo de vibração, para que a inclinação da curva referente à função matemática que calcula os desvios, não seja de

maneira abrupta. Por exemplo, para as vibrações de estiramento C=O e O-H, a constante é igual a 0.035, mas foi atribuído o valor de 0.05, para as demais vibrações.

A figura IV-16 ilustra um gráfico mostrando o comportamento das posições da banda de carbonila extraídas de 52 espectros de I.V. de ácidos carboxílicos da biblioteca SADTLER. A região do espectro em que essa vibração é normalmente encontrada, está entre 1720 e 1686 cm^{-1} no entanto, a faixa máxima de absorção para essa ocorrência encontra-se entre 1745 e 1660 cm^{-1} . Apenas quando a posição da banda de carbonila em análise está localizada fora desse limite, é aplicada a regra com a referida função matemática, encarregada de calcular o score proporcional ao seu desvio, caso contrário, este "score" é igual a 100.

Para a intensidade e a largura da banda de absorção, primeiro verifica-se a existência ou não, de alguma informação existente na base de conhecimento "dba_links", através do predicado "posicao_tipo_vibracao", implementado do seguinte modo:

```
posicao_tipo_vibracao(Grupo_Funcional,Vibracao_Caracteristica,[Faixa_Absorcao],  
intensidade([X]),largura([Y])).
```

onde, as variáveis X e Y representam as listas de informações para a intensidade e largura, respectivamente, como por exemplo:

```
intensidade(["muito_forte","forte"]),largura(["larga","aguda"]).
```

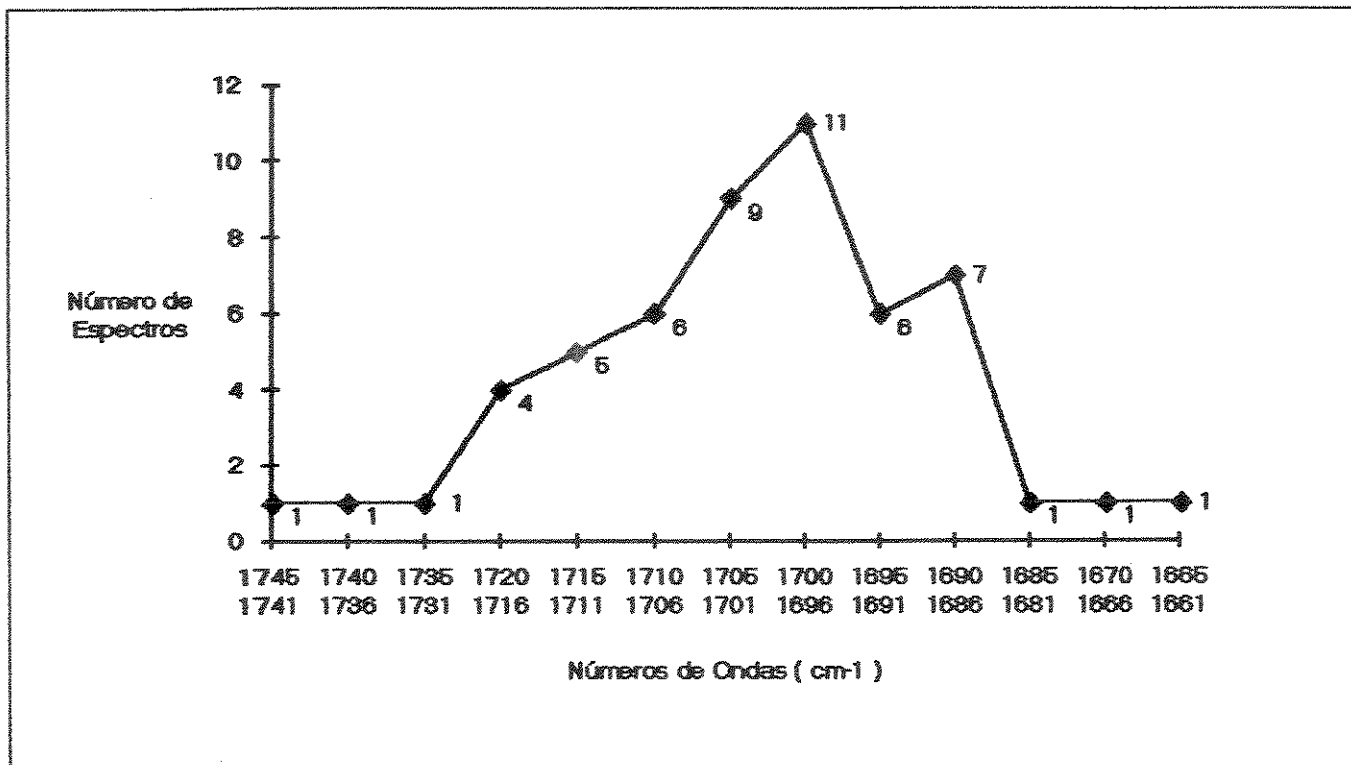


FIGURA IV-16 - Comportamento das posições da banda de carbonila extraídas de 52 espectros de I.V. de ácidos carboxílicos. A faixa de absorção onde essa banda mais ocorre, encontra-se entre 1720 a 1686 cm^{-1} .

Os argumentos para a intensidade e largura foram determinados com base nos limites da porcentagem de transmitância (%T) e da faixa de absorção em números de onda (cm^{-1}) respectivamente, como relacionados na tabela IV-1. Se ambas são bem sucedidas, isto é, encontram-se dentro dos limites estipulados, seus "scores" são iguais a 100. Ao contrário, aplica-se as regras com uma função matemática, agora direcionadas para os cálculos dos "scores" para a intensidade e a largura da banda de absorção.

Com a determinação dos "scores" para a posição, a intensidade e a largura de uma banda de absorção, necessita-se calcular o seu "score" médio. Normalmente, a posição de uma banda de absorção pode ser encontrada em qualquer tabela de correlação da literatura, porém os dados para a intensidade e largura nem sempre são disponíveis. Desta forma, implementou-se uma regra com as seguintes atribuições de pesos: a) posição

= 4, b) intensidade = 2 e, c) largura = 1. Por exemplo, se na base de conhecimento "dba_links" não constar informações, a respeito da intensidade de uma certa vibração, o "score" médio da banda de absorção é calculado, somando o "score" da posição vezes 4, mais o "score" da largura e divide por 5, como mostra as regras abaixo:

```

calcular_score_banda(ScorePosicao,ScoreIntensidade,ScoreLargura),ScoreBanda) :-
    ScorePosicao = (ScorePosicao * 4),
    ScoreIntensidade = (ScoreIntensidade * 2),
    Score = (ScorePosicao + ScoreIntensidade + ScoreLargura),
    verifica_N_score(ScoreIntensidade,ScoreLargura,Nsoma),
    ScoreBanda = (round (Score / Nsoma)).           % ScoreBanda é o valor do score calculado

verifica_N_score(0,0,4):-!.                       % dados não disponíveis para intensidade e largura (posição)
verifica_N_score(0,_,5) :-!.                       % dado não disponível para intensidade (posição e largura)
verifica_N_score(_,0,6) :-!.                       % dado não disponível para largura (posição e intensidade)
verifica_N_score(_,_,7) :-!.                       % dados disponíveis para posição, intensidade e largura

```

As figuras IV-17 e IV-18 ilustram os gráficos, que indicam os comportamentos das intensidades e larguras da banda de carbonila extraídas de 52 espectros de I.V. de ácidos carboxílicos. Através da figura IV-17 pode-se confirmar que a maioria dos ácidos carboxílicos verificados, apresentam uma intensidade menor ou igual a 20 %T, que é considerada pelo programa como uma intensidade forte e/ou muito_forte. Na figura IV-18 a largura calculada a $\frac{1}{4}$ da altura para a maioria dos ácidos carboxílicos, prevaleceram de 30 a 60 cm^{-1} , considerando-se como uma média. Essas informações estão coerentes com os dados das tabelas de correlações. Portanto, a posição, a intensidade e a largura das bandas de absorção, foram inseridas como argumentos na base de conhecimento "dba_links", através do predicado "posicao_tipo_vibracao".

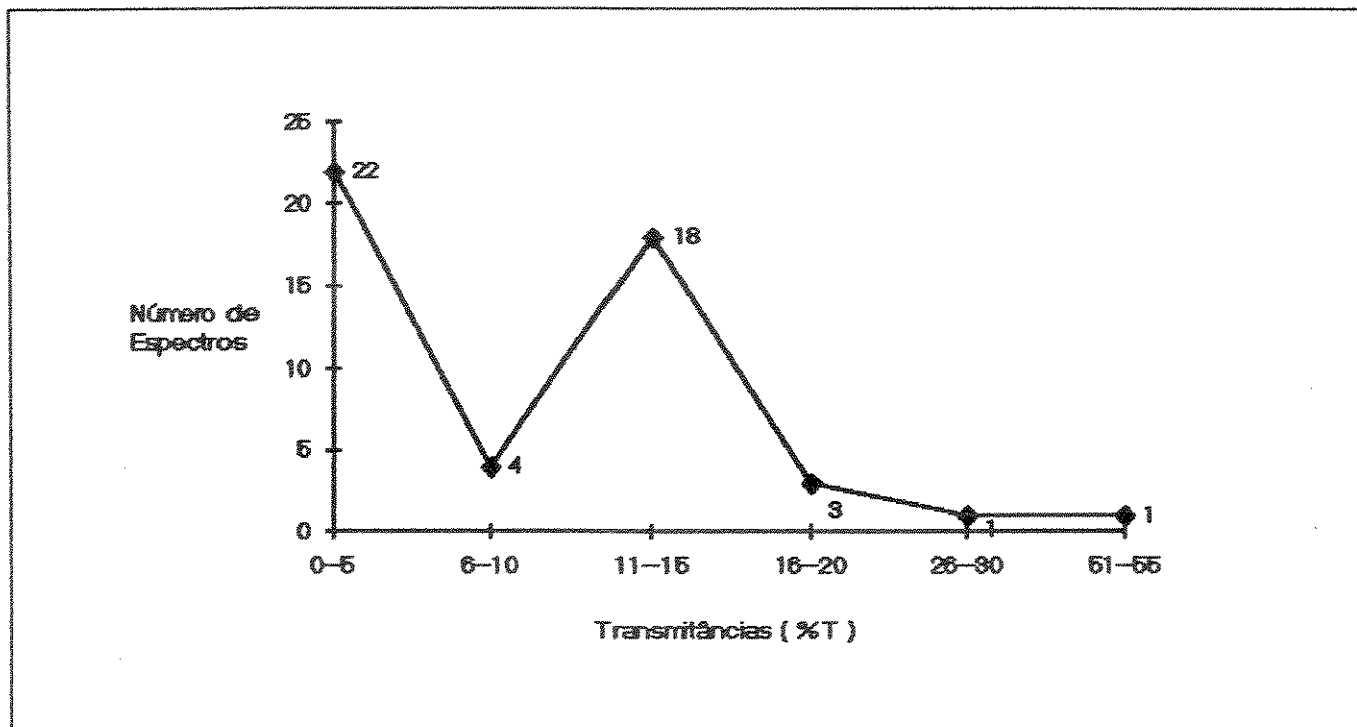


FIGURA IV-17 - Comportamento das intensidades (% T) das bandas de carbonila extraídas de 52 espectros de I.V. de ácidos carboxílicos. As bandas mais intensas são menores ou iguais a 20 %T, consideradas no programa como forte e muito forte.

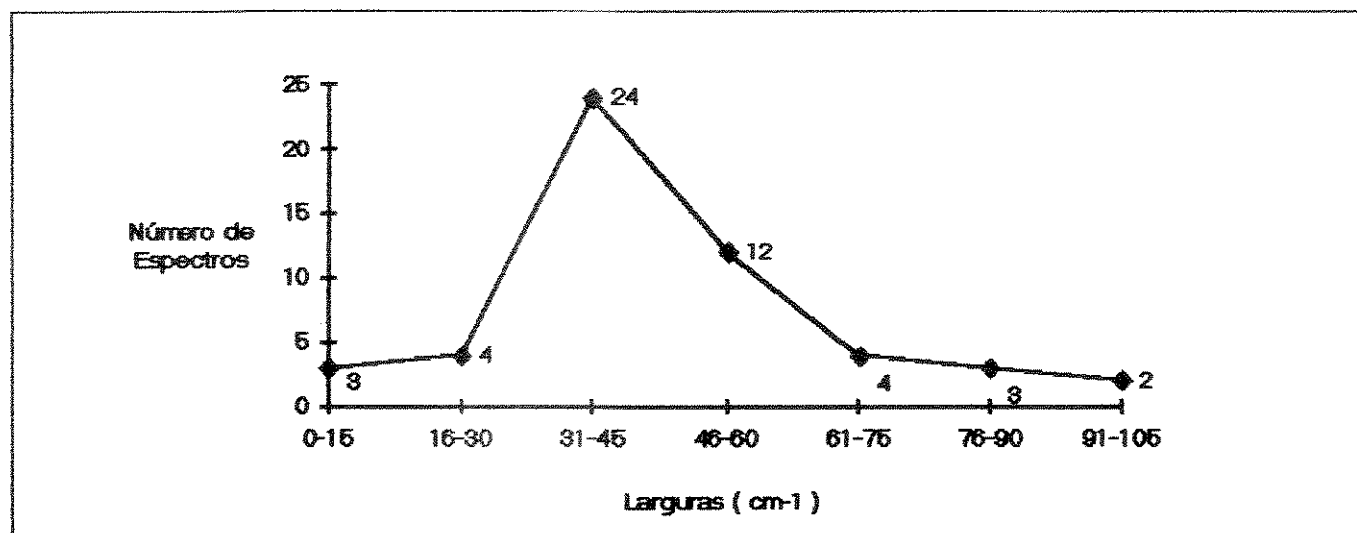


FIGURA IV-18 - Comportamento das larguras das bandas de carbonila extraídas de 52 espectros de I.V. de ácidos carboxílicos. A largura calculada a $\frac{1}{4}$ da altura, mostra que essas bandas de absorção estão entre 30 a 60 cm^{-1} e, foram consideradas no programa como média.

Quando as metas são bem sucedidas pelos procedimentos acima descritos, os fatos obtidos dos grupos funcionais contêm suas vibrações características, de modo que nessas vibrações podem constar uma ou mais bandas de absorção.

Em seguida, se o INTERMOD ainda não descartou a hipótese da existência do grupo funcional haleto de acila do composto em análise, então é necessário buscar qual ou quais bandas de absorção correspondentes as vibrações de estiramento C-X, apresentam os maiores "scores" calculados. Por exemplo, se foram encontradas vibrações de estiramento das ligações C-F e C-Cl, efetua-se separadamente, a busca da banda ou das bandas de absorção com maiores scores, eliminando as demais vibrações C-X.

Após a resolução desse problema, ou seja, confirmar ou não a presença do haleto de acila, é necessário calcular o "score" de cada grupo funcional que ainda não foi descartado pelo módulo interpretador. Então, calcula-se o "score" de cada vibração característica, que é o "score" médio calculado para as bandas de absorção, relacionadas com tal vibração. Depois, calcula-se o "score" do grupo funcional, que é a média dos "scores" das suas vibrações características. Se essa meta for bem sucedida, os fatos são inseridos durante a execução na base de dados interna "terceira_varredura", através do predicado padrão "assertz" e salvos no arquivo DOS "scanned3.lnk". A figura IV-19 mostra uma parte desse arquivo referente ao espectro do ácido 3-bromopropiônico, que relaciona os fatos obtidos até este momento da interpretação, isto é, o grupo funcional com suas vibrações características e os "scores" calculados para a posição, intensidade e largura de cada banda de absorção encontrada. O predicado padrão "fail" foi colocado para forçar o retrocesso, isto é, para consultar os fatos de todos os grupos funcionais, existentes na base de dados interna "segunda_varredura" que posteriormente, são removidos pelo predicado padrão "retract".


```

terceira_varredura("ácido_carboxílico_(-COOH)",
[tipo_vibricao("estiramento_C=O",faixa_absorcao([1770,1700]),
[banda(largura("media",1741,1676,65,"cm-1"),intensidade("muito_forte",9,"%T"),
posicao(1708),scores(score_posicao(100),score_intensidade(100),score_largura(0),
banda_score(100),desvios(desvio_posicao("Dentro",0,"cm-1"),
desvio_intensidade("Dentro",0,"%T"),desvio_largura("nao_disponivel",0,"cm-1")))]),
tipo_vibricao("estiramento_O-H",faixa_absorcao([3550,2500]),
[banda(largura("larga",3314,2744,570,"cm-1"),intensidade("media",37,"%T"),
posicao(2982),scores(score_posicao(100),score_intensidade(100),score_largura(100),
banda_score(100),desvios(desvio_posicao("Dentro",0,"cm-1"),
desvio_intensidade("Dentro",0,"%T"),desvio_largura("Dentro",0,"cm-1")))]),
tipo_vibricao("estiramento_C-O",faixa_absorcao([1440,1200]),
[banda(largura("media",1445,1429,16,"cm-1"),intensidade("media",39,"%T"),
posicao(1435),scores(score_posicao(100),score_intensidade(100),score_largura(0),
banda_score(100),desvios(desvio_posicao("Dentro",0,"cm-1"),
desvio_intensidade("Dentro",0,"%T"),desvio_largura("nao_disponivel",0,"cm-1")))],
banda(largura("aguda",1408,1395,13,"cm-1"),intensidade("forte",21,"%T"),
posicao(1400),scores(score_posicao(100),score_intensidade(100),score_largura(0),
banda_score(100),desvios(desvio_posicao("Dentro",0,"cm-1"),
desvio_intensidade("Dentro",0,"%T"),desvio_largura("nao_disponivel",0,"cm-1")))],
banda(largura("aguda",1278,1268,10,"cm-1"),intensidade("forte",18,"%T"),
posicao(1273),scores(score_posicao(100),score_intensidade(100),score_largura(0),
banda_score(100),desvios(desvio_posicao("Dentro",0,"cm-1"),
desvio_intensidade("Dentro",0,"%T"),desvio_largura("nao_disponivel",0,"cm-1")))],
banda(largura("media",1252,1232,20,"cm-1"),intensidade("media",45,"%T"),
posicao(1241),scores(score_posicao(100),score_intensidade(100),score_largura(0),
banda_score(100),desvios(desvio_posicao("Dentro",0,"cm-1"),
desvio_intensidade("Dentro",0,"%T"),desvio_largura("nao_disponivel",0,"cm-1")))]),
score_grupo(100))

```

FIGURA IV-19 - Parte do arquivo DOS "scanned3.lnk" relacionando os fatos inseridos durante a execução na base de dados interna "terceira_varredura", referente a interpretação do espectro de I.V. do ácido 3-bromopropiônico.

Até o momento, as três varreduras feitas no espectro de I.V. acima descrito, são necessárias para se chegar a conclusões cada vez mais restritas a respeito dos dados espectrais, ou seja, o INTERMOD atua inicialmente buscando as informações da base de

dados interna "dba_bands" oriundas do READMOD, para extrair de uma maneira simbólica, os fatos espectroscópicos. A figura IV-20 ilustra como é a atuação desse módulo para a extração destes fatos.

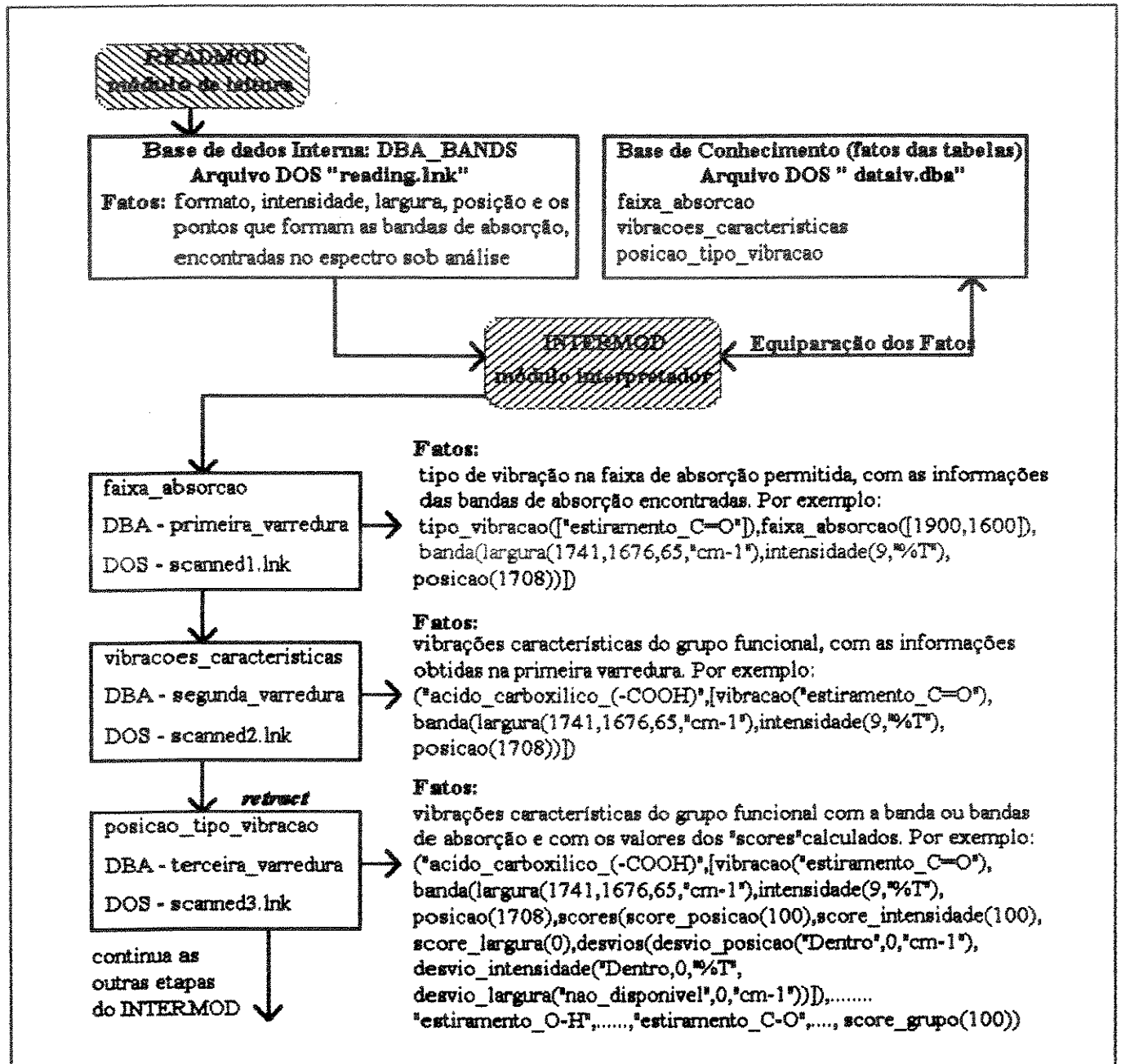


FIGURA IV-20 - Fatos espectroscópicos extraídos de maneira simbólica pelo INTERMOD, até a terceira varredura feita em um espectro de I.V.. O predicado padrão "retract" é utilizado para a remoção dos fatos da base de dados interna "segunda_varredura".

Através de uma avaliação nos espectros de compostos com diferentes grupos funcionais, foi tomada uma decisão com relação aos valores mínimos permitido para os seus "scores", de modo que o valores determinados para as vibrações características e para os grupos funcionais devem ser maiores ou iguais a 75 e 90, respectivamente. Com esta determinação, torna-se o INTERMOD mais rígido para filtrar qual ou quais grupos funcionais são mais prováveis de pertencerem ao composto em análise.

Nesse caso, inicialmente a regra implementada consulta os fatos da base de dados interna "*terceira_varredura*", verificando os valores dos "scores" para cada vibração característica e finalmente, o "score" do grupo funcional, de forma que os mesmos devem ser maiores ou iguais aos "scores" mínimos estipulados no programa. Ao contrário, descarta-se a possibilidade do grupo funcional em análise pertencer ao composto. No entanto, os fatos dos grupos funcionais bem sucedidos nessa meta, são os mesmos do arquivo DOS "*scanned3.lnk*" (veja exemplo da figura IV-19), portanto estes grupos funcionais podem ser considerados como os mais prováveis de estarem presentes na molécula. Desta forma, os fatos obtidos são inseridos, durante a execução, na base de dados interna "*provável_grupo_funcional*" e armazenados no arquivo DOS "*probable.lnk*". Também, foi aplicado o predicado padrão "*fail*" para forçar o retrocesso, isto é, verificar quais grupos funcionais apresentam os "scores" maiores ou iguais ao limites determinados. Então, a cada passagem do grupo funcional analisado, os fatos da "*terceira_varredura*" são removidos da sua base de dados interna, pelo predicado padrão "*retract*".

A figura IV-21 mostra a saída do arquivo DOS "*provável.lnk*" com os prováveis grupos funcionais obtidos na interpretação do espectro de I.V. do ácido 3-bromopropiônico. Através desse arquivo, observa-se que os fatos apresentam a hipótese de grupos funcionais mais prováveis de compostos carbonilados e não carbonilados.

```

provavel_grupo("acido_carboxilico_(-COOH)",[tipo_vibricao("estiramento_C=O",
faixa_absorcao([1770,1700]),
[banda(largura("media",1741,1676,65,"cm-1"),intensidade("muito_forte",9,"%T"),posicao(1708),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))]),
tipo_vibricao("estiramento_O-H",faixa_absorcao([3550,2500]),
[banda(largura("larga",3314,2744,570,"cm-1"),intensidade("media",37,"%T"),posicao(2982),
scores(posicao_score(100),score_intensidade(100),score_largura(100),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("Dentro",0,"cm-1")))]),
tipo_vibricao("estiramento_C-O",faixa_absorcao([1440,1200]),
[banda(largura("media",1445,1429,16,"cm-1"),intensidade("media",39,"%T"),posicao(1435),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))],
banda(largura("aguda",1408,1395,13,"cm-1"),intensidade("forte",21,"%T"),posicao(1400),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))],
banda(largura("aguda",1278,1268,10,"cm-1"),intensidade("forte",18,"%T"),posicao(1273),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))],
banda(largura("media",1252,1232,20,"cm-1"),intensidade("media",45,"%T"),posicao(1241),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))]),score_grupo(100))

provavel_grupo("ester_(-CO-O-)",[tipo_vibricao("estiramento_C=O",
faixa_absorcao([1730,1710]),
[banda(largura("media",1741,1676,65,"cm-1"),intensidade("muito_forte",9,"%T"),posicao(1708),
scores(posicao_score(94),score_intensidade(100),score_largura(100),banda_score(97),
desvios(desvio_posicao("DIREITA",2,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("Dentro",0,"cm-1")))]),
tipo_vibricao("estiramento_C-O",faixa_absorcao([1300,1050]),
[banda(largura("aguda",1278,1268,10,"cm-1"),intensidade("forte",18,"%T"),posicao(1273),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))]),score_grupo(99))

```

continua na próxima página (arquivo DOS "probable.lnk")

continuação (arquivo DOS "probable.lnk")

```
provavel_grupo("cetona_(-CO-)",[tipo_vibracao("estiramento_C=O",
faixa_absorcao([1725,1705]),
[banda(largura("media",1741,1676,65,"cm-1"),intensidade("muito_forte",9,"%T"),posicao(1708),
scores(posicao_score(100),score_intensidade(100),score_largura(100),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("Dentro",0,"cm-1")))]),
tipo_vibracao("estiramento_C-C",faixa_absorcao([1300,1100]),
[banda(largura("aguda",1278,1268,10,"cm-1"),intensidade("forte",18,"%T"),posicao(1273),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))],
banda(largura("media",1252,1232,20,"cm-1"),intensidade("media",45,"%T"),posicao(1241),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))],
banda(largura("aguda",1145,1141,4,"cm-1"),intensidade("media",34,"%T"),posicao(1143),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))]), score_grupo(100))
```

```
provavel_grupo("alcool_(R-OH)",[tipo_vibracao("estiramento_O-H",
faixa_absorcao([3200,2500]),
[banda(largura("larga",3314,2744,570,"cm-1"),intensidade("media",37,"%T"),posicao(2982),
scores(posicao_score(100),score_intensidade(100),score_largura(100),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("Dentro",0,"cm-1")))]),
tipo_vibracao("estiramento_C-O",faixa_absorcao([1260,1010]),
[banda(largura("media",1252,1232,20,"cm-1"),intensidade("media",45,"%T"),posicao(1241),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))],
banda(largura("aguda",1145,1141,4,"cm-1"),intensidade("media",34,"%T"),posicao(1143),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))]), score_grupo(100))
```

```
provavel_grupo("alcano_(CnH2n+2)",[tipo_vibracao("aliphatic_estiramento_C-H",
faixa_absorcao([2970,2850]),
[banda(largura("media",2958,2921,37,"cm-1"),intensidade("media",42,"%T"),posicao(2928),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))]),
```

continua na próxima página

continuação

```

tipo_vibricao("bending_C-H",faixa_absorcao([1460,1355]),
[banda(largura("media",1445,1429,16,"cm-1"),intensidade("media",39,"%T"),posicao(1435),
scores(posicao_score(100),score_intensidade(100),score_largura(97),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("Out",1,"cm-1")))),
banda(largura("aguda",1408,1395,13,"cm-1"),intensidade("forte",21,"%T"),posicao(1400),
scores(posicao_score(100),score_intensidade(100),score_largura(100),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("Dentro",0,"cm-1")))))]],score_grupo(100))

provavel_grupo("eter_(R-O-R)",[tipo_vibricao("estiramento_C-O",
faixa_absorcao([1275,1020]),
[banda(largura("aguda",1278,1268,10,"cm-1"),intensidade("forte",18,"%T"),posicao(1273),
scores(posicao_score(100),score_intensidade(100),score_largura(0),banda_score(100),
desvios(desvio_posicao("Dentro",0,"cm-1"),desvio_intensidade("Dentro",0,"%T"),
desvio_largura("nao_disponivel",0,"cm-1")))))]],score_grupo(100))

```

FIGURA IV-21 - Arquivo DOS "probable.lnk" relacionando os fatos (grupos funcionais mais prováveis) com as informações das bandas de absorção correspondentes as vibrações características, inseridos na base de dados interna "*provavel_grupo_funcional*", referente a interpretação do espectro de I.V. do ácido 3-bromopropiônico.

Como os espectros da biblioteca SADTLER utilizados neste trabalho são de compostos orgânicos puros de relativa complexidade, foi possível implementar regras para tomar uma decisão quanto a separação de compostos carbonilados dos não carbonilados. Este procedimento é similar ao de um usuário de espectroscopia no I.V., que durante a análise do espectro, observa a presença ou não da vibração de estiramento C=O, podendo descartar vários grupos funcionais que não ocorrem no composto sob análise. Desta forma, estas regras possibilitam apenas o retorno dos grupos funcionais carbonilados, que foram encontrados como mais prováveis de pertencerem ao composto sob análise.

Se nos fatos da base de dados interna "*provavel_grupo_funcional*" (por exemplo, figura IV-21) existir qualquer grupo funcional de composto carbonilado, descarta-

se os demais grupos funcionais não carbonilados. Esta separação é realizada porque a vibração de estiramento C=O prevalece no processo de filtragem do INTERMOD. Ao contrário, os grupos funcionais de compostos não carbonilados, ficam armazenados na base de dados interna "*funcao_grupo*". Neste caso, quando se atinge a meta para os grupos funcionais mais prováveis, normalmente, os carbonilados já foram descartados nos procedimentos anteriores, principalmente com relação a confirmação da presença da vibração de estiramento C=O.

Nas regras implementadas para se atingir tais metas (carbonilados ou não carbonilados), os fatos de cada grupo funcional que estão colocados na base de dados interna "*provavel_grupo_funcional*" são analisados recursivamente, verificando se apresentam ou não a vibração de estiramento C=O. Nesse caso, os fatos bem sucedidos são inseridos na base de dados interna "*funcao_grupo*" e os fatos da "*provavel_grupo_funcional*" são removidos, pelo predicado padrão "*retract*".

No caso da figura IV-21 os grupos funcionais que ficam armazenados durante a execução, na base de dados interna "*funcao_grupo*", são os grupos carbonilados (ácido carboxílico, éster e cetona), descartando-se os demais grupos não carbonilados (álcool, alcano e éter) em análise. O programa EXPERTIV está direcionado para identificar grupos funcionais de compostos orgânicos puros. A inclusão dos grupos funcionais éster e cetona é devida as posições das vibrações que os caracterizam, isto é, vibrações de estiramento C=O e C-O para o éster e, C=O e de C-C em -C-(CO)-C- para a cetona. Essas vibrações ocorreram na mesma região das vibrações de estiramento C=O e C-O do ácido 3-bromopropiônico.

Em seguida, foi implementada uma cláusula para verificar se um fenol foi identificado como um dos grupos funcionais mais prováveis do composto em análise. Além das vibrações que o caracteriza, os fatos da "*funcao_grupo*" são ligados por intermédio de uma regra, que procura na outra base de dados interna "*primeira_varredura*", ainda não removida, (veja exemplo na figura IV-14) a presença da

vibração de estiramento C=C de grupo aromático (região de absorção em 1650-1450 cm^{-1}). Se essa presença foi confirmada, aplica-se a função matemática apenas para a intensidade da banda de absorção, onde o "score" calculado deve ser maior ou igual a 90. Portanto, como a regra foi imposta de maneira recursiva, ou seja, procede-se para todos os grupos funcionais inseridos na "funcao_grupo", confirmando ou não a presença do grupo fenólico, os fatos da "funcao_grupo" são removidos a cada grupo funcional analisado pelo predicado padrão "retract" e, os fatos obtidos são inseridos na base de dados interna "real_funcao_grupo", pelo predicado padrão "assertz".

Quando o INTERMOD atinge até esta meta, pode-se concluir que os fatos obtidos são realmente do grupo ou dos grupos funcionais mais prováveis, pertencentes ao composto orgânico em análise. Desta forma, uma outra regra é imposta para abrir o arquivo DOS "possible.se", que relaciona em ordem, qual ou quais os grupos funcionais foram identificados com sucesso, com o seu "score" e o nome do respectivo arquivo DOS "possibl(X).se". Aqui, a variável X representa o número do grupo funcional em questão. Por exemplo, essa regra ao ser bem sucedida, na interpretação do espectro de I.V. do ácido 3-bromopropiônico (veja figura IV-21), cria o arquivo DOS "possible.se" da seguinte forma:

```
ácido_carboxílico,100,possibl1.se  
éster,99,possibl2.se  
cetona,100,possibl3.se
```

Os arquivos DOS "possibl(X).se" relacionam a vibração ou as vibrações características do grupo funcional, com a posição, a intensidade e a largura da banda ou bandas de absorção, respectivamente. No caso do exemplo acima citado, o arquivo DOS

"possibl1.se" relaciona-os da seguinte forma:

```
estiramento_C=O,1708,9,65  
estiramento_O-H,2982,37,570  
estiramento_C-O,1435,39,16  
estiramento_C-O,1400,21,13  
estiramento_C-O,1273,18,10  
estiramento_C-O,1241,45,20
```


Os arquivos DOS "possible.se" e "possibl(X).se" são utilizados como arquivos de entrada pelo programa PROGEIV, para mostrar ao usuário na tela do monitor do microcomputador, abaixo do gráfico do espectro de I.V. (selecionado para a interpretação), as informações espectrais que levaram a conclusão dos grupos funcionais mais prováveis de estarem ligados ao composto. A regra foi inserida de maneira recursiva para inserir os fatos durante o tempo de execução na base de dados interna "*grupo_funcional*" através do predicado padrão "*assertz*", salvá-los no arquivo DOS "interpr.lnk" e, remover os fatos provenientes da "*real_funcao_grupo*" através do predicado padrão "*retract*".

Até o momento, os procedimentos descritos atingiram o objetivo principal do programa EXPERTIV, que é a identificação dos grupos funcionais mais prováveis em um espectro de I.V. de um determinado composto orgânico, após a interpretação do seu espectro de I.V..

Com os grupos funcionais mais prováveis, ainda armazenados em tempo de execução, na base de dados interna "*grupo_funcional*", implementou-se no INTERMOD algumas regras para verificar através da posição da banda de absorção, a existência de possíveis tipos específicos do composto orgânico em análise, como por exemplo um ácido carboxílico α - β -insaturado. O apêndice B apresenta a base de conhecimento "*dba_links*", arquivada sob o nome DOS "dataiv.dba", onde consta o predicado "*posicao_tipo_especifico*", com quatro argumentos (objetos), que são os fatos espectroscópicos extraídos da literatura, implementados da seguinte maneira:

posicao_tipo_especifico(Grupo_Funcional,Vibricao_Caracteristica,[Faixa_Absorcao],
TipoEspecifico).

Alguns exemplos de tipos específicos dos grupos funcionais estudados neste trabalho e inseridos como argumentos desse predicado são:

```
posicao_tipo_especifico("anidrido_de_acido_(-CO-O-CO)","estiramento_C=O",[1790,1770],
"(alfa_beta_insaturado_6_membros)").
```

```
posicao_tipo_especifico("acido_carboxilico_(-COOH)","estiramento_C=O",[1730,1700],
"(alfa_Br)_ou_(alfa_Cl)_dimer)").
```

```
posicao_tipo_especifico("alcool_(R-OH)","estiramento_O-H",[3630,3610],(alcool_secundario)").
```

Inicialmente, a cláusula para atingir essa meta liga os fatos existentes na base de dados interna *"grupo_funcional"* (veja figura IV-21), através de uma regra que verifica o tipo específico associado ao grupo funcional em análise, principalmente a partir das posições das bandas de absorção, relacionadas com as suas vibrações características. Essa cláusula tem quatro sub-metas e foi implementada da seguinte forma:

busca_tipo_especifico :-

```
    grupo_funcional(GrupoFuncional,ListaVibracoes,_),           % base de dados interna
    encontre_tipo_especifico(GrupoFuncional,ListaVibracoes,ListaEspecificica),
    verifique_anidrido_especifico(Case,GrupoFuncional,ListaEspecificica,NovaEspecificica),
    descarta_lista vazia(NovaEspecificica,RealEspecificica),
    verifique_caso_especifico(Case,GrupoFuncional,RealEspecificica),
    fail,!.                                                       % predicado fail para forçar retrocesso
```

A primeira sub-meta *encontre_tipo_especifico* procede com outras regras, implementadas para verificar todas as bandas de absorção, que estão relacionadas ao mesmo modo vibracional e também, aplicar a função matemática que calcula o *"score"*, apenas para a posição da banda de absorção, visto que, nem sempre há informações nas tabelas de correlações dos livros sobre espectroscopia no I.V., quanto à sua intensidade e largura. Se o *"score"* é maior ou igual a 90 os tipos específicos podem estar associados aos compostos.

A segunda sub-meta *verifique_anidrido_especifico* foi implementada para consultar sobre a presença do grupo funcional anidrido de ácido, caso o mesmo seja um

fato na base de dados interna "*grupo_funcional*". O anidrido de ácido é caracterizado pelas seguintes vibrações, colocados em uma lista da seguinte forma: ["estiramento_C=O", "estiramento_C=O", "estiramento_CO"]. Nesse caso, a sub-meta é bem sucedida apenas se o tipo específico do grupo funcional confirmado para a primeira banda de carbonila (vibração de estiramento C=O), também é verdadeira para a sua segunda banda de carbonila.

Um exemplo dessa sub-meta bem sucedida foi na interpretação do espectro do anidrido de ácido crotônico ($\text{CH}_3\text{CH}=\text{CH}(\text{CO})-\text{O}-(\text{CO})\text{CH}=\text{CHCH}_3$). As posições das duas vibrações de estiramento C=O encontradas pelo READMOD foram 1779 e 1720 cm^{-1} , respectivamente. O tipo específico retornado para esse ácido foi o "acrílico_ou_benzóico" visto que, as referidas posições estão incluídas nas regiões de absorção desse tipo específico, ou seja, de 1795 a 1775 cm^{-1} e de 1735 a 1715 cm^{-1} , respectivamente.

A terceira sub-meta *descarta_lista vazia* foi imposta para não retornar informações com a lista vazia ([]) de tipos específicos, ou seja, retornar apenas os fatos existentes de um tipo específico do grupo funcional que foi bem sucedido para a sua vibração característica.

A quarta sub-meta *verifique_caso_especifico* confirma ou não a existência de tipos específicos associados ao composto em análise. Quando estes fatos são sucedidos para um determinado grupo funcional, estes são inseridos em tempo de execução, na base de dados interna "*tipo_especifico*", através do predicado padrão "*assertz*" e salvos no arquivo DOS "*specific.lnk*". A figura IV-22 mostra esse tipo de arquivo, referente a interpretação do anidrido de ácido crotônico.

```

grupo_funcional_especifico("anidrido_acido_(-CO-O-CO)",
[tipo_vibracao("estiramento_C=O",faixa_absorcao([1795,1775]),
[banda(largura("media",1808,1768,40,"cm-1"),intensidade("muito_forte",7,"%T"),
posicao(1779),[score_tipo_especifico(faixa_absorcao([1795,1775]),
"(acrilico_ou_benzolco)",posicao_score(100),desvio_posicao("Dentro",0,"cm-1"))]]),
tipo_vibracao("estiramento_C=O",faixa_absorcao([1735,1715]),
[banda(largura("medium",1731,1713,18,"cm-1"),intensidade("very_strong",12,"%T"),
posicao(1720),[score_tipo_especifico(faixa_absorcao([1735,1715]),
"(acrilico_ou_benzolco)",score_posicao(100),desvio_posicao("Dentro",0,"cm-1"))]]))]]))

```

FIGURA IV-22 - Fatos da base de dados interna "tipo_especifico", salvos no arquivo DOS "specific.lnk", referente a interpretação do espectro de I.V., do anidrido de ácido crotônico. Os tipos específicos "acrílico_ou_benzóico" associado ao anidrido de ácido estão coerentes para as duas vibrações de estiramento C=O.

Até este momento, além das informações que podem ser obtidas nesse módulo interpretador, outras regras foram implementadas para encontrar outros grupos presentes no composto em análise, como por exemplo uma ligação dupla C=C de compostos olefinicos, halogênios, ou grupos como: aromático, amino, nitro, tio, nitrila e alcino.

Os procedimentos para que esta meta seja bem sucedida são análogos aos descritos para a identificação do grupo funcional, mas com algumas imposições, isto é, alguns grupos funcionais encontrados como mais prováveis, não podem ser analisados como outros grupos possíveis. Alguns exemplos são citados como: a) se o grupo funcional é um alceno, omite-se a análise para a vibração de estiramento C=C de alifáticos insaturados como um outro grupo com ligação_dupla_C=C, b) se o grupo funcional é um aromático, descarta-se a análise para o grupo aromático, alceno e fenol e, c) para um grupo funcional haleto de acila, descarta-se a análise do grupo halogênio (vibrações de estiramento C-X onde, X pode ser F, Cl, Br e I). Caso contrário, não há coerência em

informar, no relatório final para o usuário, que existe um outro grupo ligado na molécula, quando o mesmo já foi identificado como grupo funcional mais provável.

A regra implementada liga os fatos da base de conhecimento "*dba_links*", através do predicado "*vibracoes_caracteristicas_outro_grupo*" em outras duas sub-metas, que verificam na base de dados interna "*primeira_varredura*", a presença ou não das vibrações que caracterizam cada grupo em análise. A primeira sub-meta é colocada para confirmar a existência das vibrações de estiramento C-X, (veja exemplo na figura IV-14) de modo que, essa regra, procede-se apenas para o grupo halogênio. A segunda sub-meta é para confirmar a presença das vibrações que caracterizam os outros grupos. Esses procedimentos estão implementados da seguinte maneira:

busca_outros_grupos :-

```
vibracoes_caracteristicas_outro_grupo(OutroGrupo,ListaVibracoes),    %predicado da "dba_links"
verificar_outro_grupo_estiramento_CX(SubGroup,["estiramento_C-F","estiramento_C-Cl",
"estiramento_C-Br","estiramento_C-I"],ListaVibracoes,VibracoesCaracteristicas),
real_vibracoes(OutroGrupo,VibracoesCaracteristicas,ListaVibracoesGrupo),
assertz(outro_grupo_provavel(OutroGrupo,ListaVibracoesGrupo)),
fail, !.
```

Em seguida, uma outra regra analisa detalhadamente cada banda de absorção, que está relacionada com o seu modo vibracional, de um determinado grupo, ligando os fatos da base de conhecimento "*dba_links*", através do predicado "*tipo_vibracao_outro_grupo*" e, procurando as bandas que apresentam os maiores "*scores*", ou seja, com valores maiores ou iguais a 90. Os fatos desse predicado são removidos pelo predicado padrão "*retract*". Conseqüentemente, os fatos dos grupos definidos como mais prováveis, são inseridos na base de dados interna "*outros_grupos*", através do predicado padrão "*assertz*" e salvos no arquivo DOS "*groups.lnk*". A figura IV-23 mostra uma parte da base de dados interna "*dba_links*", referente ao grupo nitro, atribuído neste trabalho.

```
vibacoes_caracteristicas_outro_grupo("nitro_(C-NO2)","["estiramento_assimétrico_
C-NO2","estiramento_simétrico_C-NO2"]);
tipo_vibricao_outro_grupo("nitro_(C-NO2)","estiramento_assimétrico_C-NO2",
[1650,1500],intensidade(["forte","muito_forte"],largura([])).
tipo_vibricao_outro_grupo("nitro_(C-NO2)","estiramento_simétrico_C-NO2",
[1370,1250],intensidade([],largura([])).
```

FIGURA IV-23 - Parte do arquivo DOS "dataiv.dba", onde estão armazenados os fatos da base de conhecimento "dba_links" para o grupo nitro.

A interpretação de um espectro de I.V. para um determinado composto orgânico é finalizada, quando o INTERMOD atinge com sucesso todas as metas requeridas como a identificação dos grupos funcionais mais prováveis e informações com relação aos tipos específicos e outros grupos ligados na molécula. Como descrito anteriormente, esta interpretação é obtida selecionando a opção número 7, referente ao menu principal do programa PROGEIV (veja figura III-3). A sub-rotina denominada "CALLPROLOG" encarrega-se desta tarefa, abrindo o arquivo de entrada "spec.spt" (espectro corrente da biblioteca, escolhido pelo usuário), no sub-diretório "sistspec\spectrum", para gerar os dados desse arquivo em um outro de saída, denominado "spec.ltd", na forma de lista Turbo Prolog. Em seguida, abre um arquivo de saída denominado "PROLOG.FLG" que confecciona o arquivo de "flag" para o retorno do programa EXPERTIV. Deste modo, fecha o arquivo "PROLOG.FLG", retornando o comando para o arquivo "sistspec.bat", que chama o programa EXPERTIV. O arquivo "sistspec.bat" está confeccionado do seguinte modo:

ECHO OFF

PROGEIV

IF EXIST PROLOG.FLG READMOD

IF EXIST PROLOG.FLG INTERMOD

IF EXIST PROLOG.FLG SISTSPEC

ECHO ON

Finalmente, é retornado para o usuário, na tela do monitor, do microcomputador, abaixo do gráfico do espectro de I.V. interpretado, as informações dos grupos funcionais com os seus tipos específicos e outros grupos mais prováveis de estarem presentes no composto, através de mensagens claras e objetivas. Também, todos os arquivos DOS referentes aos fatos obtidos durante a interpretação do espectro, podem ser impressos antes de uma próxima interpretação.

IV.2 - RESULTADOS E DISCUSSÕES

O programa EXPERTIV foi desenvolvido para identificar, principalmente, o grupo ou grupos funcionais mais prováveis de estarem presentes em um determinado composto orgânico, através da interpretação dos espectros de I.V.. Para que esse objetivo seja alcançado, inicialmente o READMOD faz a leitura das ocorrências do espectro em análise e em seguida, o INTERMOD atua como um filtro. Desta forma, os grupos funcionais incluídos neste trabalho, apenas são eliminados quando os fatos (formato, largura, intensidade e posição de uma banda de absorção) obtidos do espectro, não equiparam definitivamente com os da base de conhecimento "*dba_links*", implementada através das tabelas de correlações da literatura. Normalmente, a interpretação de um espectro de I.V. pelo programa EXPERTIV, com um microcomputador do tipo descrito na parte experimental, é efetuada com um tempo de processamento de 20 segundos em média, dependendo da complexidade do composto em análise, podendo se estender até 45 segundos.

Quando o espectro de um composto carbonilado é interpretado pelo programa EXPERTIV, retorna-se em uma primeira análise, como o grupo ou grupos funcionais mais prováveis, apenas os que tem a função carbonila, descartando os demais. Por exemplo, na interpretação de um ácido carboxílico que tem um grupo aromático ligado a esse composto, o programa pode retornar como grupos funcionais mais prováveis o próprio ácido carboxílico, um éster e uma cetona. No entanto, em uma outra etapa o programa identifica o grupo aromático como um outro grupo ligado ao composto ácido, inserindo os fatos obtidos no arquivo DOS "*groups.lnk*".

Alguns critérios de avaliação foram incluídos no programa, como por exemplo os valores dos "*scores*" para a banda de absorção correspondente a uma vibração característica e para o grupo funcional analisado, que devem ser maiores ou iguais a 75 e 90, respectivamente. Essa tomada de decisão é importante para o processo de filtragem do

INTERMOD, caso contrário, o programa retornaria os diversos grupos funcionais incluídos neste trabalho, com os seus devidos "scores", podendo variar de 0 a 100. A tabela IV-3 mostra como seria retornado os resultados da interpretação do ácido 2,2-dimetilpropiónico, se não fossem atribuídos os valores limites dos "scores".

TABELA IV-3 - Resultados obtidos da interpretação do espectro de I.V. do composto ácido 2,2-dimetilpropiónico, indicando como seria retornado os resultados dessa interpretação para o usuário, sem as atribuições dos valores dos "scores" no INTERMOD. G.F. = grupo funcional, ν = vibração de estiramento, δ = vibração de deformação e * = valores dos "scores" dos halogênios: F = 90 e Cl, Br e I = 80.

G.F.	Scores das Bandas de Absorção (vibrações características)							Score do G.F.
	ν C=O	ν O-H	ν C-O	ν C-C	ν C-H	ν C-X	sobretom δ C-H	
carbonilados								--
ácido carboxílico	100	100	100	-	-	-	-	100
éster	100	80	-	-	-	-	-	90
cetona	100	-	-	100	-	-	-	100
aldeído	86	-	-	-	1	-	26	38
haleto de acila	35	-	-	-	-	*	-	73
não carbonilados	ν C=C	ν O-H	ν C-O	ν C=C	ν C-H	δ C-H	sobretom δ C-H	--
álcool	-	100	37	-	-	-	-	69
fenol	-	43	100	-	-	-	-	72
alceno	82	-	-	-	100	-	-	91
alcano	-	-	-	-	34	100	-	67
éter	-	-	100	-	-	-	-	100

O programa EXPERTIV interpretou trezentos e cinquenta espectros de I.V. de compostos orgânicos obtidos da biblioteca SADTLER e distribuídos em quatorze grupos funcionais, sendo que seis grupos funcionais são carbonilados. Como estes compostos são conhecidos foi possível avaliar com mais precisão o desempenho do programa quanto a identificação do grupo funcional em análise e o retorno de outros grupos funcionais também encontrados como mais prováveis. A tabela IV-4 relaciona os grupos funcionais incluídos neste trabalho com os respectivos números de espectros interpretados e os arquivos DOS `spxxx.spt`, onde xxx pode ser de 000 a 999. Uma visão completa dos resultados obtidos quanto a identificação correta para estes compostos, é indicada na tabela IV-5.

Através desse programa o usuário pode fazer uso dos arquivos DOS obtidos após a interpretação de um determinado espectro de I.V., para verificar com detalhes os relatórios de cada etapa do programa. Por exemplo, quando não é identificado corretamente o grupo funcional de um composto, pode-se recorrer aos arquivos DOS "scanned1.lnk" ou "scanned2.lnk", que indicam porque esta meta não foi bem sucedida. O primeiro arquivo contém os fatos das vibrações de estiramento ou deformação de um composto, com as suas respectivas faixas máximas de absorção (veja figura IV-14) e o último, relaciona os fatos extraídos para cada grupo funcional, com suas vibrações características e estas com a banda ou as bandas de absorção ocorridas na sua devida região do espectro (veja figura IV-15). Desta forma, o usuário pode avaliar onde e porque certos grupos funcionais foram descartados da interpretação. Esses arquivos DOS devem ser utilizados, quando necessários, antes da interpretação de um outro espectro. Caso contrário, o mesmo é eliminado pelo espectro corrente.

O desempenho do programa EXPERTIV para identificar a presença dos grupos funcionais mais prováveis de um determinado composto é descrito abaixo, separadamente para cada grupo funcional analisado.

TABELA IV.4 - Relação dos grupos funcionais analisados pelo programa EXPERTTV, indicando o número total de espectros interpretados para cada grupo funcional e seus respectivos arquivos DOS "spexxx.spt".

Grupos Funcionais	Total de Espectros de I.V. Interpretados	Arquivos DOS - spexxx.spt -
1. anidrido de ácido	18	073-075, 220-234
2. ácido carboxílico	28	000-026, 480
3. éster	44	027-069 e 072
4. cetona	24	273-286, 453-462
5. aldeído	28	091-115, 128, 272, 464
6. haleto de acila	18	235-245, 248-254
7. álcool	25	415-432, 434-440
8. fenol	26	313-331, 441-447
9. alceno	26	255-271, 397-401, 410-413
10. aromático	26	177-184, 187-199, 402, 404-407
11. alcino	17	287-302, 414
12. nitrila	25	157-176, 466-470
13. alcano	24	133-156
14. éter	21	117-127, 129-132 e 448-452

TABELA IV-5 - Resultados obtidos das interpretações dos espectros de I.V. para cada grupo funcional analisado pelo programa EXPERTIV, indicando a sua porcentagem de identificação correta. G.F. = Grupos Funcionais.

G.F. analisados	Total de espectros interpretados	Identificação correta da presença do G.F. no composto	% de identificação correta
1. anidrido de ácido	18	16	88.9
2. ácido carboxílico	28	26	92.3
3. éster	44	42	95.5
4. cetona	24	22	91.7
5. haleto de acila	18	17	94.4
6. aldeído	28	27	96.4
7. álcool	25	24	96.0
8. fenol	26	26	100.0
9. alceno	26	24	92.3
10. aromático	26	24	92.3
11. alcino	17	12	70.6
12. nitrila	25	23	92.0
13. alceno	24	22	91.6
14. éter	21	18	85.7
TOTAL	350		91.4

- COMPOSTOS CARBONILADOS

O programa EXPERTIV interpretou cento e sessenta compostos carbonilados distribuídos em seis grupos funcionais, como indicado na tabela IV-4. Através destas interpretações observou-se que o programa é capaz de identificar a presença dos grupos funcionais mais prováveis de estarem presentes no composto em análise, como mostra a tabela IV-6.

A figura IV-24 ilustra o desempenho do programa EXPERTIV quanto a identificação dos grupos funcionais carbonilados mais prováveis, indicando o número de compostos em que foi reconhecida a presença do grupo funcional e sua respectiva porcentagem de identificação correta.

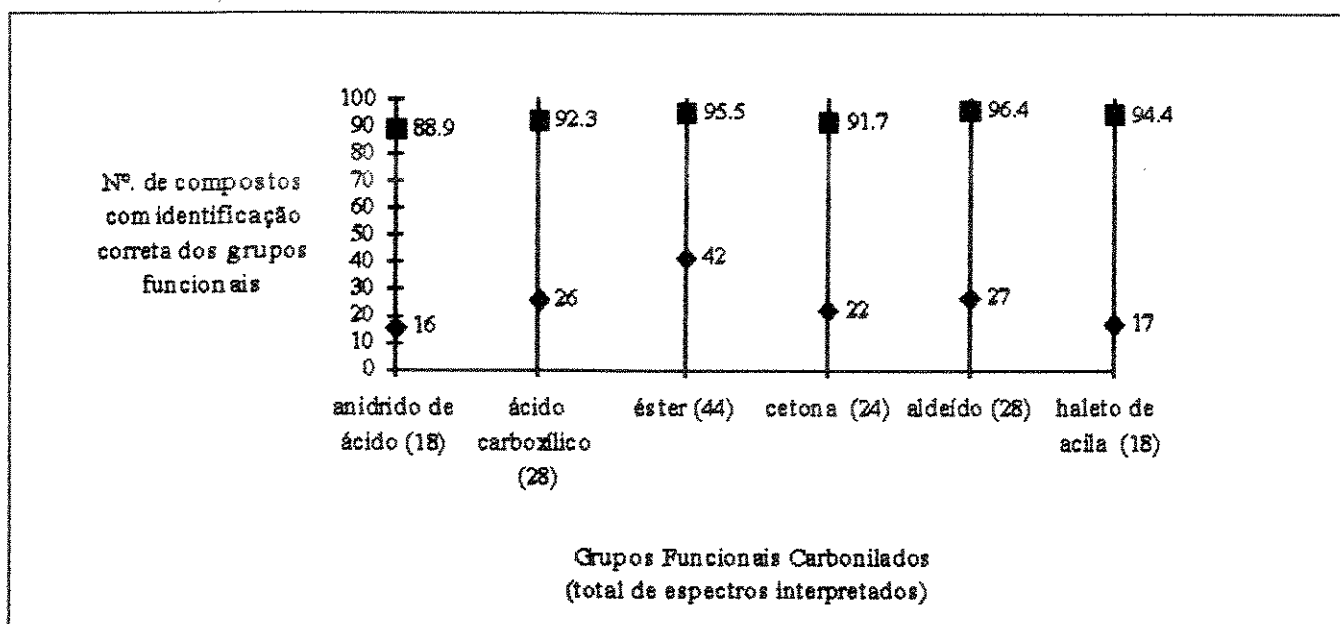


FIGURA IV-24 - Desempenho do programa EXPERTIV na interpretação de cento e sessenta espectros de I.V. de compostos carbonilados, para identificar os grupos funcionais mais prováveis. ♦ indica o número de compostos em que o programa identificou corretamente a presença do grupo funcional carbonilado em análise e ■ indica a porcentagem de identificação desses grupos.

TABELA IV-6 - Resultados obtidos das interpretações dos espectros de I.V. de compostos carbonilados, relacionando a identificação dos grupos funcionais mais prováveis. G.F. = grupos funcionais, E.I. = espectros interpretados e I.C. = identificação correta para o grupo funcional em análise.

G.F. carbonilados identificados nos compostos								
G.F. analisados	Nº E.I.	% I.C.	anidrido de ácido	ácido carboxílico	éster	cetona	haleto de acila	aldeído
anidrido de ácido	18	88.9	16	1	10	16	9	-
ácido carboxílico	28	92.3	1	26	9	24	2	-
éster	44	95.5	2	4	42	44	16	1
cetona	24	91.7	-	-	6	22	1	-
haleto de acila	18	94.4	1	1	10	14	17	-
aldeído	28	96.4	-	2	6	21	1	27

- ANIDRIDO DE ÁCIDO

Em dezoito espectros de I.V. interpretados de anidridos de ácidos, foram identificados corretamente a sua presença em dezesseis compostos, com um índice de acerto de 88.9%. No espectro spe225.spt do anidrido cinâmico ($C_6H_5-CH=CH-(C=O)-O-(C=O)-CH=CH-C_6H_5$), a segunda vibração de estiramento C=O foi encontrada pelo READMOD em 1703 cm^{-1} , de modo que o valor do "score" calculado para essa banda de absorção foi igual a 24, muito abaixo do valor necessário que é maior ou igual 75. Na interpretação do espectro spe227.spt do anidrido glicínico ($H_2N-CH_2-(CO)-O-(CO)-CH_2-NH_2$) foram encontradas as vibrações de estiramento C=O em 1704 e 1682 cm^{-1} , respectivamente. Essas posições estão abaixo da região mínima inserida na base de conhecimento "dba_links", de acordo com os dados da literatura, que é de 1735 a 1715 cm^{-1} .

Nos espectros dos anidridos de ácidos também foi identificada a presença de outros grupos funcionais como mais prováveis, retornando a cetona, o éster e o haleto de acila. As vibrações que caracterizam esses grupos funcionais, praticamente ocorrem nas mesmas regiões das vibrações dos anidridos de ácidos analisados. Por exemplo, no grupo cetônico as vibrações de estiramento C=O e de C-C em $-C-(CO)-C$, no éster as vibrações de estiramento C=O e C-O e, no haleto de acila as vibrações de estiramento C=O e C-X, onde X pode ser F, Cl, Br e I, podem sobrepor as vibrações características do anidrido de ácido analisado. Portanto, esses grupos funcionais às vezes não são descartados pelo programa. No caso do haleto de acila há uma probabilidade maior devido a vibração de estiramento C-X, onde o programa busca informações para os quatro halogênios citados.

A tabela IV-7 mostra os resultados das interpretações dos espectros de I.V. de anidridos de ácidos, com os grupos funcionais mais prováveis. A figura IV-25 ilustra a situação do programa EXPERTIV, indicando o número de compostos em que foi

identificada a presença de grupos funcionais mais prováveis, com as suas respectivas porcentagens de identificação.

TABELA IV-7 - Resultados obtidos de dezoito espectros de I.V. de anidridos de ácidos, interpretados pelo programa EXPERTIV, com o retorno dos grupos funcionais mais prováveis. * Referente aos compostos não identificados como anidridos de ácidos.

Grupos funcionais retornados das interpretações dos espectros de I.V. de anidridos de ácidos	Nº de compostos identificados
anidrido de ácido, éster, cetona e haleto de acila	7
anidrido de ácido e cetona	4
anidrido de ácido, éster e cetona	2
anidrido de ácido, cetona e haleto de acila	1
anidrido de ácido, éster e haleto de acila	1
anidrido de ácido, ácido carboxílico, éster, cetona e haleto de acila	1
* cetona	2

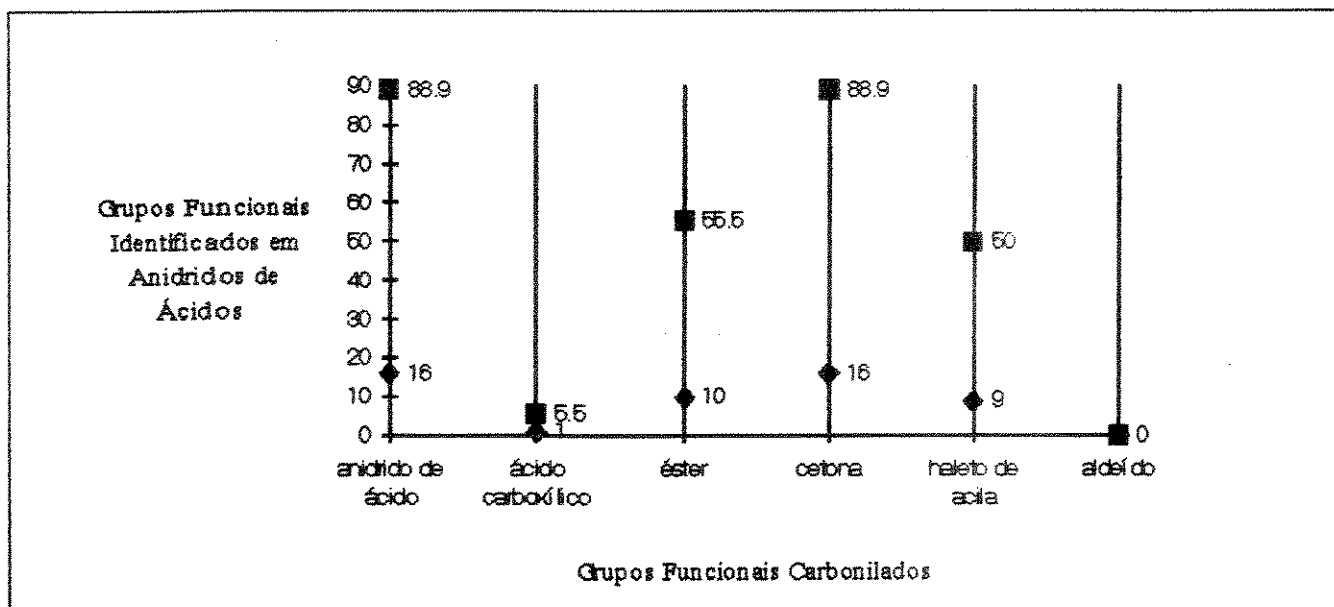


Figura IV-25 - Desempenho do programa EXPERTIV nas interpretações de dezoito espectros de I.V. de anidridos de ácidos. ◆ indica o número de compostos em que o programa identificou o grupo funcional carbonilado e ■ indica a porcentagem de identificação dos grupos funcionais.

- ÁCIDO CARBOXÍLICO

Para os ácidos carboxílicos foram interpretados vinte e oito espectros de I.V. pelo programa EXPERTIV, que identificou corretamente a sua presença em vinte e seis compostos, com um índice de acerto de 92.3%. O espectro spe013.spt referente ao ácido trifluoracético (CF_3COOH) apresenta duas bandas de absorção para a vibração de estiramento $\text{C}=\text{O}$ em 1819 e 1788 cm^{-1} com intensidades de 42 e 52 %T, respectivamente. Essas intensidades foram determinantes no processo de filtragem do INTERMOD, pois as informações (fatos) na base de conhecimento "*dba_links*" são forte ou muito forte (0 a 35 %T), de modo que a diferença de 7 %T diminuiu o valor do "*score*" calculado da primeira carbonila, não alcançando portanto o valor mínimo admitido pelo programa. O programa retornou como grupo funcional mais provável o álcool e o éter. No entanto, mesmo não obtendo êxito em identificar esse ácido carboxílico, foi observado no arquivo DOS "*groups.lnk*" a identificação do flúor como um outro possível grupo ligado no ácido, através da banda de absorção posicionada em 1196 cm^{-1} , com uma intensidade muito forte correspondente a 10 %T.

O espectro spe018.spt do ácido 7-aminoeptanóico ($\text{H}_2\text{N}-\text{CH}_2-(\text{CH}_2)_4\text{CH}_2-(\text{CO})-\text{OH}$) também não foi identificado como um ácido carboxílico, mas como um grupo cetônico. Isto deve-se principalmente a posição da banda de absorção da carbonila encontrada em 1633 cm^{-1} , estando abaixo da região mínima inserida na base de conhecimento "*dba_links*" que é de 1675 a 1655 cm^{-1} , portanto uma diferença em 22 cm^{-1} , que resultou um "*score*" igual a 52. Esse fator foi predominante para descartar esse grupo funcional no decorrer da análise feita pelo INTERMOD. A largura dessa banda de absorção é aguda com 11 cm^{-1} e a intensidade é muito forte com 13 %T, estando dentro das condições requeridas para tais reconhecimentos. Também, esse espectro tem uma linha base calculada pelo READMOD em 66 %T, de modo que a sua interpretação apenas foi completada, alterando especificamente o limite mínimo atribuído em 70 %T.

Dos vinte e oito espectros de I.V. interpretados para os ácidos carboxílicos observou-se que o programa identificou a presença de outros grupos funcionais carbonilados, que foram retornados como mais prováveis. O anidrido de ácido retornou apenas uma vez, conforme o espectro spe024.spt referente ao ácido tetrafluorftálico ($F_4-C_6-((CO)-OH)_2$), que apresenta duas bandas de absorção encontradas em 1732 e 1726 cm^{-1} , respectivamente, correspondentes as vibrações de estiramento $C=O$, característica esta relevante para a identificação dessa função anidrido de ácido. O grupo éster foi identificado em nove compostos devido a posição da carbonila e bandas de absorção para a vibração de estiramento $C-O$. O grupo cetônico foi identificado em vinte e quatro compostos, predominando pela sua ampla faixa de absorção da carbonila e a ocorrência da sua vibração de estiramento $C-C$ em $-C-(CO)-C-$ dentro da mesma região da vibração de estiramento $C-O$ dos ácidos. O grupo haleto de acila foi identificado em dois compostos, conforme os espectros spe014.spt e spe482.spt referentes aos ácidos tricloroacético ($CCl_3-(CO)-OH$) e dibromoacético ($Br_2-CH-(CO)-OH$), pois além da ocorrência da carbonila encontrada, os halogênios permitiram tais identificações.

A tabela IV-8 mostra os grupos funcionais que foram retornados como mais prováveis, com os seus respectivos números de compostos identificados. A figura IV-26 ilustra a atuação do programa EXPERTIV na interpretação dos espectros de ácidos carboxílicos, indicando o número de compostos em que foram identificados os grupos funcionais com a sua porcentagem de identificação.

TABELA IV-8 - Resultados obtidos de vinte e oito espectros de I.V. de ácidos carboxílicos interpretados pelo programa EXPERTIV, com o retorno dos grupos funcionais mais prováveis. * Referentes aos compostos não identificados como ácidos carboxílicos.

Grupos funcionais retornados das interpretações dos espectros de I.V. de ácidos carboxílicos	Nº de compostos identificados
ácido carboxílico	3
ácido carboxílico e cetona	13
ácido carboxílico, cetona e éster	7
ácido carboxílico, cetona, éster e haleto de acila	2
ácido carboxílico, cetona e anidrido de ácido	1
* cetona	1
* álcool e éter	1

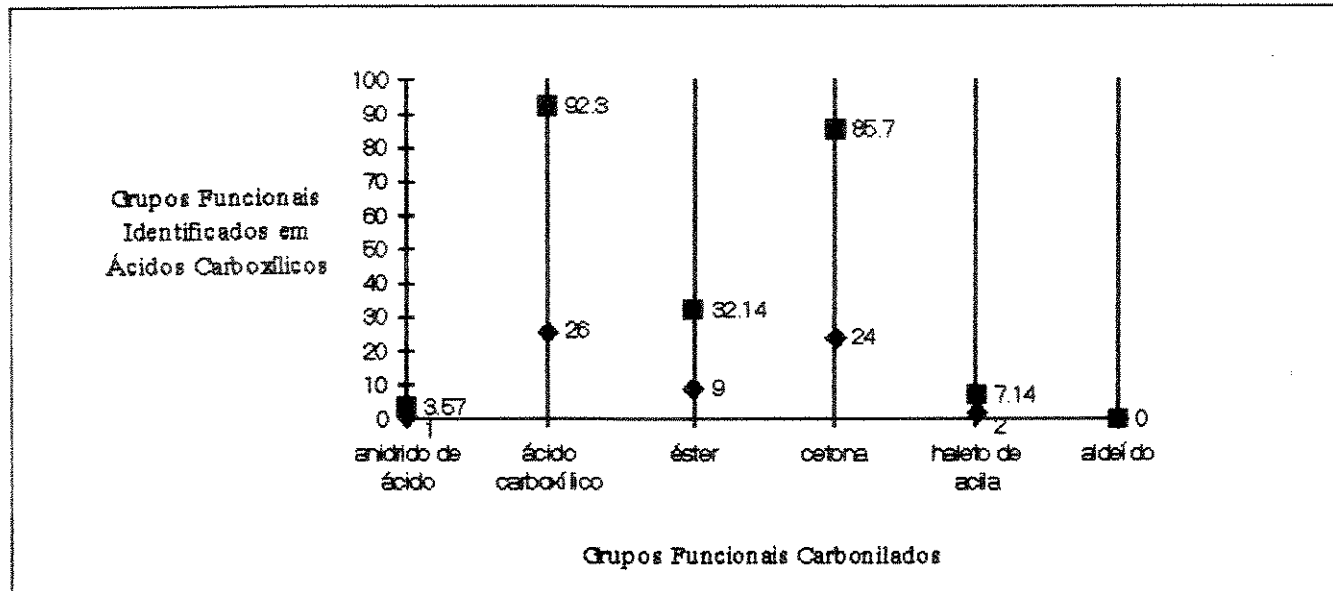


Figura IV-26 - Desempenho do programa EXPERTIV nas interpretações de vinte e oito espectros de I.V. de ácidos carboxílicos. ♦ indica o número de compostos em que o programa identificou o grupo funcional carbonilado e ■ indica a porcentagem de identificação dos grupos funcionais.

- ÉSTER

Em quarenta e quatro espectros de I.V. interpretados para os compostos que tem o grupo funcional éster, o programa EXPERTIV identificou corretamente quarenta e dois compostos, obtendo-se desta forma um índice de acerto de 95.5%. Os dois espectros não identificados foram o spe039.spt e spe068.spt referentes aos compostos propionato de *n*-metila ($\text{CH}_3\text{CH}_2\text{-(CO)-O-CH}_3$) e laurato de *n*-butila ($\text{CH}_3\text{CH}_2(\text{CH}_2)_8\text{CH}_2\text{-(CO)-O-CH}_2(\text{CH}_2)_2\text{CH}_3$) que apresentam bandas de absorção na região de 1300-1050 cm^{-1} para suas vibrações de estiramento C-O, com intensidades de média a fraca. Essas ocorrências espectrais não estão de acordo com os fatos da base de conhecimento "*dba_links*". Portanto, durante as análises desses espectros, o grupo éster foi descartado pelo processo de filtragem do INTERMOD, retornando como grupos funcionais mais prováveis a cetona e haleto de acila para o primeiro espectro e a cetona para o segundo espectro.

Também foi verificado que o programa constatou a presença de outros grupos funcionais carbonilados, retornando-os como mais prováveis o anidrido de ácido em dois compostos, justificados pela ocorrência de duas absorções para a carbonila, conforme os espectros spe031.spt do formiato de *n*-butila ($\text{H-(CO)-O-CH}_2\text{CH}_2\text{CH}_2\text{CH}_3$) e spe037.spt do dibromoacetato de etila ($\text{Br}_2\text{CH-(CO)-OCH}_2\text{CH}_3$). O ácido carboxílico foi retornado em quatro compostos devido principalmente a ocorrência da vibração de estiramento O-H, com uma banda de absorção intensa e larga, de forma que em três compostos a identificação está correta. O haleto de acila foi identificado em dezessete compostos porque a absorção da carbonila ocorre dentro da região permitida para o grupo éster e a vibração de estiramento C-X, que tem uma probabilidade maior do programa identificá-la no espectro, porque X pode ser igual F, Cl, Br e I. Através da interpretação de cinco ésteres com um halogênio presente no composto, observou-se que todos esses halogênios foram identificados corretamente. O grupo cetônico prevaleceu mais uma vez, retornando também como mais provável em todos ésteres analisados. Este fato atribui-se a sua ampla

faixa de absorção de 1860 a 1590 cm^{-1} para a vibração de estiramento C=O, que é praticamente idêntica a do éster e a vibração de estiramento C-C em -C-(CO)-C- de 1300 a 1100 cm^{-1} , ocorrendo na mesma região da vibração de estiramento C-O.

A tabela IV-9 relaciona os resultados obtidos das interpretações dos espectros de I.V. dos ésteres, relacionando o número de compostos que foram identificados os grupos funcionais carbonilados como mais prováveis. A figura IV-27 ilustra o desempenho do programa EXPERTIV para os ésteres, indicando o número de compostos em que o grupo funcional foi identificado e sua respectiva porcentagem de identificação.

TABELA IV-9 - Resultados obtidos de quarenta e quatro espectros de I.V. de ésteres interpretados pelo programa EXPERTIV, com o retorno dos grupos funcionais identificados como mais prováveis. * Referentes aos compostos não identificados como ésteres.

Grupos funcionais retornados das interpretações dos espectros de I.V. de ésteres	Nº de compostos identificados
éster e cetona	20
éster, cetona e haleto de acila	15
éster, cetona e ácido carboxílico	4
éster, cetona e anidrido de ácido	2
éster, cetona e aldeído	1
* cetona	1
* cetona e haleto de acila	1

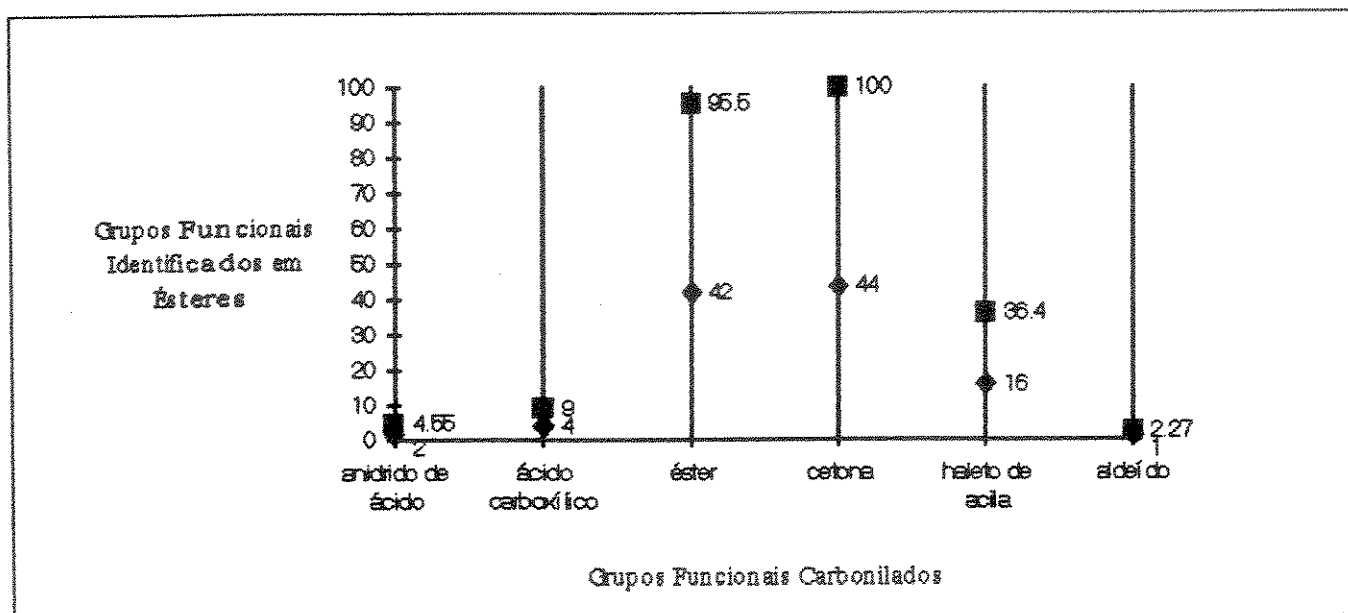


FIGURA IV-27 - Desempenho do programa EXPERTIV nas interpretações de quarenta e quatro espectros de I.V. de compostos com o grupo funcional éster. ♦ indica o número de compostos em que o programa identificou o grupo funcional carbonilado e ■ indica a porcentagem de identificação dos grupos funcionais.

- CETONA

Nos vinte e quatro espectros de I.V. de compostos com grupos cetônicos interpretados pelo programa EXPERTIV, foi identificada a sua presença em vinte e dois compostos, com um índice de acerto de 91.7%. Nos espectros spe275.spt e spe460.spt dos compostos 3-3-dimetil-2-butanona ($((\text{CH}_3)_3\text{C}-(\text{CO})\text{CH}_3)$) e 1,3-dicloro-2-propanona ($\text{Cl}-\text{CH}_2-(\text{CO})-\text{CH}_2-\text{Cl}$), as vibrações de estiramento $\text{C}=\text{O}$ encontram-se dentro das condições exigidas para a cetona. No entanto, suas vibrações de estiramento $\text{C}-\text{C}$ em $-\text{C}(\text{CO})-\text{C}-$ apresentam-se com intensidades fracas. Consequentemente, os valores de seus "scores" calculados ficaram abaixo do limite mínimo permitido no INTERMOD portanto, o grupo cetônico foi eliminado do restante da análise. Os retornos das interpretações dos

referidos espectros foram o grupo funcional alceno. Neste caso, mesmo tendo o programa identificado a presença da carbonila, as vibrações que caracterizam os grupos funcionais carbonilados analisados neste trabalho, não atingiram o valor mínimo de "score", razão pela qual não foram retornados como mais prováveis.

Através dos resultados obtidos das vinte e quatro interpretações dos espectros de I.V. para as cetonas, verificou-se que o programa EXPERTIV também identificou a presença de outros grupos funcionais carbonilados como mais prováveis, retornando o ácido carboxílico em dois compostos, o éster em seis e o haleto de acila em apenas um. A tabela IV-10 indica esses resultados relacionando o número de espectros interpretados e o retorno dos grupos funcionais identificados como mais prováveis.

TABELA IV-10 - Resultados obtidos de vinte e quatro espectros de I.V. de compostos cetônicos interpretados pelo programa EXPERTIV, com o retorno dos grupos funcionais identificados como mais prováveis. * Referentes aos compostos não identificados como cetona.

Grupos funcionais retornados das interpretações dos espectros de I.V. de cetonas	Nº de compostos identificados
cetona	16
cetona e éster	5
cetona, éster, aldeído e haleto de acila	1
* alceno	2

A figura IV-28 ilustra o desempenho do programa EXPERTIV para os compostos cetônicos, indicando o número de compostos em que foram identificados a presença de grupos funcionais mais prováveis e sua respectiva porcentagem de identificação.

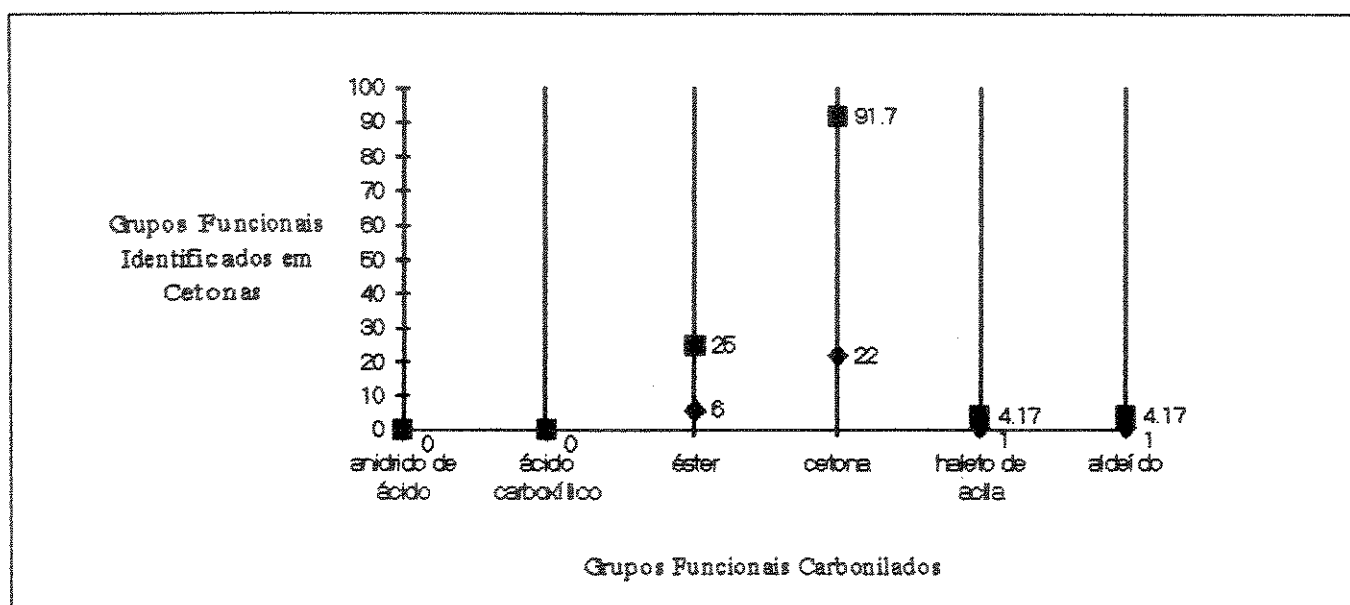


FIGURA IV-28 - Desempenho do programa EXPERTIV nas interpretações de vinte e quatro espectros de I.V. de compostos cetônicos. ◆ indica o número de compostos em que o programa identificou o grupo funcional carbonilado e ■ indica a porcentagem de identificação dos grupos funcionais.

- ALDEÍDO

O programa EXPERTIV interpretou vinte e oito espectros de I.V. de aldeídos e identificou corretamente sua presença em vinte e sete compostos, obtendo-se um índice de acerto de 96.4%. Para o espectro spe093.spt do composto crotonaldeído ($\text{CH}_3\text{-CH=CH-CHO}$), a vibração de estiramento C=O foi encontrada pelo READMOD em 1627 cm^{-1} , estando bem abaixo da região mínima inserida na base de conhecimento "*dba_links*", que é de $1675\text{ a }1655\text{ cm}^{-1}$. Desta forma, o valor do "*score*" calculado para essa posição não atingiu o valor necessário no programa e o INTERMOD descartou a possibilidade desse grupo funcional estar presente no composto.

Das análises destes espectros foi possível observar que o programa identificou a presença de outros grupos funcionais carbonilados como mais prováveis, retornando o grupo cetônico em vinte e um compostos, o éster em seis, o ácido carboxílico em dois e o haleto de acila em apenas um. A forma na qual os grupos funcionais foram retornados, pode ser vista na tabela IV-11.

Com relação a outros grupos funcionais não carbonilados ligados aos compostos com função aldeídica, o programa mostrou um bom desempenho para identificá-los. Por exemplo, conforme as observações feitas nos seus arquivos DOS "groups.lnk", o programa identificou corretamente o grupo aromático presente nos seguintes compostos: difenil-acetaldeído, benzaldeído, 2,5-dimetil-benzaldeído, *p*-isopropil-benzaldeído, *o*-tolualdeído, *m*-tolualdeído, *p*-tolualdeído e cinamaldeído. Essa mesma atuação foi obtida, agora com os grupos aromático e nitro, presentes nos seguintes compostos: *o*-nitrobenzaldeído, *m*-nitrobenzaldeído e *p*-nitrobenzaldeído. Isto indica mais uma vez que o programa é capaz de reconhecer os grupos funcionais quando presentes em um determinado composto orgânico, mesmo sendo polifuncional.

A atuação do programa EXPERTIV para os aldeídos pode ser observada na figura IV-29, que indica o número de compostos onde identificou-se a presença do grupo funcional mais provável, com sua respectiva porcentagem de identificação.

TABELA IV-11 - Resultados obtidos de vinte e oito espectros de I.V. de aldeídos interpretados pelo programa EXPERTIV, com os grupos funcionais carbonilados identificados como mais prováveis. * Referente ao composto não identificado como aldeído.

Grupos funcionais retornados das interpretações dos espectros de I.V. de aldeídos	Nº de compostos identificados
aldeído	7
aldeído e cetona	11
aldeído, cetona e éster	6
aldeído, cetona e ácido carboxílico	2
aldeído, cetona e haleto de acila	1
* cetona	1

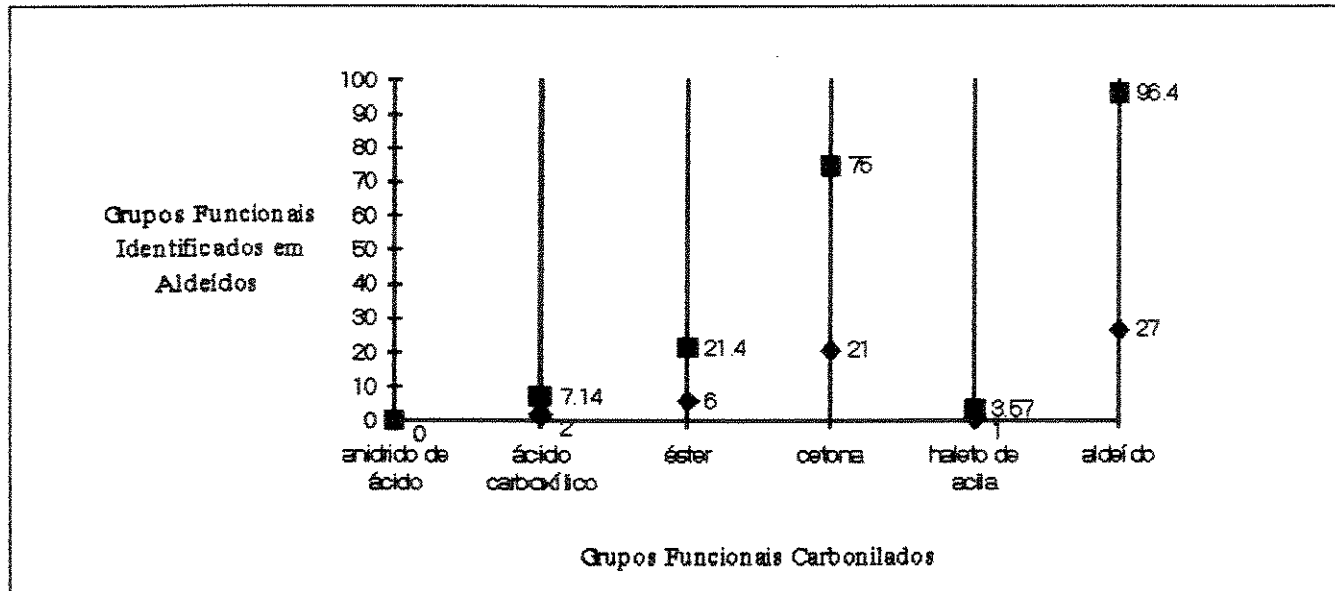


FIGURA IV-29 - Desempenho do programa EXPERTIV nas interpretações de vinte e oito espectros de I.V. de aldeídos. ♦ indica o número de compostos em que o programa identificou o grupo funcional carbonilado e ■ indica a porcentagem de identificação dos grupos funcionais.

- HALETOS DE ACILA

Para os compostos haletos de acila o programa EXPERTIV interpretou dezoito espectros de I.V., de forma que a sua presença foi identificada em dezessete compostos, com um índice de acerto de 94.4%. O espectro spe252.spt do composto brometo de benzofila ($C_6H_5-(CO)-Br$) não foi identificado pelo programa, retornando a cetona como grupo funcional mais provável.

Através dos resultados obtidos das interpretações dos espectros de haletos de acila, verificou-se que o programa também identificou a presença de outros grupos funcionais carbonilados. O grupo cetônico predominou em quatorze compostos, depois o éster em dez, o anidrido de ácido em quatro e o ácido carboxílico em apenas um composto. O anidrido de ácido foi identificado em quatro compostos devido principalmente as ocorrências de duas bandas de absorção, encontradas para a carbonila. A tabela IV-12 mostra os resultados obtidos das interpretações dos haletos de acila, com os seus respectivos retornos dos grupos funcionais carbonilados identificados como mais prováveis.

Esses resultados estão ilustrados na figura IV-30 que indica o número de compostos em que o programa EXPERTIV identificou a presença dos grupos funcionais mais prováveis e sua respectiva porcentagem de identificação.

TABELA IV-12 - Resultados obtidos de dezoito espectros de I.V. de haletos de acila interpretados pelo programa EXPERTIV, com os grupos funcionais carbonilados identificados como mais prováveis. * Referente ao composto não identificado como haleto de acila.

Grupos funcionais retornados das interpretações dos espectros de I.V. de haletos de acila	Nº de compostos identificados
haleto de acila	1
haleto de acila, cetona e éster	6
haleto de acila e cetona	3
haleto de acila e éster	2
haleto de acila, éster, cetona e anidrido de ácido	2
haleto de acila, cetona e anidrido de ácido	1
haleto de acila, ácido carboxílico e cetona	1
haleto de acila e anidrido de ácido	1
* cetona	1

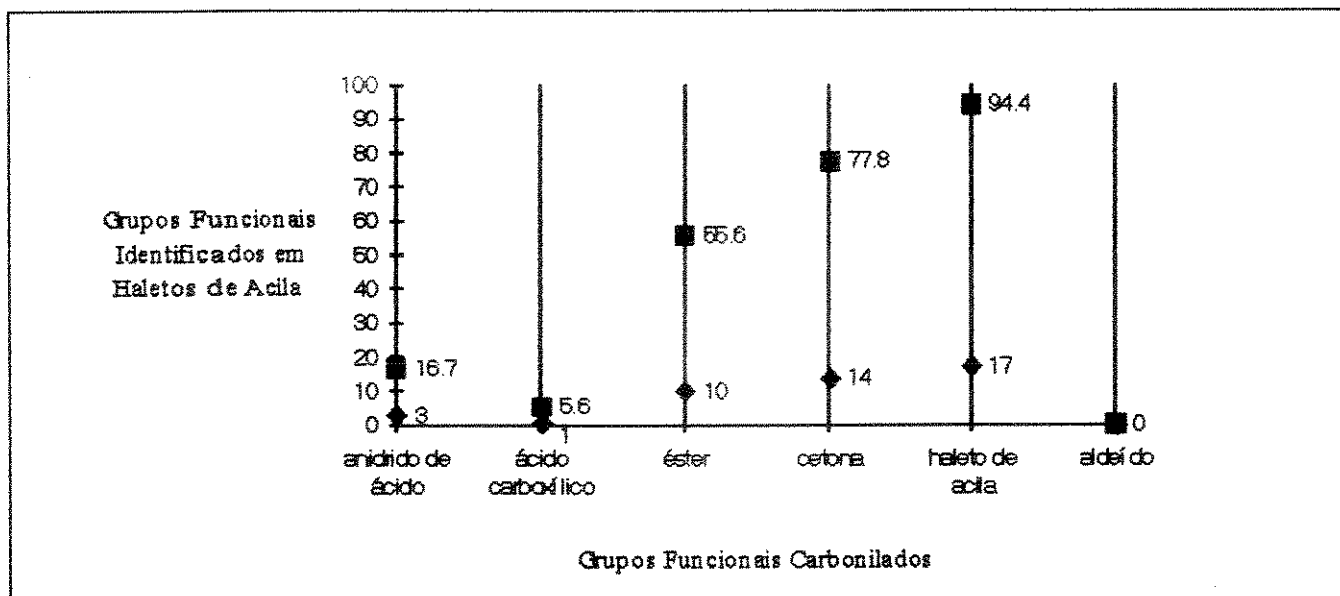


FIGURA IV-30 - Desempenho do programa EXPERTIV nas interpretações de dezoito espectros de I.V. de haletos de acila. ♦ indica o número de compostos em que o programa identificou o grupo funcional carbonilado e ■ indica a porcentagem de identificação dos grupos funcionais.

- COMPOSTOS NÃO CARBONILADOS

O programa EXPERTIV interpretou cento e noventa espectros de I.V. de compostos não carbonilados, distribuídos em oito grupos funcionais como mostrado na tabela IV-4. Uma avaliação completa quanto a identificação dos grupos funcionais ligados ao composto em análise, pode ser observada na tabela IV-13. A atuação do programa nas interpretações destes espectros é descrita abaixo para cada grupo funcional não carbonilado.

TABELA IV-13 - Resultados obtidos do programa EXPERTIV nas interpretações de cento e noventa espectros de I.V. de compostos não carbonilados, para identificar os grupos funcionais mais prováveis. G.F. = grupo funcional, E.I. = espectros interpretados e I.C. = identificação correta da presença do grupo funcional.

Grupos Funcionais Identificados nos Compostos										
G.F. analisado	Nº E.I.	% I.C.	álcool	fenol	aromático	alceno	alcino	nitrila	alcano	éter
álcool	25	96.0	24	8	1	6	--	--	22	18
fenol	26	100.0	26	26	--	20	--	--	17	21
aromático	26	92.3	--	--	24	19	--	--	26	9
alceno	26	92.3	--	--	6	24	--	--	20	6
alcino	17	70.6	--	--	8	4	12	11	13	4
nitrila	25	92.0	--	--	14	14	18	23	16	8
alcano	24	100.0	1	--	--	3	--	--	24	--
éter	21	85.7	--	--	8	7	--	--	14	18

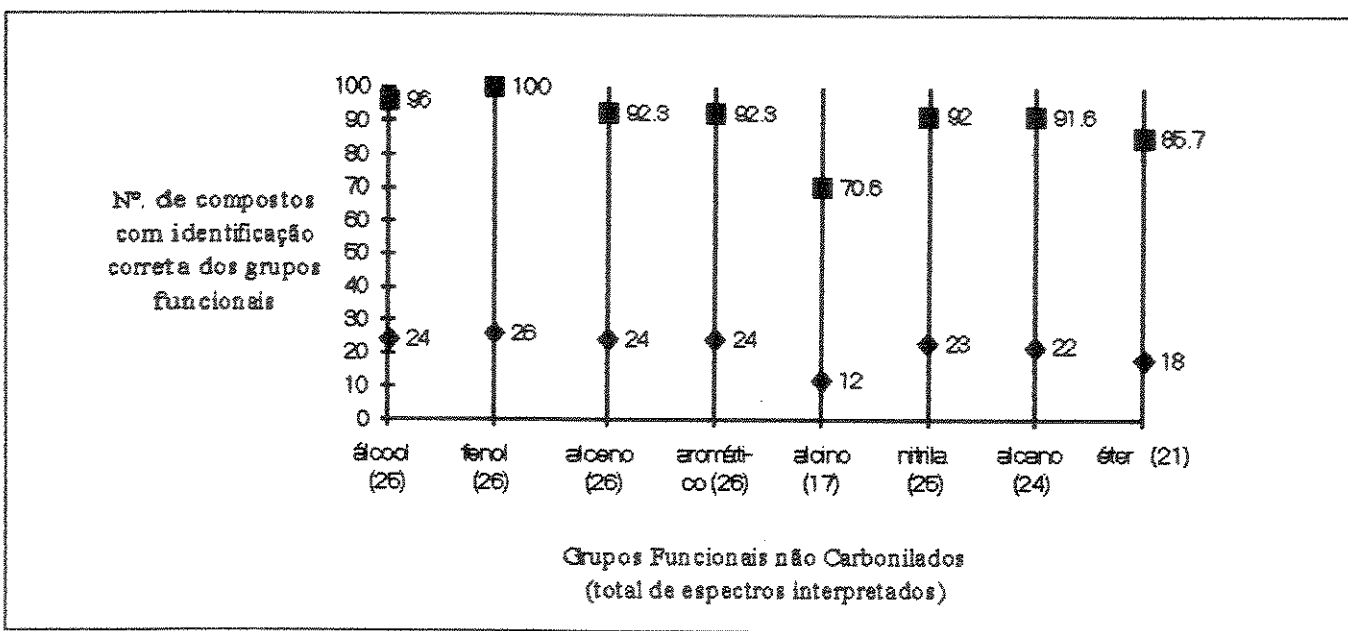


FIGURA IV-31 - Desempenho do programa EXPERTIV nas interpretações de cento e noventa espectros de I.V. de compostos não carbonilados. ◆ indica o número de compostos que o programa identificou corretamente a presença do grupo funcional não carbonilado em análise e ■ indica a porcentagem de identificação dos grupos funcionais.

- ÁLCOOL

Dos vinte e cinco espectros de I.V. de álcoois interpretados pelo programa EXPERTIV identificou-se corretamente vinte e quatro, obtendo-se desta forma um índice de acerto de 96.0%. Apenas o espectro spe419.spt referente ao composto 1,3-diamino-2-propanol ($\text{H}_2\text{N}-\text{CH}_2-\text{CH}(\text{OH})-\text{CH}_2-\text{NH}_2$) não foi identificado como um álcool, mas retornou um grupo alcano. Através do arquivo DOS "scanned2.lnk" observou-se que foi encontrada uma banda de absorção em 3346 cm^{-1} , com largura de 452 cm^{-1} e intensidade, muito forte, de 10 %T, que corresponde a vibração de estiramento O-H. Todavia, a banda de absorção encontrada na região entre 1280 a 1000 cm^{-1} para a vibração de estiramento

C-O, apresenta uma intensidade muito fraca de 82 %T, de modo que na base de conhecimento "*dba_links*" os fatos da intensidade são média e forte, o que proporcionou o descarte desse grupo funcional.

O programa também identificou a presença de outros grupos funcionais como mais prováveis de estarem ligados aos compostos alcoólicos analisados. O grupo aromático retornou apenas do 2-buten-1-ol, o qual foi identificado incorretamente devido a dificuldade em distinguir a vibração de estiramento C=C de alifático insaturado daquela do aromático. O grupo alceno foi identificado em seis espectros de álcoois, sendo que em cinco destes existe a ligação dupla C=C como: 2-metil-2-propen-1-ol, 2-buten-1-ol, 3-hexen-1-ol, 1-octen-3-ol, *o*-metoxi- α -metilbenzenol. Como a vibração de estiramento C-O do éter e do álcool, ocorre praticamente na mesma região do espectro com intensidades forte e muito forte, o éter foi retornado como mais provável em dezoito compostos e o grupo alceno predominou em vinte um.

Os resultados das interpretações dos espectros de I.V. de álcoois são apresentados na tabela IV-14, que relaciona o retorno dos grupos funcionais identificados como mais prováveis, com o número de compostos identificados.

TABELA IV-14 - Resultados obtidos de vinte e cinco espectros de I.V. de álcoois interpretados pelo programa EXPERTIV, com o retorno dos grupos funcionais identificados como mais prováveis.* Referente ao composto não identificado como álcool.

Grupos funcionais retornados das interpretações dos espectros de I.V. de álcoois	Nº de compostos identificados
álcool, fenol, alceno e éter	6
álcool e alceno	5
álcool, alceno e éter	4
álcool, alceno, alceno e éter	3
álcool e éter	2
álcool, alceno e alceno	1
álcool, fenol e éter	1
álcool, fenol, alceno, alceno e éter	1
álcool, aromático, alceno, alceno e éter	1
* alceno	1

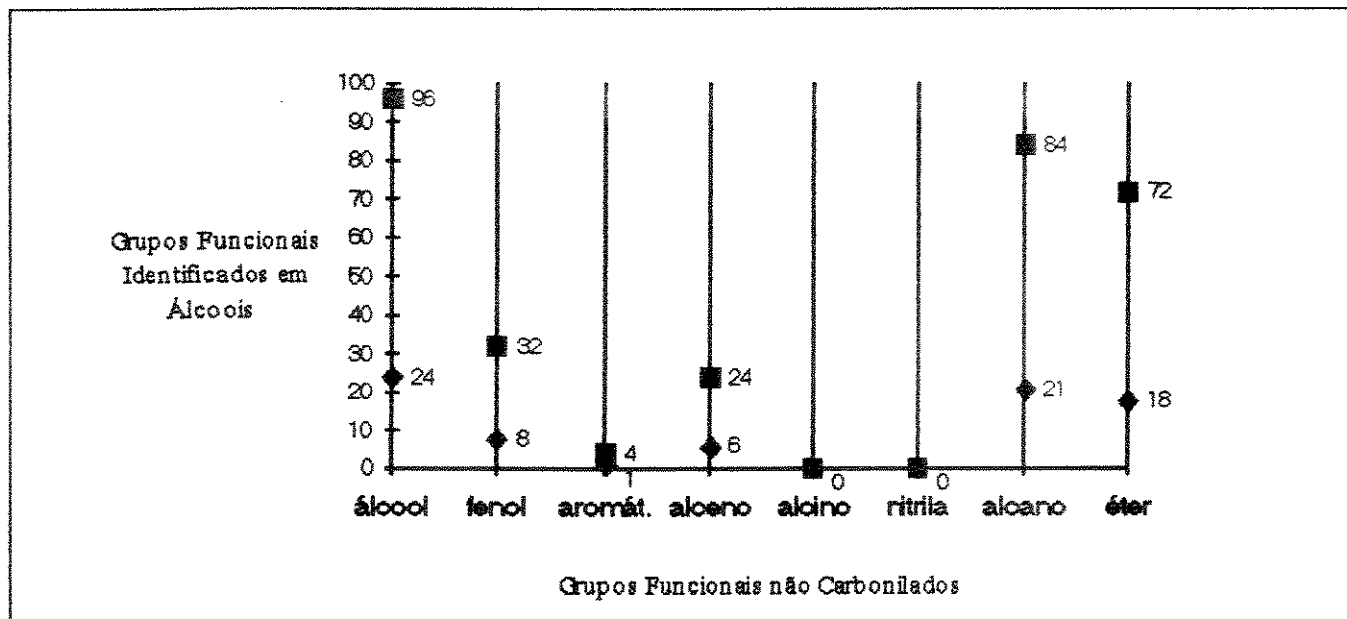


FIGURA IV-32 - Desempenho do programa EXPERTIV nas interpretações de vinte e cinco espectros de I.V. de compostos alcoólicos. ◆ indica o número de compostos que o programa identificou o grupo funcional e ■ indica a porcentagem de identificação dos grupos funcionais.

- FENOL

Os vinte e seis espectros de I.V. de compostos fenólicos interpretados pelo programa EXPERTIV, foram identificados corretamente. É importante ressaltar que além das vibrações de estiramento O-H e C-O que caracterizam esse grupo, o INTERMOD busca informações a respeito da vibração de estiramento C=C de aromático na região de 1650 a 1450 cm^{-1} , mostrando assim que a confirmação da presença desse grupo funcional no composto é válida. Uma outra observação é que em todos os fenóis analisados, também retorna o álcool como um dos grupos funcionais mais prováveis existente no composto. Entretanto, dos vinte e cinco espectros de I.V. interpretados para os álcoois, apenas em oito compostos foi indicada a existência do grupo fenólico. Isso revela a importância da regra implementada no módulo interpretador para confirmar a presença do grupo fenol. Também foi identificada a presença do grupo alceno em vinte compostos fenólicos, devendo-se principalmente as vibrações de estiramento C-H e C=C de compostos insaturados alcenos e de aromáticos que ocorrem praticamente nas mesmas regiões do espectro.

Na região acima de 2000 cm^{-1} os espectros de I.V. de álcoois, fenóis e ácidos carboxílicos podem apresentar uma ou mais bandas de absorção, de modo que a mais importante é a vibração de estiramento O-H, característica desses grupos funcionais. Como descritas no READMOD, as regras implementadas para distinguir um álcool ou fenol de um ácido carboxílico exigiram maiores cuidados e a confiabilidade nos resultados somente foi possível, após várias avaliações das interpretações nos seus setenta e nove espectros. Em nenhum dos compostos alcóolicos e fenólicos analisados neste trabalho foi identificado a presença do ácido carboxílico e vice-versa.

- ALCENO

O programa EXPERTIV interpretou vinte e seis espectros de I.V de alcenos, identificando a sua presença em vinte e quatro compostos, portanto com um índice de acerto de 92.3%. Os espectros spe266.spt e spe267.spt dos compostos *trans*-2,5-dimetil-3-hexeno e *trans*-3-hepteno foram identificados como alcanos. A identificação do grupo alceno nesses dois espectros não foi possível devido a não ocorrência da banda de absorção na região de 3090 a 2960 cm^{-1} , correspondente a vibração de estiramento C-H de grupos alifáticos insaturados. Essa avaliação pode ser feita através das informações do arquivo DOS "scanned1.lnk", que relaciona os fatos obtidos das ocorrências espectrais do composto em análise, como o tipo de vibração e sua respectiva faixa máxima de absorção (veja figura IV-14).

Nas interpretações dos compostos insaturados olefinicos o programa também identificou a presença de outros grupos funcionais como mais prováveis. Os grupos aromático e éter foram identificados em seis compostos e, o alceno em vinte compostos. A tabela IV-15 relaciona os resultados obtidos das interpretações dos compostos olefinicos, indicando o retorno dos grupos funcionais mais prováveis com o número de compostos identificados.

TABELA IV-15 - Resultados obtidos de vinte e seis espectros de I.V. de compostos olefinicos interpretados pelo programa EXPERTIV, com o retorno dos grupos funcionais identificados como mais prováveis. * Referente aos compostos que não foram identificados como alceno.

Grupos funcionais não carbonilados retornados das interpretações dos espectros de I.V. de alcenos	Nº de compostos identificados
alceno	3
alceno e alceno	12
alceno, alceno e éter	3
alceno, aromático e alceno	2
alceno, aromático e éter	2
alceno e aromático	1
alceno, aromático, alceno e éter	1
* alceno	2

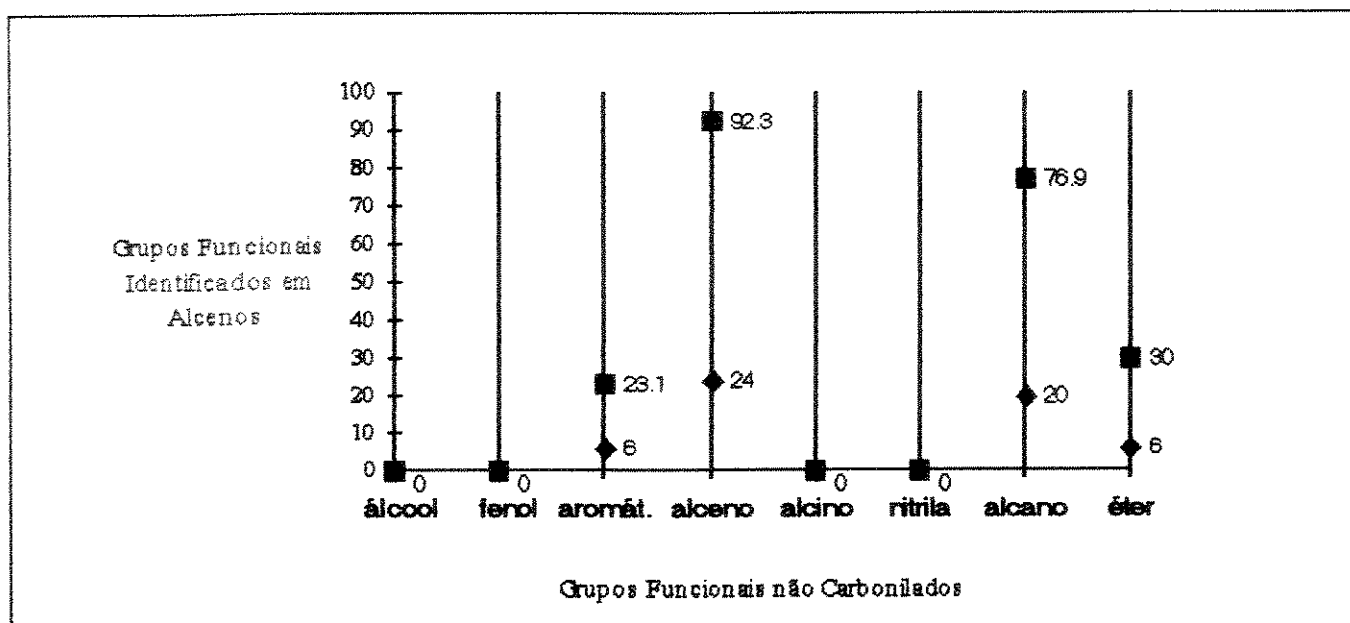


FIGURA IV-33 - Desempenho do programa EXPERTIV nas interpretações de vinte e seis espectros de I.V. de compostos olefinicos. ◆ indica o número de compostos que o programa identificou o grupo funcional e ■ indica a porcentagem de identificação dos grupos funcionais.

- AROMÁTICO

O programa EXPERTIV interpretou vinte e seis espectros de I.V. de compostos aromáticos, de modo que em vinte e quatro compostos a sua presença foi identificada corretamente, com um índice de acerto 92.3%. Conforme os dados da tabela IV-13 pode-se verificar a identificação da presença de outros grupos funcionais nestes compostos. A tabela IV-16 mostra em detalhes os resultados das interpretações destes compostos aromáticos com os grupos funcionais retornados como mais prováveis.

TABELA IV-16 - Resultados obtidos de vinte e seis espectros de I.V. de compostos aromáticos interpretados pelo programa EXPERTIV, com o retorno dos grupos funcionais identificados como mais prováveis. * Referentes aos compostos não identificados como grupo aromático.

Grupos funcionais não carbonilados retornados das interpretações dos espectros de I.V. de compostos aromáticos	Nº de compostos identificados
aromático, alceno e alceno	16
aromático, alceno e éter	4
aromático, alceno, alceno e éter	3
aromático e alceno	1
* alceno e éter	2

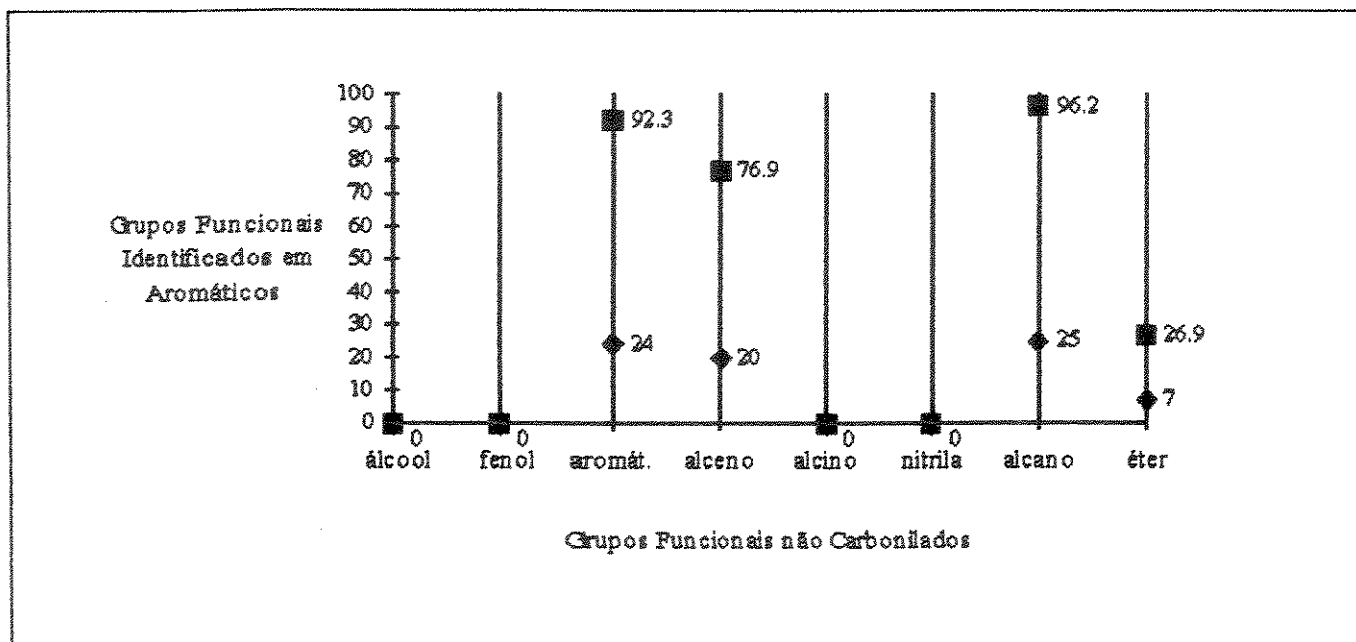


FIGURA IV-34 - Desempenho do programa EXPERTIV nas interpretações de vinte e seis espectros de I.V. de compostos aromáticos. ♦ indica o número de compostos que o programa identificou o grupo funcional não carbonilado e ■ indica a porcentagem de identificação dos grupos funcionais.

- ALCINO

Dos dezessete espectros de I.V. interpretados de compostos acetilênicos, o programa EXPERTIV identificou a presença da ligação C=C em doze compostos, portanto com um índice de acerto de 70.6%. O fato desse índice encontrar-se abaixo dos outros grupos funcionais estudados neste trabalho, deve-se principalmente a não ocorrência da vibração de estiramento C=C na região do espectro entre 2260 a 2090 cm^{-1} , conforme as observações feitas nos espectros originais. A tabela IV-17 relaciona os cinco compostos acetilênicos não identificados corretamente, mas retornando outros grupos funcionais como mais prováveis de estarem ligados nestes compostos.

TABELA IV-17 - Relação dos compostos acetilênicos não identificados pelo programa EXPERTIV, com o retorno dos grupos funcionais identificados como mais prováveis.

Compostos acetilênicos não identificados pelo programa EXPERTIV	Grupos funcionais identificados como mais prováveis
difenilalcino	alceno e aromático
2-butino	nitrila e alceno
4,4-dimetil-2-pentino	alceno e éter
9,9-bis(2-propinil)fluoreno	aromático
1,4-dicloro-2-butino	aromático e éter

Nas interpretações de oito espectros de compostos acetilênicos que contêm o grupo aromático, o programa identificou a presença desse grupo em sete compostos, portanto com um índice de acerto de 87.5%. Os resultados foram considerados satisfatórios visto que, foi identificado corretamente para os espectros dos seguintes compostos: etinilbenzeno, 1-propinilbenzeno, 5-fenil-1-pentino, 6-fenil-2-hexino, 5-hexinilfluoreno, difenilacetileno e 9,9-bis(2-propinil)fluoreno. Na interpretação dos dois últimos compostos citados, o programa não identificou o grupo alcino, mas obteve informações sobre o grupo aromático.

- NITRILA

Em vinte e cinco espectros de I.V. interpretados de compostos nitrilados, o programa identificou corretamente a sua presença em vinte e três compostos, com um índice de acerto de 92.0%. A análise destes compostos revelaram alguns fatos que confirmam o desempenho do programa EXPERTIV como por exemplo, na interpretação do espectro spe469.spt do benzoilacetonitrila ($C_6H_5(C=O)C=N$), que retornou a cetona como o grupo funcional mais provável. Essa identificação deve-se ao processo de filtragem do INTERMOD, onde os compostos carbonilados predominam sobre os não carbonilados. Uma banda de absorção aguda de 11 cm^{-1} , com intensidade muito forte de 11 %T e localizada em 1679 cm^{-1} , caracterizou a vibração de estiramento C=O para o grupo cetônico. No entanto, foram identificados como outros grupos ligados ao composto uma ligação dupla C=C de compostos insaturados, um grupo aromático e um grupo nitrila com uma banda de absorção aguda, correspondente a vibração de estiramento C=N posicionada em 2246 cm^{-1} .

Também, o programa EXPERTIV não identificou o grupo nitrila na interpretação do espectro spe468.spt, referente ao composto o 2,6-dicloro-fenil-acetonitrila ($Cl_2-C_6H_3-CH_2-C=N$), mas retornou como grupos funcionais mais prováveis o alceno, o aromático, o alceno e o éter. No arquivo DOS "groups.lnk" verificou-se a identificação do cloro como um outro grupo ligado a esse composto. Através dos resultados obtidos das interpretações dos espectros acima citados, pode-se concluir que mesmo o programa não reconhecendo o grupo nitrila como a função principal, ele foi capaz de identificar outros grupos funcionais nesses compostos, auxiliando o usuário da espectroscopia no I.V. nas suas conclusões.

A presença de outros grupos funcionais ligados nestes compostos foi confirmada pelo programa. Por exemplo, em quatorze compostos nitrilados foi identificada a existência dos grupos aromático e alceno, em dezoito o grupo *alceno*, em dezesseis o grupo alceno e

em oito o grupo éter. O grupo aromático foi confirmado nos compostos: 1-ciclo-exeno-1-acetonitrila, fenilacetonitrila, *m*-clorofenilacetonitrila, 2-etil-2-fenil-butironitrila, benzonitrila, 2,3-dimetil-benzonitrila, *p*-bromo-benzonitrila, 4-fluorbenzonitrila, *p*-nitrobenzonitrila, *o*-tolunitrila, α -bromo-*o*-tolunitrila, α -bromo-*p*-tolunitrila e 2,6-diclorofenilacetonitrila. Quanto a identificação do grupo acetileno apresentar um índice elevado, deve-se principalmente a sua vibração de estiramento C=C, que se localiza praticamente na mesma região da vibração de estiramento C=N, característica dos compostos nitrilados.

- ALCANO

Para os hidrocarbonetos alifáticos saturados ou alcanos, o programa EXPERTIV interpretou vinte e quatro espectros de I.V. e os resultados foram significativamente positivos, ou seja, identificou-se corretamente sua presença em todos os seus compostos analisados. Também verificou-se que não foi identificado a presença de outros grupos funcionais ligados aos mesmos. Isto demonstra que o programa tem capacidade de eliminar os demais grupos funcionais estudados neste trabalho, exibindo um relatório que explica como e porque se chegou a conclusão que o composto é um hidrocarboneto alifático saturado.

- ÉTER

Em vinte e um espectros de I.V. interpretados para compostos de éteres foram identificados corretamente dezoito compostos, com um índice de acerto de 85.7 %. No espectro spe122.ltd do éter hexadecil ($\text{CH}_3(\text{CH}_2)_{14}\text{CH}_2\text{-O-CH}_2(\text{CH}_2)_{14}(\text{CH}_3)$) para a vibração de estiramento C-O foi encontrada uma banda de absorção aguda de 14 cm^{-1} , posicionada em 1131 cm^{-1} , mas com uma intensidade fraca de 63 %T. Na base de conhecimento "*dba_links*" o fato correspondente a essa intensidade é de forte a muito forte (0 a 35 % T). Portanto, a eliminação do grupo funcional foi devido a diferença de 28 %T, valor esse considerável para o "*score*" obtido abaixo do valor determinado no programa, que retornou o alcano como grupo funcional mais provável.

Nas interpretações dos espectros de éteres também foi identificada a presença de outros grupos funcionais. Por exemplo, o grupo alcano foi identificado em dezenove compostos, o alceno em cinco e, o acetileno e o nitrila em apenas um composto. Em oito compostos de éteres que contêm o grupo aromático foi identificada a presença do grupo aromático em sete compostos, com um índice de acerto de 87.5 %. Os compostos que retornaram o grupo aromático foram éter etilvinílico, éter alilfenílico, éter *o*-toluilalílico, éter *p*-toluilalílico, éter fenilpropílico, éter butilfenílico, éter *p-terc*-butilfenildodecílico e éter fenílico.

Conforme os resultados acima descritos para os trezentos e cinquenta espectros de I.V., pode-se concluir que a principal proposta do programa EXPERTIV foi bem sucedida, isto é, o programa mostrou a sua capacidade em identificar qual ou quais os grupos funcionais mais prováveis, existentes em um determinado composto orgânico.

Além disso, o programa busca informações relacionadas aos possíveis tipos específicos associados ao grupo funcional do composto em análise, como por exemplo se um composto cetônico é de seis membros, se um ácido carboxílico é α,β -insaturado ou α -halogenado ou, se um álcool é terciário, etc.. Como descrito no INTERMOD esta meta pode ser atingida se o valor do "score" calculado para a posição da banda de absorção, correspondente a uma certa vibração de estiramento ou de deformação em análise, é maior ou igual a 90. Normalmente, os livros de espectroscopia no I.V. não contêm dados sobre a intensidade e largura da banda de absorção, para a maioria dos tipos específicos de um grupo funcional. Isto nos condicionou a implementar uma regra utilizando-se uma função matemática para calcular o "score", apenas para a posição da banda de absorção. Este procedimento melhorou substancialmente o desempenho do programa EXPERTTV, pois se elimina um grupo funcional do resto da análise, apenas quando a posição da banda de absorção, obtida do espectro de I.V. pelo READMOD é equiparada com todas as posições de seus tipos específicos existentes na base de conhecimento "dba_links", obtendo-se, assim, um valor do "score" menor do que o citado acima.

Os resultados obtidos na identificação dos tipos específicos quando presentes no composto em análise são bastante positivos, considerando-se as dificuldades encontradas em extrair da literatura maiores informações das vibrações que os envolvem. A tabela IV-18 indica alguns destes fatos obtidos com base na posição da vibração de estiramento C=O de diversos tipos específicos, relacionando os grupos funcionais, os nomes de alguns compostos carbonilados e o seu tipo específico, identificados pelo programa como mais provável. A figura IV-35 ilustra as fórmulas estruturais de alguns dos compostos relacionados na referida tabela.

TABELA IV-18 - Relação de grupos funcionais de compostos carbonilados, com os seus tipos específicos identificados pelo programa EXPERTIV, baseado na posição da vibração de estiramento C=O. G.F. = grupos funcionais, ν = vibração de estiramento, T.E = tipos específicos, *as letras correspondem às fórmulas estruturais representadas na figura IV-35.

G.F. / Compostos	T.E.	G.F. / Compostos	T.E.
Anidrido de Ácido	—	Ácido Carboxílico	—
crotônico ^{(a)*}	acrílico ou benzóico	2-bromoheptanoico	α -Br, α -Cl dímero
tetracloroftálico ^(b)	anel de 5 membros	isoftálico ^(h)	Ar-COOH
succínico ^(c)	anel de 5 membros	fenilacético ⁽ⁱ⁾	Ar-COOH
benzóico	acrílico ou benzóico	p-tolilacético	Ar-COOH
Éster	—	Cetona	—
fumarato de dietila ^(d)	Ar-CO-O ou -C=C-CO-O-	ciclopropilmetilcetona	ciclopropil-CO-R
dibromoacetato de <i>n</i> -etila	α -halo	1,3-difenil-2-propanona	Ar-CO-
2-hepteniato de <i>n</i> -metila	Ar-CO-O ou -C=C-CO-O-	2-etilciclohexanona	anel de 6 membros
ftalato de mono-metila	Ar-CO-O ou -C=C-CO-O-	2-butilciclohexanona	anel de 6 membros
tartarato de dietila	Ar-CO-O ou -C=C-CO-O-	1,3-difluor-2-propanona	-CX ₂ -CO ou -CX-CO-CX
Haletos de Acila	—	Aldeído	—
cloreto de 3-bromopropionila ^(e)	-CO-Cl saturado	p-isopropilbenzaldeído ^(j)	Ar-CHO
cloreto de benzóila ^(f)	-CO-Cl insaturado	3,7-dimetil-2,6-octadienal	α - β -insaturado
cloreto de o-tolúila ^(g)	-CO-Cl insaturado	cinamaldeído	α - β -insaturado
cloreto de undecanoíla	-CO-Cl saturado	cinamaldeído	α - β -insaturado

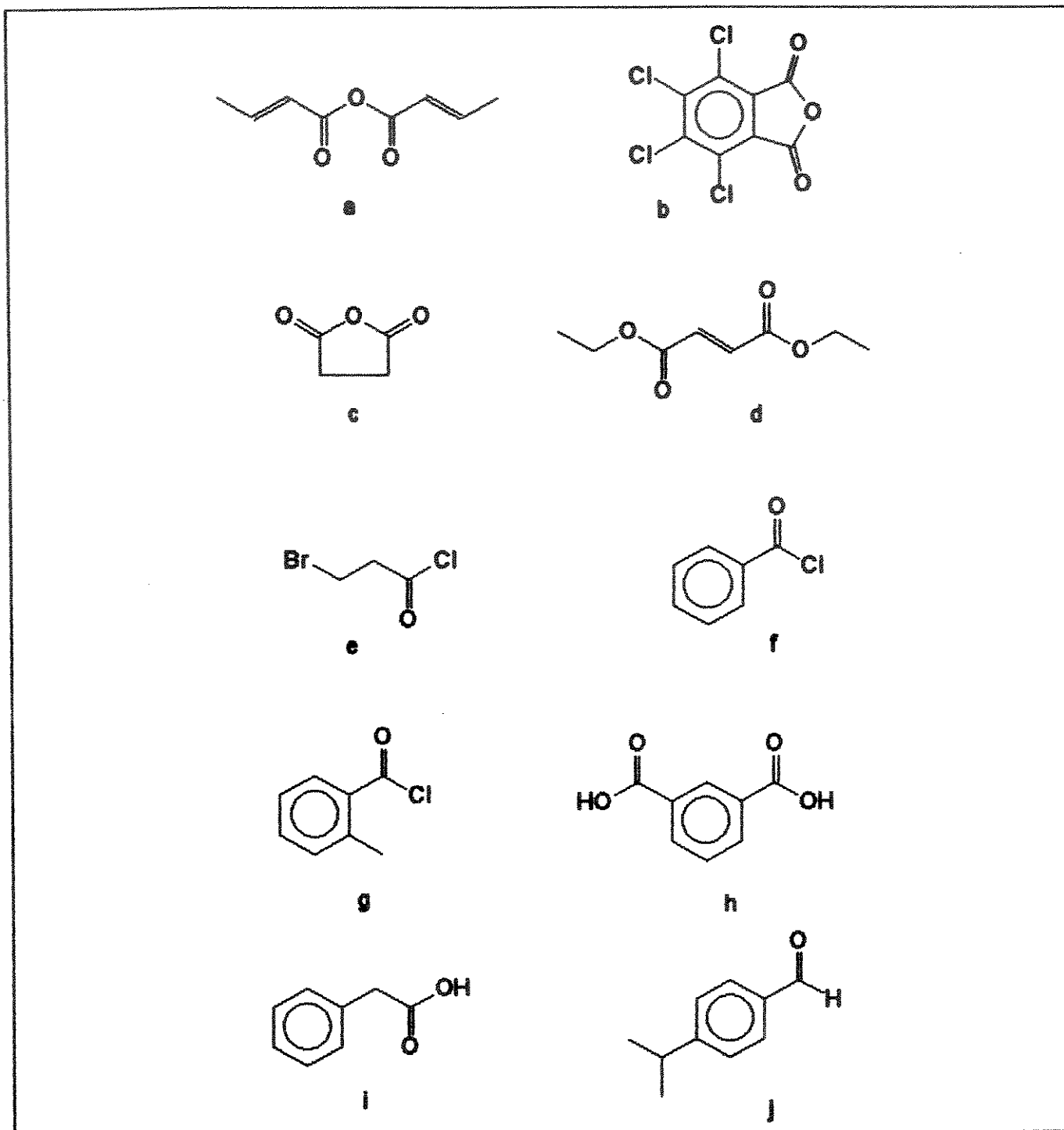


FIGURA IV-35 - Fórmulas estruturais de alguns compostos da tabela IV-18, para as quais o programa EXPERTIV identificou os grupos funcionais e os tipos específicos, através dos seus espectros de I.V..

Durante o desenvolvimento do programa EXPERTIV foi possível implementar mais algumas regras, desta vez buscando novamente os fatos armazenados no arquivo DOS "scanned1.lnk" (veja exemplo na figura IV-14), que contêm informações a respeito das vibrações que caracterizam outros grupos existentes no composto em análise. Desta forma, o usuário pode obter um relatório completo sobre o conjunto de grupos funcionais mais prováveis que formam o composto, até então desconhecido.

Outros grupos com suas vibrações características inseridos no programa foram: 1) tio (ν C-S, ν S-H e ν C=S), 2) amina primária (ν N-H, ν C-N e δ N-H no plano), 3) amina secundária (ν N-H e ν C-N), 4) nitro (ν C-NO₂ assimétrico e ν C-NO₂ simétrico), 5) ligação dupla C=C de compostos olefinicos (ν C=C e ν C-H), 6) halogênio (ν C-F, ν C-Cl, ν C-Br e ν C-I), 7) aromático (ν C=C, ν C-H e δ C-H fora do plano), 8) alcinos (ν C≡C) e 9) nitrila (ν C≡N), onde ν é igual a vibração de estiramento e δ é igual a vibração de deformação. O programa retorna os grupos de 6 a 9 apenas se os mesmos não foram identificados como grupo funcional mais provável. Por exemplo, no caso do grupo halogênio não pode constar na base de dados interna "*funcional_group*" o haleto de acila, para o grupo aromático o próprio grupo aromático e o fenol. A tabela IV-19 relaciona os resultados de alguns compostos orgânicos existentes na biblioteca do programa PROGEIV, indicando o grupo funcional, o seu tipo específico e outros grupos identificados como mais prováveis. Os grupos fluoreto, cloreto e nitro foram retornados de muitos compostos aos quais não pertenciam, devido a ocorrência das suas vibrações na região de "*impressão digital*". Por isso, necessita-se, talvez além da posição e intensidade da banda de absorção, outras informações como por exemplo a largura, formato, etc., para tomar decisões mais confiáveis na identificação dos mesmos. Infelizmente, estes dados não estão disponíveis na literatura e somente uma compilação de um número significativos de espectros destes tipos de compostos, poderia levar ao estabelecimento de regras confiáveis.

No caso de compostos não carbonilados as informações sobre o conjunto dos possíveis grupos funcionais quando ligados na molécula, também revelaram resultados

muito satisfatórios. O programa retorna praticamente todos os grupos funcionais que formam a estrutura dos mesmos, permitindo dessa maneira um relatório final para o usuário mais completo e com maior credibilidade.

TABELA IV-19 - Resultados obtidos das interpretações dos espectros de I.V. de alguns compostos orgânicos pelo programa EXPERTIV, com o retorno dos grupos funcionais, os tipos específicos e outros grupos identificados como mais prováveis. G.F. = grupos funcionais, * as letras correspondem às fórmulas estruturais representadas na figura IV-36.

Composto	G.F. mais prováveis	Tipos Específicos	Outros Grupos
Ácido Carboxílico	-----	-----	-----
2-bromo-3-metilbutírico ^{(a)*}	ácido carboxílico	α -Br ou α -Cl , dímero	brometo
2-bromoheptenóico	ácido carboxílico e cetona	α -Br ou α -Cl , dímero	-
difenilacético ^(b)	ácido carboxílico	C=C-COOH	aromático
<i>p</i> -tolilacético ^(c)	ácido carboxílico e cetona	Ar-COOH	-
isoftálico	ácido carboxílico e cetona	Ar-COOH	aromático
tetrafluorftálico	ácido carboxílico, anidrido de ácido e cetona	C=C-COOH	aromático
2-cloropropiônico	ácido carboxílico, éster e cetona	α -Br ou α -Cl , dímero	brometo
fenilacético	ácido carboxílico e cetona	Ar-COOH	aromático
Éster	-----	-----	-----
propionato de alila	éster e cetona		C=C (alceno)
<i>p</i> -nitrobenzoato de <i>n</i> -etila ^(d)	éster e cetona	Ar-CO-O ou -C=C-CO-O-	NO ₂
ftalato de dipropila ^(e)	éster e cetona	-	aromático
fumarato de dietila	éster e cetona	Ar-CO-O ou -C=C-CO-O-	-

continua na próxima página

continuação da tabela IV-19

Composto	G.F. mais prováveis	Tipos Específicos	Outros Grupos
Cetona	-----	-----	-----
2-cicloexilcicloexanona	cetona	anel de 6 membros	-
4-etilcicloexanona	cetona	anel de 6 membros	-
cicloexanona	cetona e éster	anel de 6 membros	-
1-cloro-2-propanona ^(f)	cetona	-CX ₂ -CO ou -CX-CO-CX	fluoreto e cloreto
Haleta de Acila	-----	-----	-----
cloreto de acetila-D ₃	haleta de acila e éster	-CO-Cl saturado	-
cloreto de pimeloila ^(g)	haleta de acila e cetona	-CO-Cl saturado	-
cloreto de suberoila ^(h)	haleta de acila e cetona	-CO-Cl saturado	-
4-clorofornilbutirato de metila	haleta de acila, ácido carboxílico e cetona	-CO-Cl saturado	-
cloreto de p-nitrobenzoila		-CO-Cl saturado	aromático
Aldeído	-----	-----	-----
difenilacetaldeído ⁽ⁱ⁾	aldeído, cetona e haleta de acila	-	aromático
2,5-dimetil-benzaldeído ^(j)	aldeído e cetona	α - β -insaturado	aromático
o-nitrobenzaldeído	aldeído, éster e cetona	Ar-CHO	aromático, NO ₂
m-nitrobenzaldeído ^(k)	aldeído, éster e cetona	Ar-CHO	aromático, NO ₂
p-nitrobenzaldeído	aldeído, éster e cetona	Ar-CHO	aromático, NO ₂
o-tolualdeído	aldeído e cetona	Ar-CHO	aromático
m-tolualdeído	aldeído, éster e cetona	Ar-CHO	aromático
10-undecenal ^(l)	aldeído	-	v C=C alceno

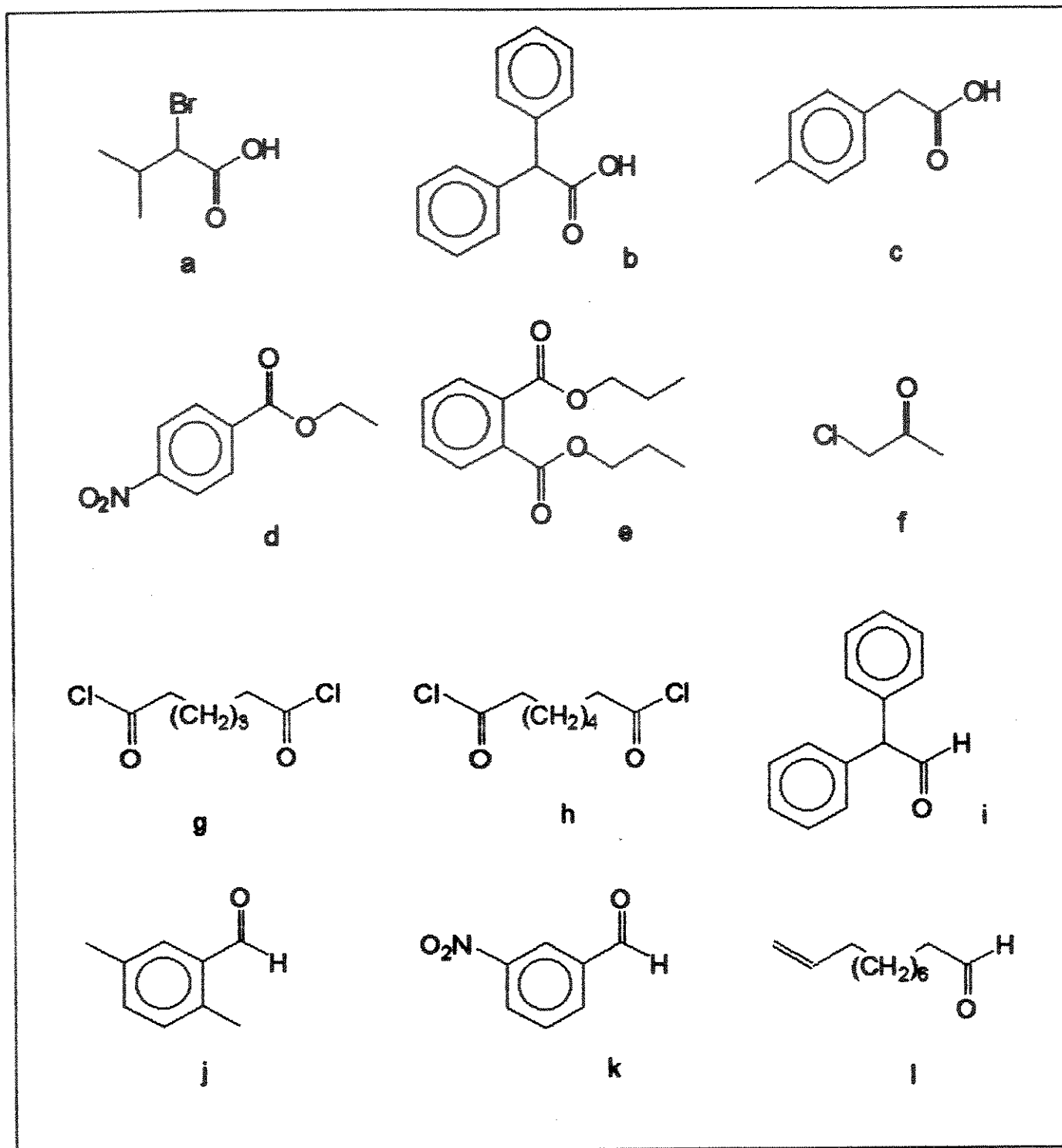


FIGURA IV-36 - Fórmulas estruturais de alguns compostos orgânicos da tabela IV-19, cujos espectros de I.V. foram interpretados pelo programa EXPERTIV.

A partir da implementação das regras que identificam os tipos específicos de um grupo funcional em análise, descritas anteriormente no INTERMOD, foi observado que o programa tem capacidade para retomar grupos funcionais mais prováveis de compostos polifuncionais. O desempenho do programa na interpretação destes compostos foi considerado satisfatório, mencionando-se que a base de conhecimento "*dba_links*" não contém as informações (fatos) necessárias para a identificação de todos os compostos polifuncionais atualmente existentes na literatura. A tabela IV-20 indica os resultados obtidos das interpretações dos espectros de I.V. de alguns compostos polifuncionais, relacionando o nome do composto, os grupos funcionais, os tipos específicos e outros grupos identificados pelo programa como mais prováveis. Na figura IV-37 consta as fórmulas estruturais de alguns compostos relacionados na referida tabela.

Numa avaliação mais ampla do desempenho do programa EXPERTIV quanto ao retorno de grupos funcionais carbonilados identificados como mais prováveis, mas que na realidade não estão presentes ao composto em análise, revela que este fato se deve principalmente a não implementação proposital de regras para eliminá-los. Este procedimento não seria uma tarefa difícil. Por exemplo, na interpretação do espectro de um ácido carboxílico em que o programa identifica o próprio ácido, a cetona e o éster, é possível efetuar o descarte dos dois últimos grupos funcionais, através de regras que procedem como um especialista da área, ou seja, identificada uma banda de absorção larga e intensa correspondendo a vibração de estiramento O-H, o composto tem uma maior probabilidade de ser um ácido carboxílico, retornando então, este último como grupo funcional mais provável. No entanto, o programa foi desenvolvido para ser aplicado em laboratórios que fazem uso da espectroscopia no I.V., com compostos orgânicos desconhecidos. Desta forma, após consultas com professores da área de química orgânica, decidiu-se que a implementação de regras para descartar grupos funcionais obtidos como mais prováveis, poderiam resultar em interpretações incorretas, como por exemplo se um composto a ser analisado é polifuncional.

TABELA IV-20 - Resultados obtidos das interpretações dos espectros de I.V. de alguns compostos polifuncionais, que o programa EXPERTIV identificou os grupos funcionais, os seus tipos específicos e outros grupos como mais prováveis. G.F. = grupos funcionais mais prováveis, * as letras correspondem às fórmulas estruturais representadas na figura IV-37.

Composto	G.F.	Tipos Específicos	Outros Grupos
formiato de isobutila ^{(a)*}	éster, aldeído, cetona	éster: (Ar-CO-O-Ar), (β -ceto-éster(ceto)), (lactona 6 memb.) cetona: (-CX ₂ -CO), (-CX-CO-CX), (α -X)	-
ftalato de mono-metila ^(b)	ácido carboxílico, éster e cetona	ácido carboxílico: (α -Br ou α -Cl dímero) éster: (Ar-CO-O-), (-C=C-CO-O-) (α, β -lactona 6 memb.), (β -ceto-éster(ceto))	NO ₂ , fluoreto e cloreto
tartarato de dietila ^(c)	ácido carboxílico, éster e cetona	ácido carboxílico: (α -Br ou α -Cl dímero) éster: (α -pirona), (-C=C-CO-O-), (Ar-CO-O-), (α, β -lactona 6 memb.), (β -ceto-éster(ceto))	aromático, fluoreto
clorofornato de p-nitrobenzofila ^(d)	éster, cetona e haleto de acila	éster: (α, β -lactona 5 memb.), (O=C-O-CH ₂ -C ₆ H ₆), (α -halo) haleto de acila: (-CO-Cl insaturado)	aromático, NO ₂
4-cloroformilbutirato de metila ^(e)	ácido carboxílico, cetona e haleto de acila	ácido carboxílico: (-COOH dímero) cetona: (-C-(C))-C-) aromático haleto de acila: (-CO-Cl saturado)	-
clorofornato de 2,2,2, tricloroetila ^(f)	éster, cetona e haleto de acila	éster: (β, γ -lactona 5 memb.) haleto de acila: (-CO-Cl insaturado)	-
clorofloxilato de n-etila ^(g)	anidrido de ácido, éster, cetona e haleto de acila	éster: (α -halo), (-CO-O-C=C), (α, β -lactona 5 memb.), (γ, δ -lactona 6 memb.) haleto de acila: (-CO-Cl saturado)	-
formiato de n-propila	éster, aldeído, cetona	éster: -C-C-CO-O-), (Ar-CO-O-), (α, β -lactona de 6 membros) cetona: (-CX ₂ -CO e -CX-CO-CX), (α -halo)	aromático, fluoreto
formiato de n-hexila	éster, aldeído, cetona e haleto de acila	éster: (Ar-CO-O-Ar), (β -ceto-éster(ceto)), (lactona de 6 membros), (α -pirona) cetona: (-CX ₂ -CO e -CX-CO-CX), (α -halo)	-
clorofornato de n-butila	éster, cetona e haleto de acila	éster: (β, γ -lactona 5 memb.), (α, β -lactona 5 memb.) cetona: (4 memb.), (-CO-CO-) 5 memb.)	-
clorofornato de fenila	éster, cetona e haleto de acila	haleto de acila: (-CO-Cl saturado)	aromático

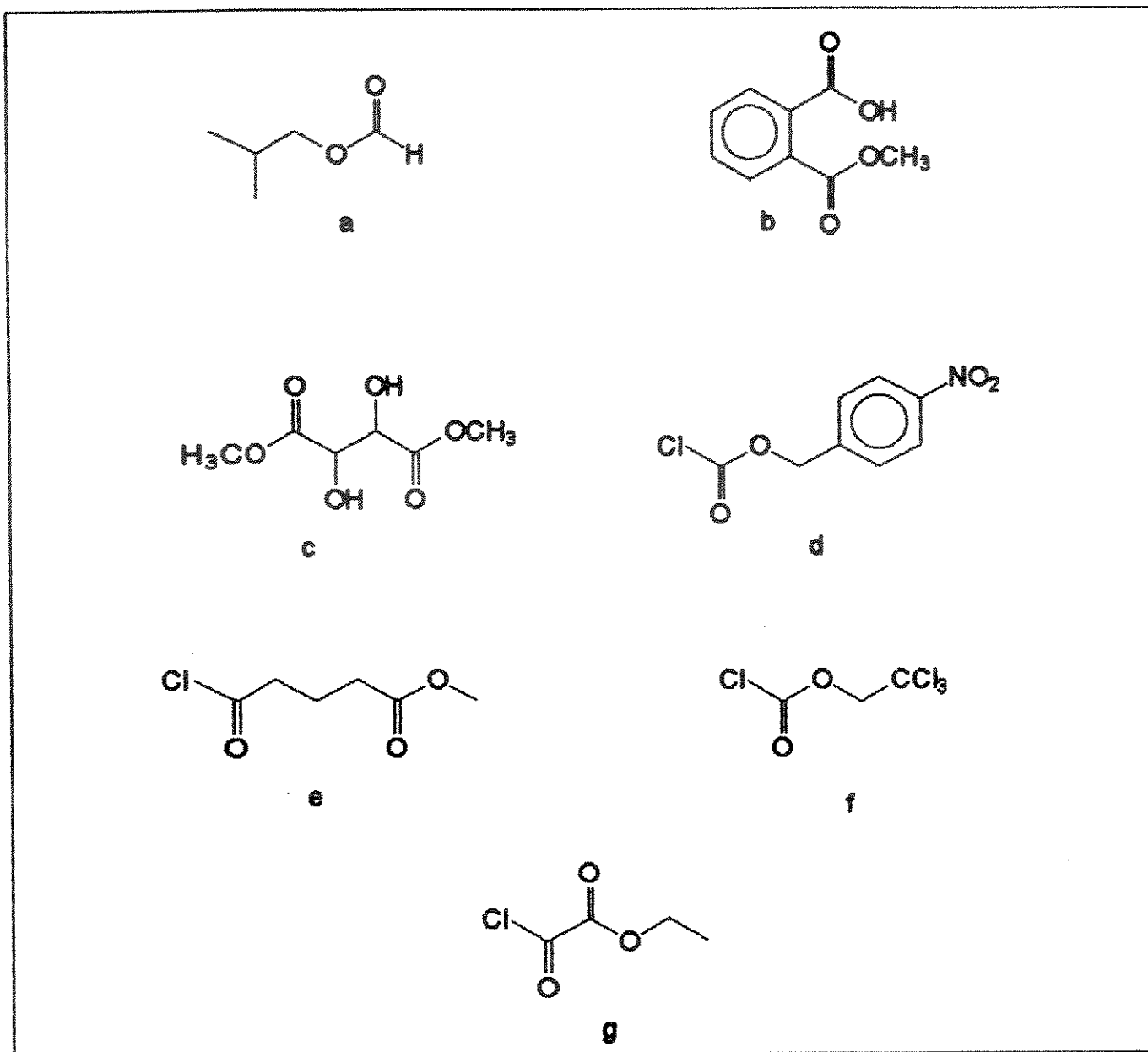


FIGURA IV-37 - Fórmulas estruturais de alguns compostos polifuncionais da tabela IV-20, que foram interpretados pelo programa EXPERTIV, através dos seus espectros de I.V..

Conforme informações a respeito da existência de um programa que atua na forma de busca em biblioteca, instalado no espectrofotômetro Perkin Elmer, modelo 1600, de um laboratório deste instituto, decidiu-se que seria importante uma avaliação do seu desempenho e do programa EXPERTIV, quanto as interpretações dos espectros de I.V.. Basicamente, o programa de busca em biblioteca apresenta na tela do monitor, entre outros recursos, menus para o usuário selecionar as opções desejadas para a interpretação do espectro de I.V., como por exemplo: a) as propriedades químicas do composto (se tem halogênio, N, P, S, etc., b) estado físico da amostra (sólida, líquida e vapor) e, c) desconhecida, caso não se conheça as características da amostra. Após essas seleções, o usuário pode solicitar a interpretação do espectro, de modo que o programa retorna os grupos funcionais ou os compostos mais semelhantes aos encontrados na biblioteca. Além dessas informações, também é retornado um número como referência para o usuário consultar separadamente nos manuais que fazem parte do banco de dados, informações relacionadas as identificações feitas pelo programa de busca em biblioteca.

A avaliação dos resultados foram feitas através das interpretações de tres espectros de I.V. existentes na biblioteca do referido espectrofotômetro, para o hexanaldeído, metiletilcetona e benzofenona. Os resultados obtidos pelo programa EXPERTIV foram considerados muito significativos, ou seja, o programa identificou corretamente todos os grupos funcionais e outros grupos presentes nesses compostos. O programa de busca em biblioteca não identificou o hexanaldeído, mas, também, retornou informações significativas para os outros compostos. A tabela IV-21 mostra os resultados obtidos das interpretações desses espectros, com os programas de busca em biblioteca e EXPERTIV.

TABELA IV-21 - Resultados obtidos das interpretações de tres espectros de I.V. com o programa de busca em biblioteca e o programa EXPERTIV. Os números representam as referências para consultar o manual do programa de busca em biblioteca e os valores dos "scores" calculados pelo programa EXPERTIV.

hexanaldeído
Programa de Busca em Biblioteca
202 grupo alquila - geral
204 grupo alquila - substituinte de cadeia longa
208 grupo alquila - substituinte metil
402 composto hidroxi ou amino - geral
1416 ácido carboxílico ou éster - hidróxi substituído
4906 composto carbonilado - classe 6 (consultar manual)
4928 composto carbonilado - alifático com cadeia longa
4934 composto carbonilado - éster ou ácido com hidróxi substituído
Programa EXPERTIV
cetona (-CO-): 100
aldeído (-CHO): 98
haleto de acila (-CO-X): 92
metiletilcetona
Programa de Busca em Biblioteca
201 grupo alquila - geral
203 grupo alquila - possivelmente substituinte de cadeia pequena
1001 cetona alifática - geral
1003 cetona com metil substituído
4902 composto carbonilado - classe 2 (consultar manual)
4925 composto carbonilado - possivelmente éster ou cetona
Programa EXPERTIV
cetona (-CO-): 100
tipos específicos:
éster(-C=C-CO-O-), ou
(Ar-CO-O-), ou (α - β -lactona 6 membros), ou (β -ceto-éster (ceto)
cetona(α -X) (Cl-C-C=CO), ou
(α -X) (Br-C-C=O), ou
(-C-(CO)-C-) alifático
benzofenona
Programa de Busca em Biblioteca
212 composto alifático ou aromático - grupo carbonila de cetona ou amida
256 composto aromático ou possivelmente insaturado
294 composto aromático substituído (consultar manual)
3107 cetona conjugada - composto quinona ou aromático
3116 cetona aromática - possivelmente benzofenona
Programa EXPERTIV
éster (-CO-O): 100
cetona (-CO-): 100
tipos específicos:
éster(β -ceto-éster(enol), ou
(Ar-CO-O), ou (-C=C-CO-O-),
cetona(HO ou NH ₂)-C-C-CO-) livre ou tropona), ou (quinona estendida)

IV.3 - PERSPECTIVAS FUTURAS

Além da resolução de alguns problemas através do programa EXPERTIV, ou seja, a identificação de grupos funcionais mais prováveis de estarem presentes em um determinado composto orgânico, que pode auxiliar o usuário da espectroscopia no I.V., ainda é possível de serem implementadas mais regras para se atingir outras metas mais específicas.

Uma das limitações do presente S.E. anteriormente descrito, está principalmente nos resultados das interpretações dos espectros de I.V. de compostos carbonilados, em que o programa às vezes retorna um ou mais grupos funcionais que não pertencem ao composto em análise. Através do processo de filtragem do INTERMOD, o programa analisa separadamente os quatorze grupos funcionais estudados neste trabalho e, normalmente retorna um ou no máximo quatro grupos funcionais identificados como os mais prováveis.

Conforme estes resultados, pode-se desenvolver um outro módulo com regras que buscam os fatos inseridos nas bases de dados internas "*funcional_group*" e "*tipo_especifico*", isto é, fatos dos grupos funcionais retornados como os mais prováveis e seus respectivos tipos específicos. Entretanto, torna-se necessário a implementação de regras para confirmar as vibrações do tipo específico identificadas no composto, como por exemplo na interpretação do espectro de I.V. do ácido 3-bromopropiônico ($\text{Br-CH}_2\text{-CH}_2\text{-(CO)-OH}$). Para esse composto o programa EXPERTIV retornou como grupos funcionais mais prováveis o ácido carboxílico, a cetona e o éster e, respectivamente os tipos específicos para um ácido α,β -halogenado, uma cetona em anel de seis ou de sete membros e ésteres -C=C-CO-O- ou Ar-CO-O- ou α,β lactona de seis membros ou β -cetoéster. Nesse caso, as regras devem confirmar as informações das vibrações existentes nos tipos específicos identificados, como a vibração de estiramento C-Br para o ácido carboxílico e as demais para os grupos cetônico e éster. Após a confirmação das vibrações

existentes nos tipos específicos anteriormente identificados, o programa retornará o grupo funcional ou grupos mais prováveis com as informações espectrais de todas as ligações do composto analisado. Alguns procedimentos a serem empregados podem ser análogos aos do INTERMOD, como por exemplo a aplicação da função matemática para calcular o "score" da posição, intensidade e largura da banda de absorção.

A partir da identificação do grupo funcional mais provável, a meta do programa pode direcionar para o esclarecimento ou determinação da estrutura do composto analisado. Uma base de dados com informações de outras técnicas espectrométricas como a de massas ou de RMN teria, então, de ser implementada para a equiparação dos fatos obtidos do grupo funcional. A atuação do S.E. sobre estes dados será facilitada grandemente se grupos funcionais prováveis de estarem presentes no composto tiverem sido previamente identificados pelo programa EXPERTIV.

CONCLUSÕES

Deste trabalho pode-se extrair as seguintes conclusões:

A mesa digitalizadora desenvolvida, de baixo custo, é adequada para acessar dados expressos em forma gráfica, digitalizá-los e armazená-los para serem tratados por programas quimiométricos, especialistas ou de construção e manutenção de bibliotecas gráficas.

O programa PROGEIV permite a elaboração de bibliotecas de espectros de menor custo e que podem ser mais úteis ao usuário, uma vez que este pode especificar os espectros que deseja armazenar e transferi-los de bibliotecas tipo Sadler, comumente encontradas nas instituições de pesquisa no Brasil. Também, o programa permite armazenar digitalmente, dados de instrumentos analógicos, que fornecem espectros registrados em papel.

O programa EXPERTIV desenvolvido permite a interpretação de espectros I.V. fornecendo os grupos funcionais mais prováveis de estarem presentes no composto e mostrando os fatos espectroscópicos que originaram as conclusões. O programa apresenta, sobre outros previamente propostos, a vantagem de incorporar um módulo de leitura do espectro. Desta forma, o programa dispensa a intervenção do usuário para o fornecimento das informações necessárias ao módulo de interpretação.

A interpretação, por sua vez, pode agora ser feita de forma automática. Embora tenha-se usado a mesa digitalizadora como ferramenta na obtenção do espectro a ser interpretado, o programa poderia ser empregado em conjunto com um espectrofotômetro I.V., que pudesse fornecer os dados já digitalizados. A estrutura modular do programa permite a fácil alteração da base de conhecimento.

O tempo de processamento necessário para o programa EXPERTIV interpretar um espectro de I.V., é de 20 segundos em média, sendo que o mesmo depende da complexidade do composto e, conseqüentemente, do espectro em análise.

Para os quatorze grupos funcionais estudados neste trabalho, o programa EXPERTTV obteve um índice médio de identificação correta acima de 90 %.

O programa EXPERTTV atua como um filtro, ou seja, descarta os grupos funcionais improváveis de estarem presentes no composto e retém os grupos funcionais que indicam uma maior probabilidade da sua presença no composto. Normalmente, o programa retorna de um a três grupos funcionais identificados como os mais prováveis. Através deste procedimento, o programa tem amplas possibilidades de expansão, podendo inclusive, ser acrescido de bases de conhecimentos de outras técnicas como RMN e massas, objetivando a elucidação da estrutura do composto.

REFERÊNCIAS

1. Klaessens, J. W. A.; Kateman, G. and Vandeginste, B. G. M. "Expert Systems and Analytical Chemistry"; *Trends Anal. Chem.*, 4(5),(1985)114-117.
2. Betteridge, D.; Mackison, R.; Mottershead, C. M.; Taylor, A. F. and Wade, A. P. "Development of an Expert System for the Selection of Sample Points for Moisture Analysis", *Anal. Chem.*, 60(1988),1534-1539
3. Monchy, A. R.; Forster, A. R.; Arretteig, J. R.; Lan Le and Deming, S. N. "Expert Systems for the Analytical Laboratory", *Anal. Chem.*, 60(1988)1355-1361-A.
4. Luinge, H. J. "Automated interpretation of vibrational spectra", *Vibrational Spectroscopy*, 1(1990)3-18.
5. Bridge, T. P.; Williams, M. H. and Fel, A. F. "Analytical Chemistry and Expert System", *Chem. Britain*, (1987)1085-1088.
6. Lee, J. R. and Isenhour, T. L. "An Expert System for Analytical Data Management", *J. Chem. Inf. Comput. Sci.*, 32(1992)148-153.
7. Hibbert, D. B. "A Prolog Program for the Calculation of Isotope Distributions in Mass Spectrometry", *Chemometrics and Intelligent Laboratory Systems*, 6(1989)203-212.
8. Harrington, P. B.; Street, T. E. and K. J. Voorhees*, "Rule-Building Expert System For Classification of Mass Spectra", *Anal. Chem.*, 61(1989)715-719.
9. Zupan, J.; Razinger, M.; Bohanec, S.; Novic, M.; Tusar M. and Lah, L. "Building Knowledge into an Expert System", *Chemometrics and Intelligent Laboratory Systems*, 4(1988)307-314.
10. Zupan, J.; Buydens, L.; Peeters, A. and Massart, D. L. " Expert Systems with Chemical Structure Handling Capability", *Chemometrics and Intelligent Laboratory Systems*, 8(1990)43-51.
11. Buydens, L.; Detaevernier, M.; Tombeur, D. and Massart, D. L. "An Expert System for the Development of Analytical Procedures UV Spectrophotometric Determination of Pharmaceutically Active Substances in Tablets", *Chemometrics and Intelligent Laboratory Systems*, 1(1986)99-108.

12. Janssens, K. and Van Espen, P. "Implementation of an Expert System for The Qualitative Interpretation of X-Ray Fluorescence Spectra", *Anal. Chim. Acta*, **184** (1986)117-132.
13. Janssens, K.; Dorrine, W. and Van Espen, P. "The Development Process of an Expert System for the Automated Interpretation of Large EPMA Data Sets", *Chemometrics and Intelligent Laboratory Systems*, **4**(1988)147-161.
14. Dessy, R. E. "Expert System - Part I", *Anal. Chem.*, **56**(1984)1200-1212A.
15. Dessy, R. E. "Expert System - Part II", *Anal. Chem.*, **56**(1984)1312-1331A.
16. Wade, A.P.; Crouch, S. R. and Betteridge, D. "Development of Microcomputer-Based Expert Systems for Analytical Chemistry", *Trends Anal. Chem.*, **7**(10),(1988)358-365.
17. Settle, F. A. Jr.; Diamondstone, B. I.; Kingston, H. M. and Pleva, M. A. "An Expert-Database System for Sample Preparation by Microwave Dissolution. 1. Selection of Analytical Descriptors", *J. Chem. Inf. Comput. Sci.*, **29**(1989)11-17.
18. Greathouse, D.; Clements, J. and Morris, K. "The Use of Expert Systems to Assist in Decisions Concerning Environmental Control", *Critical Reviews in Environmental Control*, **19**(4),(1989)341-357.
19. Moldoveanu, S. and Rapson, C. A. "Spectral Interpretation for Organic Analysis Using an Expert System", *Anal. Chem.*, **59**(1987)1207-1212.
20. Bormam, S. A. "Expert Systems for Liquid Chromatography", *Anal. Chem.*, **58**(1986) 1192-1200A.
21. Moors, M. and Massart, D. L. "A small expert system for solid phase extraction", *Trends Anal. Chem.*, **9**(5)(1990)164-169.
22. Schoenmakers, P. J.; Dunand, N.; Cleland, A.; Musch, G. and Blaffert, Th. "An Expert System for the Optimization of Columns, Operating Conditions and Instrumentation for High-Pressure Liquid Chromatography", *Chromatographia*, **26**(1988)37-44.
23. van Leeuwen, J. A.; Buydens, L. M. C.; Vandeginste, B. G. M. and Kateman, G. "Expert Systems in Chemical Analysis", *Trends Anal. Chem.*, **9**(2),(1990) 49-54.
24. Buydens, L. M. C.; van Leeuwen, J. A.; Mulholland, M.; Vandeginste, B. G. M. and Kateman, G. "An Expert System for the Validation of High-Performance Liquid, Chromatographic Methods", *Trends Anal. Chem.*, **9**(2),(1990)58-62.

25. Schoenmakers, P. J.; Dunand, N.; Cleland, A.; Musch, G. and Blaffert, Th. "An Expert System for the Optimization of Columns, Operating Conditions and Instrumentation for High-Pressure Liquid Chromatography", *Chromatographia*, 26(1988)37-44.
26. Peeters, A.; Buydens, L.; Massart, D. L. and Schoenmakers, P. J. "An Expert System for the Selection of Criteria for Selectivity Optimization in High-Pressure Liquid Chromatography", *Chromatographia*, 26(1988)101-109.
27. Buydens, L.; Peeters, A. and Massart, D. L. "Knowledge Representation for the Selection of Methods in Analytical Chemistry" *Chemometrics and Intelligent Laboratory Systems*, 5(1988)73-79.
28. Williams, S. S. "Application of expert systems programming to high-performance liquid chromatography", *Trends Anal. Chem.*, 9(2),(1990)63
29. Tischler, M. A. and Fox, E. A. "An Expert System for Selecting Liquid Chromatographic Separation Methods", *Comput. Chem.*, 11(4),(1987)235-240
30. Gunasingham, H.; Srinivasan, B. and Ananda, A. L. "Design of a Prolog-Based Expert System for Planning Separation of Steroids by High-Performance Liquid Chromatography", *Anal. Chim. Acta*, 182(1986)193-202
31. Hippe, Z. "Problems in the Application of Artificial Intelligence in Analytical Chemistry", *Anal. Chim. Acta*, 150(1983)11-21.
32. Delaney, M. F.; Warren, F. V. Jr.; and Hallowell, J. R. Jr. "Quantitative Evaluation of Library Searching Performance", *Anal. Chem.*, 55(1983)1925-1929
33. Lowry, S. R.; Huppler, D. A. and Anderson, C. R. "Data Base Development and Search Algorithms for Automated Infrared Spectral Identification", *J. Chem. Inf. Comput. Sci.*, 25(1985)235.
34. Ruprecht M. and Clerc, J. T. "Performance Analysis of a Simple Infrared Library Search Systems", *J. Chem. Inf. Comput. Sci.*, 25(1985)241.
35. Harrington P. B. and Isenhour, T. L. "A Quantitative Measure of the Reliability of Searches of Spectral Libraries", *Anal. Chim. Acta*, 197(1987)105-119.
36. Dupuis, P. F.; Cleij, P.; Van'T Klooster, H. A. and Dijkstra, A. "Information Theory Applied to Feature Selection of binary-Coded Infrared Spectra for Automated Interpretation by Retrieval of Reference Data", *Anal. Chim. Acta*, 112(1979)83-93.

37. Zupan, J. and Munk, M. E. "Hierarchical Tree Based Storage, Retrieval, and Interpretation of Infrared Spectra", *Anal. Chem.*, **57**(1985)1609-1616.
38. Hangac, G.; Wieboldt, R. C.; Lam, R. B. and Isenhour, T. L. "Compression of an Infrared Spectral Library by Karhunen-Loeve Transformation", *Appl. Spectrosc.*, **36**(1),(1982) 40-47.
39. Settle, F. A. Jr. and Pleva, M. A. "Expert system development tools for chemists", *Chemometrics and Intelligent Laboratory Systems*, **11**(1991)13-26.
40. Harrington, P. B. and Isenhour, T. L. "Compression of Infrared Libraries by Eigenvector Projection", *Appl. Spectrosc.*, **41**(3),(1987)449-453.
41. Bjerga, J. M. and Small, G. W., "Automated Selection of Library Subsets for Infrared Spectral Searching", *Anal. Chem.*, **62**(1990)226-233.
42. Marengo M. E. and Salles, H. "Spectre: An Approach to an Expert System in Chemometrics", *Anal. Chim. Acta*, **210**(1988)33-50.
43. Yuzhu, H. Peeters, A.; Musch, G. and Massart, D. L. "Integration of Optimization Methodology with Expert Systems", *Anal. Chim. Acta*, **223**(1989)1-17.
44. Hasenoehrl, E. J.; Perkins, J. H. and Griffiths, P. R., "Expert System Based on Principal Components Analysis for the Identification of Molecular Structures from Vapor-Phase Infrared Spectra. 2. Identification of Carbonyl-Containing Functionalities.", *Anal. Chem.*, **64**(1992)656-663.
45. Lindsay, R. K.; Buchanan, G. B.; Feigenbaum, E. A. and Lederberg, J. "Applications of Artificial Intelligence for Organic Chemistry - The Dendral Project"- McGraw-Hill, New York, 1980.
46. Gray, N. A. B. "Dendral and Meta-Dendral - The Myth and the Reality", *Chemometrics and Intelligent Laboratory Systems*, **5**(1988)11-32.
47. Luinge, H. J. and Van Der Maas, J. H. "Artificial Intelligence for the Interpretation of Combined Spectral Data", *Anal. Chim. Acta*, **223**(1989)135-147.
48. Shelley, C. A and Munk, M. E. "CASE, A Computer Model of the Structure Elucidation Process", *Anal. Chim. Acta*, **133**(1981)507-516.

49. Gribov, L. A.; Elyashberg, M. E. and Raikhshtat, M. M. "A New Approach to the Determination of Molecular Spatial Structure Based on the of Spectra and Computers", *J.Mol.Struct.*, **53**(1979)81-96.
50. Huixiao, H. and Xinquan, X. "ESSESA: An Expert System for Elucidation of Structures from Spectra. 1. Knowledge Base of Infrared Spectra and Analysis and Interpretation Programs", *J.Chem.Inf.Comput.Sci.*, **30**(1990)203-210.
51. Huixiao, H. and Xinquan X. "ESSESA: An Expert System for Structure Elucidation from Spectra. 3. LNSCS for Chemical Knowledge Representation.", *J.Chem.Inf.Comput.Sci.*, **32**(1992)116-120.
52. Huixiao H. and Xinquan X. "ESSESA, an expert system for structure elucidation from spectral analysis", *Anal.Chim.Acta*, **262**(1992)179.
53. Farkas, M.; Markos, J.; Szepesvary, P.; Bartha, I.; Szalontai, G. and Simon Z. "A Computer-Aided System for Organic Functional Group Determinations", *Anal.Chim.Acta*, **133**(1981)19-29.
54. Funatsu, K.; Susuta, Y. and Sasaki, S. "Application of Infrared Data Analysis Based on Symbolic Logic in Automated Structure Elucidation by Chemics", *Anal.Chim.Acta*, **220**(1989)155-169.
55. Hidetsugu Abe, Tohru Okuyama, Iwao Fujiwawara, and Sasaki, S. "A Computer Program for Generation of Constitutionally Isomeric Structural Formulas", *J.Chem.Inf.Comput.Sci.*, **24**(1984)220-229.
56. Kimito Funatsu, Nobuyoshi Miyabayashi, and and Sasaki, S. "Further Development of Structure Generation in the Automated Structure Elucidation System" *J.Chem.Inf.Comput.Sci.*, **28**(1988)18-28.
57. Debska, B.; Duliban, J.; Guzowska-Swider, B. and Hippe, Z., "Computer-Aided Structural Analysis of Organic Compounds by an Artificial Intelligence System", *Anal.Chim.Acta*, **133**(1981)303-318.
58. Woodruff, H. B. and Smith, G. M., "Computer Program for the Analysis of Infrared Spectra", *Anal.Chem.*, **52**(1980)2321-2327.
59. Tomellini, S. A.; Stevenson, J. M. and Woodruff, H. B., "Rules for Computerized Interpretation of Vapor-Phase Infrared Spectra", *Anal.Chem.*, **56**(1984)67-70.

60. Tomellini, S. A. and Hartwick, R. A. "Automated Rule Generation for the Program for the Analysis of Infrared Spectra (PAIRS)", *Anal. Chim. Acta*, **162**(1984)227-240
61. Tomellini, S. A.; Hartwick, R. A. and Woodruff, H. B. "Automatic Tracing and Presentation of Interpretation Rules Used by PAIRS: Program for the Analysis of IR Spectra", *Appl. Spectroscopy*, **39**(2),(1985)331-333
62. Ying, Li-Shi and Levine, S. P. "Self-Training, Self-Optimizing Expert System for Interpretation of the Infrared Spectra of Environmental Mixtures", *Anal. Chem.*, **59** (1987)2197-2203.
63. Puskar, M. A. and Levine, S. P. "Qualitative Screening of Hazardous Waste Drum Mixtures", *Environ. Sci. Technol.*, **21**(1987)90-96.
64. Puskar, M. A. and Levine S. P. "Infrared Screening Technique for Automated Identification of Bulk Organic Mixtures", *Anal. Chem.*, **58**(1986)1981-89.
65. Puskar, M. A. and Levine, S. P. "Computerized Infrared Spectral Identification of Compounds Frequently Found at Hazardous Waste Sites", *Anal. Chem.*, **58**(1986) 1156-62.
66. Smith, G. M.; Woodruff, H. B.; Tomellini, S. A.; Saperstein, D. D.; Stevenson, J. M. "Automated Interpretation of Infrared Spectra with an Instrument Based Minicomputer", *Anal. Chem.*, **53**(1981)2367-2369
67. Woodruff, H. B. and Smith G. M., "Generating Rules for PAIRS - A Computerized Infrared Spectral Interpreter", *Anal. Chim. Acta*, **133**(1981)545-553.
68. Woodruff, H. B. and Munk, M. E. "Computer-Assisted Interpretation of Infrared Spectra", *Anal. Chim. Acta*, **95**(1977)13-23
69. Woodruff, H. B. and Munk, M. E. "A Computerized Infrared Spectral Interpreter as a Tool in Structure Elucidation of Natural Products", *J. Org. Chem.*, **42**(10),(1977)1761-1767.
70. Trulson, M. O. and Munk M. E., "Table-Driven Procedure for Infrared Spectrum Interpretation", *Anal. Chem.*, **55**(1983)2137-2142.
71. Frankel, D. S. "Pattern Recognition of Fourier Transform Infrared Spectra of Organic Compounds", *Anal. Chem.*, **56**(1984)1011-14.

72. Blaffert, Th. "EXPERTISE - An Expert System for Infrared Spectra Evaluation", *Anal. Chim. Acta*, 191(1986)161-168.
73. Luinge, H. J.; Kleywegt, G. H.; van't Klooster, H. A. and Van Der Maas, J. H. "Artificial Intelligence Used for the Interpretation of Combined Spectral Data. 3. Automated Generation of Interpretation Rules for Infrared Spectral Data", *J. Chem. Inf. Comput. Sci.*, 27(1987)95-99.
74. Luinge, H. J. and van der Maas, J. H., "AEGIS, an Algorithm for the Exhaustive Generation of Irredundant Structures", *Chemometrics and Intelligent Laboratory Systems*, 8(1990)157-165.
75. Luinge, H. J. "A Knowledge-based system for structure analysis from infrared and mass spectral data", *Trends Anal. Chem.*, 9(2),(1990)66-69.
76. Koeijer, J. A.; Luinge, H. J. and van der Maas, J. H. "Application of an automated interpretation system for infrared spectra - Part I. Automated qualitative analysis of polymers", *Vibrational Spectroscopy*, 4(1993)285-299.
77. Luinge, H. J.; Koeijer, J. A. and van der Maas, J. H. "Application of an automated interpretation system for infrared spectra - Part II. Characterization of aromatic sulphone copolymers", *Vibrational Spectroscopy*, 4(1993)301-308.
78. Lerner, J. M. and Lu, T. "Practical Neural Networks Aid Spectroscopic Analysis", *Photonics Spectra*, (1993)93-98.
79. Robb, E. W. and Munk, M. E. "A Neural Network Approach to Infrared Spectrum Interpretation", *Mikrochim. Acta*, 1(1990)131-155.
80. Munk, M. E.; Madison, M. S. and Robb, E. W. "Neural Network Models for Infrared Spectrum Interpretation", *Mikrochim. Acta*, 2(1991)505-514.
81. Meyer M. and Weigelt T. "Interpretation of Infrared Spectra by Artificial Neural Networks", *Anal. Chim. Acta*, 265(1992)183-190.
82. Wythoff, B. J.; Levine, S. P. and Tomellini, S. A. "Spectral Peak Verification and Recognition Using a Multilayered Neural Network", *Anal. Chem.*, 62(1990)2702-2709.
83. Gastmans, J. P.; Furlan, M.; Lopes, M. N.; Borges, J. H. G. e Emerenciano, V. P. "A Inteligência Artificial Aplicada a Química de Produtos Naturais", *Química Nova*, 13(1990)10-16.

84. Gastmans, J. P.; Furlan, M.; Lopes, M. N.; Borges, J. H. G. e Emerenciano, V. P. "A Inteligência Artificial Aplicada à Química de Produtos Naturais. O Programa SISTEMAT - Parte II. Organização do Programa e Aplicativos", *Química Nova*, 13(1990)75.
85. Emerenciano, V. P. "Determinação Estrutural Computadorizada, Princípios, Métodos, Perspectivas e Mitos", *Química Nova*, 16(1993)551-559.
86. Steel, "*Common Lisp*"
87. "*Turbo Prolog - User's Guide*" - Versão 2.0 - Borland International, Inc. 1988.
88. "*Turbo Prolog - Reference Guide*" - Versão 2.0 - Borland International, Inc. 1988.
89. Clocksin, W. F.; Mellish, C. S. "*Programming in PROLOG*", 2a Springer-Verlag, Berlin Heidelberg New York Tokyo, 2^o- 1Ed., 1984.
90. Smith, P. "*Expert System Development in Prolog and Turbo-Prolog*", John Wiley & Sons Ltd., New York, 1988.
91. Robinson, P. R. "*Turbo Prolog - Guia do Usuário*", McGraw-Hill, 1987.
92. Cabrol, D. and Forrest, T. P. "A Different Kind of Language: Prolog, Programing in Logic", *J. Chem. Educ.*, 63(2),(1986)131-135.
93. Kleywegt, G. H.; Luinge H. J. and Schuman, B. J. P. "Prolog for Chemists. Part 1", *Chemometrics and Intelligent Laboratory Systems*, 4(1988)273-297
94. Kleywegt, G. H.; Luinge H. J. and Schuman, B. J. P. "Prolog for Chemists. Part 2 **", *Chemometrics and Intelligent Laboratory Systems*, 5(1989)117-128.
95. Musch, G.; Thielemans, A. and Massart, D. L. "PROLOG: Arithmetic Relations", *Trends Anal. Chem.*, 5(3),(1986)58-59.
96. Massart, D. L. and Smeyers-Verbeke, J. "PROLOG, An Artificial Intelligence Language", *Trends Anal. Chem.*, 4(2),(1985)50-51.
97. Massart, D. L. and De Smet, M. "PROLOG, Conjunctive Queries and Rules", *Trends Anal. Chem.*, 4(5),(1985)111-112.
98. Postma, G. J.; Vandeginste, B. G. M.; van Halen, C. J. G. and Kateman, G. "A Comparison of Lisp and Prolog", *Trends Anal. Chem.*, 6(1),(1987)2-5.

99. Postma, G. J.; Vandeginste, B. G. M.; van Halen, C. J. G. and Kateman, G. "Implementation of a Teaching Program for IR Spectrometry in Lisp and Prolog", *Trends Anal Chem.*, 6(2),(1987)27-30.
100. Souza, P. S. and Pasquini, C. "Simple Graphical Data Tablet for Spectral Input", *Laboratory Microcomputer*, 9(1),77-80(1990).
101. Digigraf Mesas Digitalizadoras Manual do Uusário - versão PAD-5/90.
102. "Microsoft. *QuickBasic - Programming in Basic*", Microsoft Corporation, 1987-1988.
103. Biblioteca SADTLER Research Laboratories Inc.
104. Silverstein, R. M.; Bassler, G. C.; Morrill, T. C. "*Identificação Espectrométrica de Compostos Orgânicos*", Ed. Guanabara.
105. Bellamy, L. J. "The IR of Complex Molecules", 3rd ed., John Wiley & Sons Inc., New York, 1975.
106. Nakanishi, K. and Solomon, P. H. "Infrared Absorption Spectroscopy" 1, 2nd ed, Holden-Day, Inc., San Francisco, 1.
107. Otto, M. "Fuzzy Theory Explained", *Chemometrics and Intelligent Laboratory Systems*, 4(1988)101-120.
108. Otto, M. "Fuzzy Expert Systems", *Trends Anal Chem.*, 9(2),(1990)69-72.
109. Zadeh, L. A. "Yes, no, and relatively - Part 1", *Chemtech*, June(1987)340-345.
110. Zadeh, L. A. "Yes, no, and relatively - Part 2", *Chemtech*, July(1987)406-410.
111. Blaffert, Th. "Computer-Assisted Multicomponent Spectral Analysis with Fuzzy Data Sets", *Anal. Chim. Acta*, 161(1984)135-148.
112. Cunha, I. B. S. and Pasquini, C. "Automatic gravimetric management of solutions. I. High performance microcomputer controled gravimetric burette", *Analyst*, 117(1992) 905-911.
113. Malcome-Lawes, D. J.; Pasquini, C. and Wong, K. H. "Instrumentation based on microcomputers. Part 5: A computer operated flow injection analyser, *Lab. Micro.*, 8(1989)44-52.

114. Malcome-Lawes, D. J.; Meyers, J. and Pasquini C. "Instrumentation Based on Microcomputers. Part 4: A computer controlled visible-light absorbance monitor", *Lab. Micro.*, 7(1988)89-95.

Palavras Chaves com as suas respectivas referências:

- a) Sistemas Especialistas: 1-3, 5-6, 8-16, 18-26, 28-30, 39, 42-44, 50-52, 62, 72, 90 e 108.
- b) Prolog: 7, 30, 87-99.
- c) Infravermelho: 33-34, 36-38, 40-41, 44, 50, 54, 58-60, 62, 64-73, 75-77, 79-81, 106.
- d) Elucidação: 48, 50-52, 54, 56 e 69.
- e) Identificação: 33, 44 e 64-65.

APÊNDICE A

MESA DIGITALIZADORA - HM

A.1 - PARTE MECÂNICA

A parte mecânica da mesa digitalizadora desenvolvida tem como base uma placa de acrílico de 27 cm de largura por 38 cm de comprimento, e 1,2 cm de espessura. Dois suportes metálicos constituídos de duas hastes cilíndricas de aço inoxidável com 6 mm de diâmetro foram presos na parte superior e inferior desta placa de acrílico.

Estas hastes atuam como guias horizontais de uma régua também feita de acrílico de 6 mm de espessura. Esta régua pode percorrer toda a extensão horizontal da mesa. Sobre esta régua encontra-se uma outra haste de metal, com as mesmas características das anteriores, que atua como guia vertical de um cursor feito também de acrílico. O cursor pode ser movido por toda a extensão vertical da mesa. Na parte inferior do cursor uma pequena placa de acrílico de 0,5 mm de espessura contém um alvo em forma de cruz que é utilizado para acompanhar o traço gráfico que deseja-se digitalizar.

Os transdutores para as posições horizontal e vertical do cursor são constituídos por dois resistores lineares colocados na posição horizontal, na mesa, e na régua, em posição vertical.

Os resistores foram construídos sobre uma base constituída de uma haste cilíndrica de metal com 8 mm de diâmetro, recoberta com tinta de resina epóxi. Com o uso de um torno mecânico, utilizando-se um passo de 0,18 mm por volta, enrolou-se sobre estas hastes um fio de níquel-cromo referência 80, com uma resistividade nominal igual a $224 \Omega \cdot m^{-1}$. A figura A-1 mostra em perspectiva, a parte mecânica do digitalizador desenvolvido.

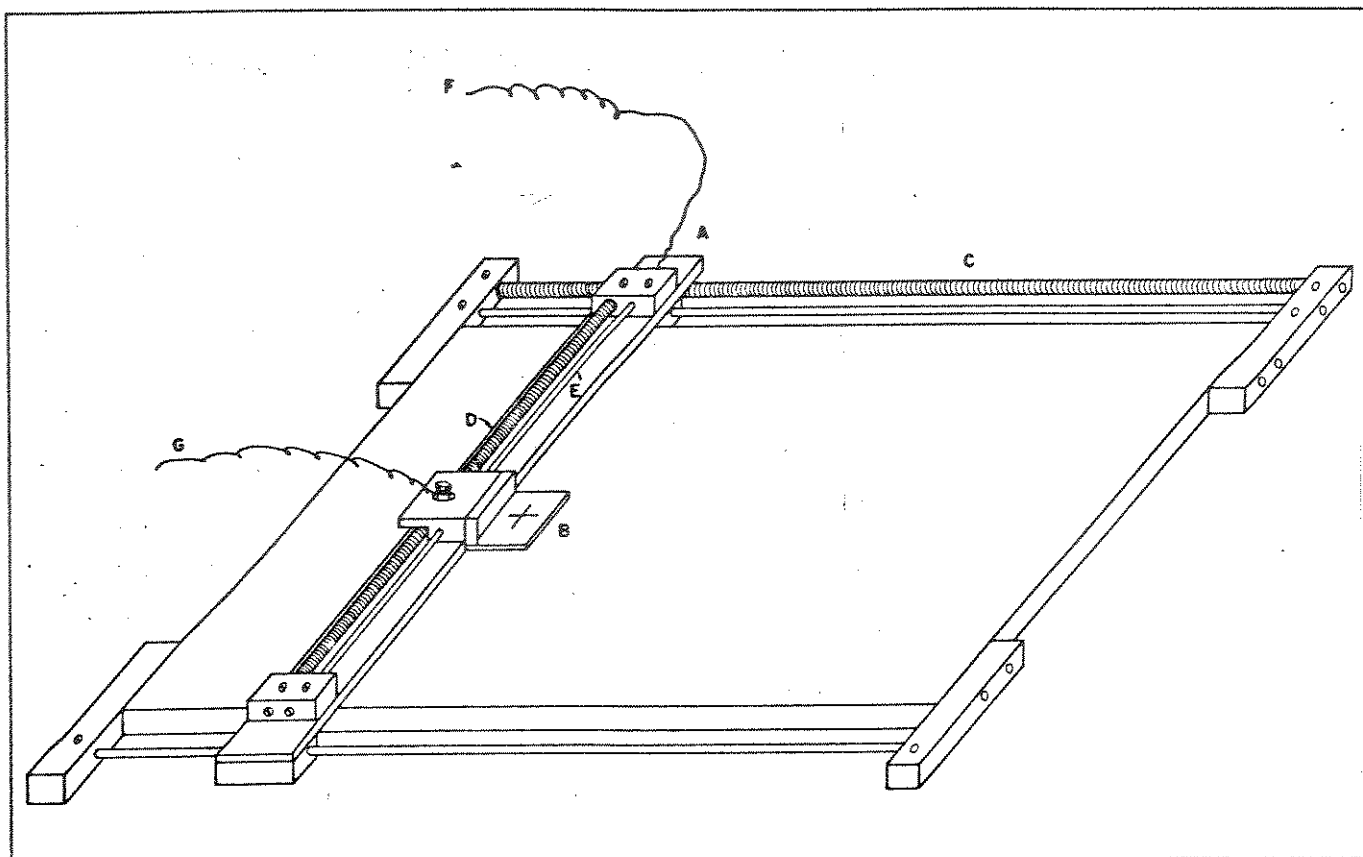


FIGURA A-1 - Vista em perspectiva da parte mecânica da mesa digitalizadora. A - régua de acrílico, B - cursor de acrílico com um alvo na forma cruz, C e D - resistores lineares, E - haste metálica, F e G - contatos elétricos para os transdutores.

Os contatos elétricos adaptados no cursor sobre o resistor vertical e na régua sobre o resistor horizontal, permitem que os valores de resistência elétrica medidos entre as extremidades dos resistores e, as posições horizontal e vertical do cursor sejam obtidos. A pressão dos contatos sobre os resistores é controlada com o auxílio de uma pequena mola e de um parafuso como mostrado na figura A-2. Os fios flexíveis, ligam os contatos e as extremidades dos resistores até o circuito transdutor e, deste, até o conversor analógico/digital. Os guias, horizontal e o vertical, são lubrificados com graxa de silicone. Desta forma, os contatos elétricos deslizam suavemente sobre os respectivos resistores.

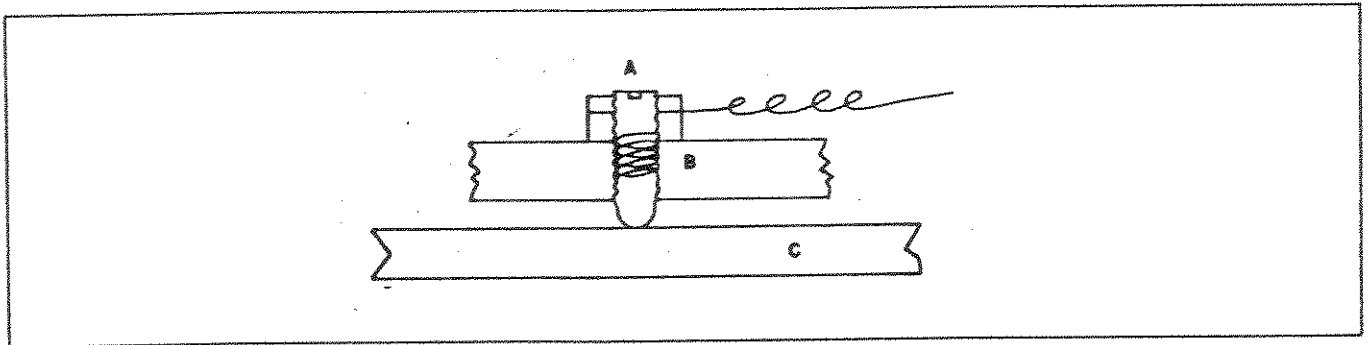


FIGURA A-2 - Detalhes do contato elétrico deslizante utilizado no digitalizador. A - parafuso para pressionar os resistores, B - mola metálica e C - resistor linear.

A.2 - PARTE ELETRÔNICA

As quatro partes principais da eletrônica da mesa digitalizadora são:

- A.2.1 - Transdutor Analógico de Posição.
- A.2.2 - Porta do Usuário para Microcomputadores IBM-XT ou AT.
- A.2.3 - Circuito Decodificador de Endereços.
- A.2.4 - Conversor Dinâmico Analógico/Digital de dois Canais.

A.2.1 - Transdutor Analógico de Posição

O circuito transdutor das coordenadas horizontal do cursor da mesa digitalizadora está descrito na figura A-3. Dois canais independentes monitoram as coordenadas do cursor, valendo-se do fato que, os ganhos dos amplificadores operacionais são efetivamente definidos entre suas extremidades e a posição vertical do cursor e horizontal da régua. Os valores da tensão a ser amplificada (definida pelos potenciômetros)

e os resistores da malha de realimentação dos operacionais, podem ser modificados, de forma a originar uma tensão adequada a ser apresentada ao sistema de conversão analógico/digital. A configuração descrita na figura A-3, origina uma variação de aproximadamente 1,0 V, quando a régua percorre toda a extensão horizontal e, quando o cursor percorre toda a extensão vertical útil. A estas variações de tensão correspondem uma variação de resistência elétrica de aproximadamente 10,0 K e 7,4 K, respectivamente, para os movimentos horizontal e vertical. A sensibilidade estimada para deslocamentos verticais e horizontais é de 34 mV/cm e de 5 mV/cm, respectivamente.

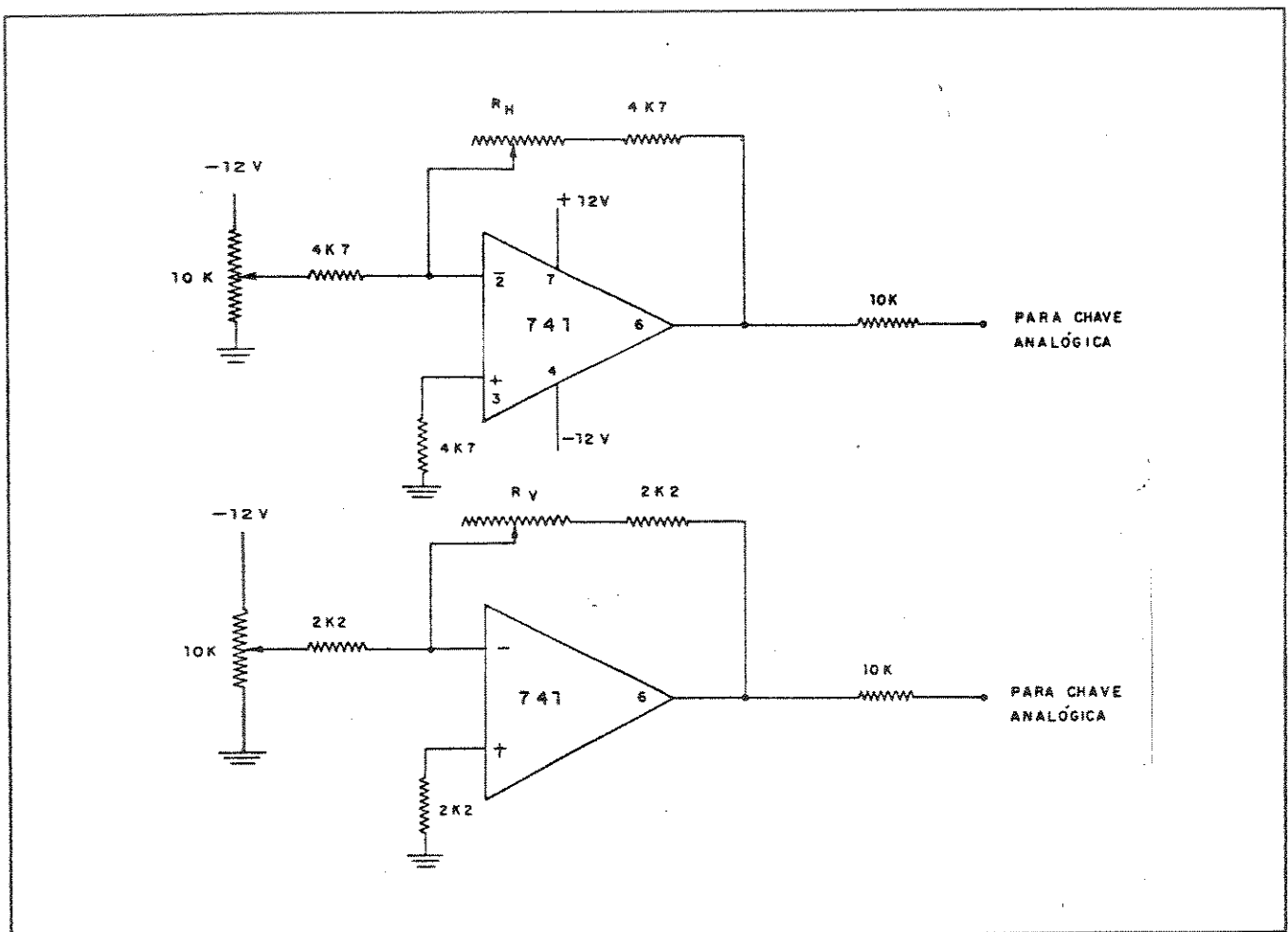


FIGURA A-3 - Circuito transdutor de posição horizontal e vertical do cursor da mesa digitalizadora.

A.2.2 - Porta do Usuário

A comunicação digital de dados entre o microcomputador e a mesa digitalizadora, é feita através de uma porta do usuário, construída neste laboratório, com base no circuito integrado 82C55A-2. Este circuito que possui 3 portas paralelas programáveis de 8 bits cada, é bastante versátil e, tem sido utilizado na implementação de comunicação assíncrona entre instrumentos e microcomputadores da linha IBM ou compatíveis[112,113]. Em resumo, o 8255 é configurado para trabalhar em regime de "handshaking" sendo que, o micro envia um sinal de "strobe" (STR) a cada conjunto de oito bits e, aguarda um pulso de "acknowledge" (ACK) proveniente do instrumento após cada operação de envio de dados ou efetivação de leitura de dados.

A.2.3 - Circuito Decodificador de Endereços

A figura A-4 descreve como foi construída a interface entre o microcomputador e os componentes do digitalizador. O desenho é uma modificação de um circuito semelhante, previamente proposto[114]. A modificação introduzida foi a substituição de um CI-4515 por um 74LS373 de custo mais baixo. Neste caso, a interface tem seu número de endereços limitado a oito. No caso da mesa digitalizadora, em sua configuração atual, somente três endereços são necessários.

Quando em situação de espera, o decodificador aguarda que um endereço seja enviado do microprocessador, seguido de um pulso na linha STR. Ao receber este pulso, os bits componentes do endereço são armazenados no CI-74373, que mantém suas saídas no modo "tri-state". Simultaneamente, um pulso ACK é enviado de volta ao micro. O byte de endereço é composto de forma que aos elementos endereçados apresentam um nível

lógico baixo em suas linhas de seleção. Eventuais correções devem ser feitas em caso da lógica de seleção apresentar-se invertida como no caso de um "latch" 74373.

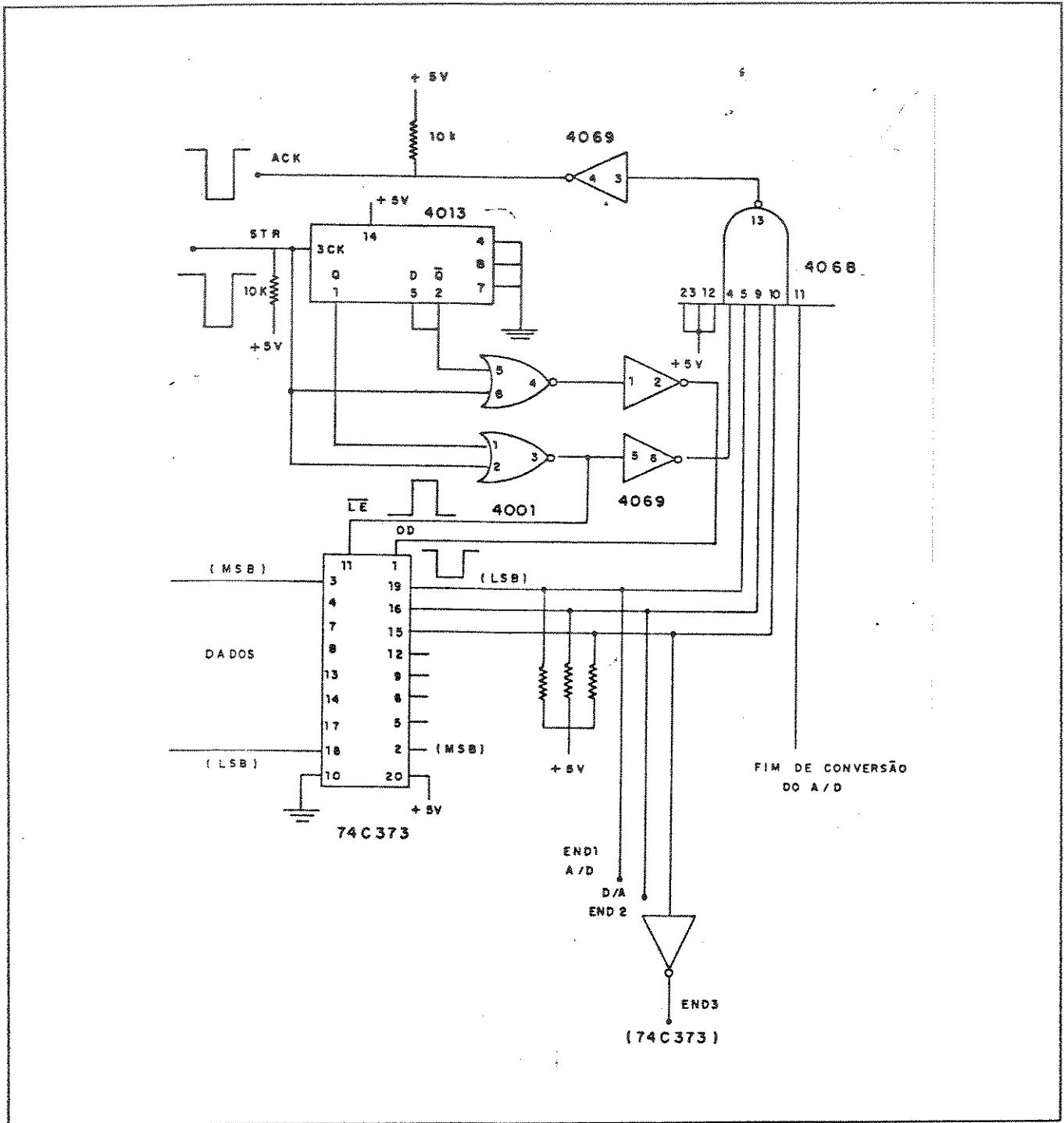


FIGURA A-4 - Diagrama do circuito decodificador de endereços.

Caso o microcomputador deseje enviar para o instrumento um certo dado, após a seleção do endereço, o byte de dados é colocado na porta de comunicação bidirecional (porta A do 8255) e a linha STR é novamente pulsada. Então, o endereço enviado torna-se transparente por um curto intervalo de tempo, nas linhas da saída do 74373. O elemento do circuito selecionado recebe a informação presente nas linhas de dados. Um segundo pulso ACK é gerado informando que a comunicação foi efetivada.

Caso o microcomputador seja instruído a realizar uma operação de leitura, (por exemplo, de um conversor analógico/digital) o processo é o mesmo até o endereçamento do elemento, do qual a leitura deverá ser realizada. A seguir, a linha STR é colocada em nível lógico baixo. Por se tratar de um endereço de leitura de dados, o pulso de ACK não é automaticamente gerado mas, deverá ser enviado pelo elemento endereçado. O protocolo de endereçamento pode, como será observado posteriormente, gerar um pulso de início de conversão, para um conversor analógico/digital. Então, o pulso ACK será gerado pelo final da operação de conversão, quando os dados estiverem prontos para serem lidos. Quando a leitura é efetivada pelo 8255, a linha STR é retornada a nível lógico alto, encerrando a comunicação e, deixando a interface a espera de um novo endereço.

Embora seja versátil, este tipo de comunicação é lenta em relação a do tipo síncrona e, algumas vezes, rotinas "assembler" tem de ser escritas de forma a aumentar a velocidade de transferência de bytes. Na presente aplicação, isto não é necessário uma vez que a variação do sinal resultante do movimento do cursor da mesa é relativamente lenta.

A.2.4 - Conversor Dinâmico Analógico/Digital com dois Canais

A última etapa de processamento dos sinais, cujas grandezas, são proporcionais as posições analógicas horizontais e verticais do cursor da mesa, é constituída da sua digitalização a qual é feita pelo circuito descrito na figura A-5. Este circuito tem como fundamento, um conversor dinâmico analógico/digital previamente descrito[114].

O circuito é constituído de um conversor A/D (ZN 448) e um conversor D/A (ZN 428) ambos de 8 bits. O sinal analógico proveniente dos transdutores de posição são apresentados ao A/D passando pelo amplificador operacional OP07 o qual amplifica a diferença existente entre este sinal e aquele gerado na saída do D/A. O ganho nominal sendo estabelecido em 47. Em conjunto com o programa de controle, o sistema dinâmico fornece uma resolução e precisão equivalentes a um A/D de 12 bits. Como o fundo de escala do A/D foi estabelecido em 2,55V, uma resolução de aproximadamente 0,6 mV pode ser esperada. O CI-4528 é um monoestável duplo, utilizado para gerar o pulso de início de conversão, quando o A/D, é endereçado e o pulso de ACK quando o A/D termina a conversão, retornando a linha do seu pino para nível alto. O programa de controle, estabelece que o menor valor que pode ser lido pelo A/D corresponde ao decimal 0 e, o maior ao decimal 245. Os valores de saída do D/A são automaticamente corrigidos pelo programa de forma a manter os valores dentro daqueles limites.

O sistema dinâmico foi acrescido da chave analógica (ADG 20) e do latch" 74373 os quais são, em conjunto, responsáveis pela seleção do canal analógico de leitura. Quando em operação, o CI-74373 é endereçado e, um dado correspondente ao canal que se deseja ler, é enviado a seguir. A chave é fechada quando um sinal de nível baixo está presente no pino de controle correspondente. Os dados 254 e 243 enviados para o endereço 251, selecionam as leituras da posição horizontal e vertical do cursor. O presente sistema de conversão pode ser expandido, sem acréscimo de nenhum componente, para monitorar até 4 canais analógicos.

O circuito de conversão dinâmico analógico/digital foi calibrado com o auxílio de um voltímetro com precisão de décimos de milivolts ligado a saída analógica do canal horizontal da mesa. As médias de 50 (cinquenta) determinações do valor lido do D/A e, do enviado para o A/D foram tomados e correlacionados com o valor da tensão obtido pelo voltímetro, para diversas posições do cursor. O resultado mostrou que o valor da tensão analógica gerada pelos transdutores pode ser obtida pela equação:

$$S = (DA\% / f) + (DI\% * f2)$$

onde; S é o sinal expresso em milivolts, DA% é o byte enviado para o conversor D/A, DI% é o byte lido do conversor A/D e f1 e f2 são constantes determinadas pela calibração, sendo iguais a 0,1049 e 0,3134; respectivamente.

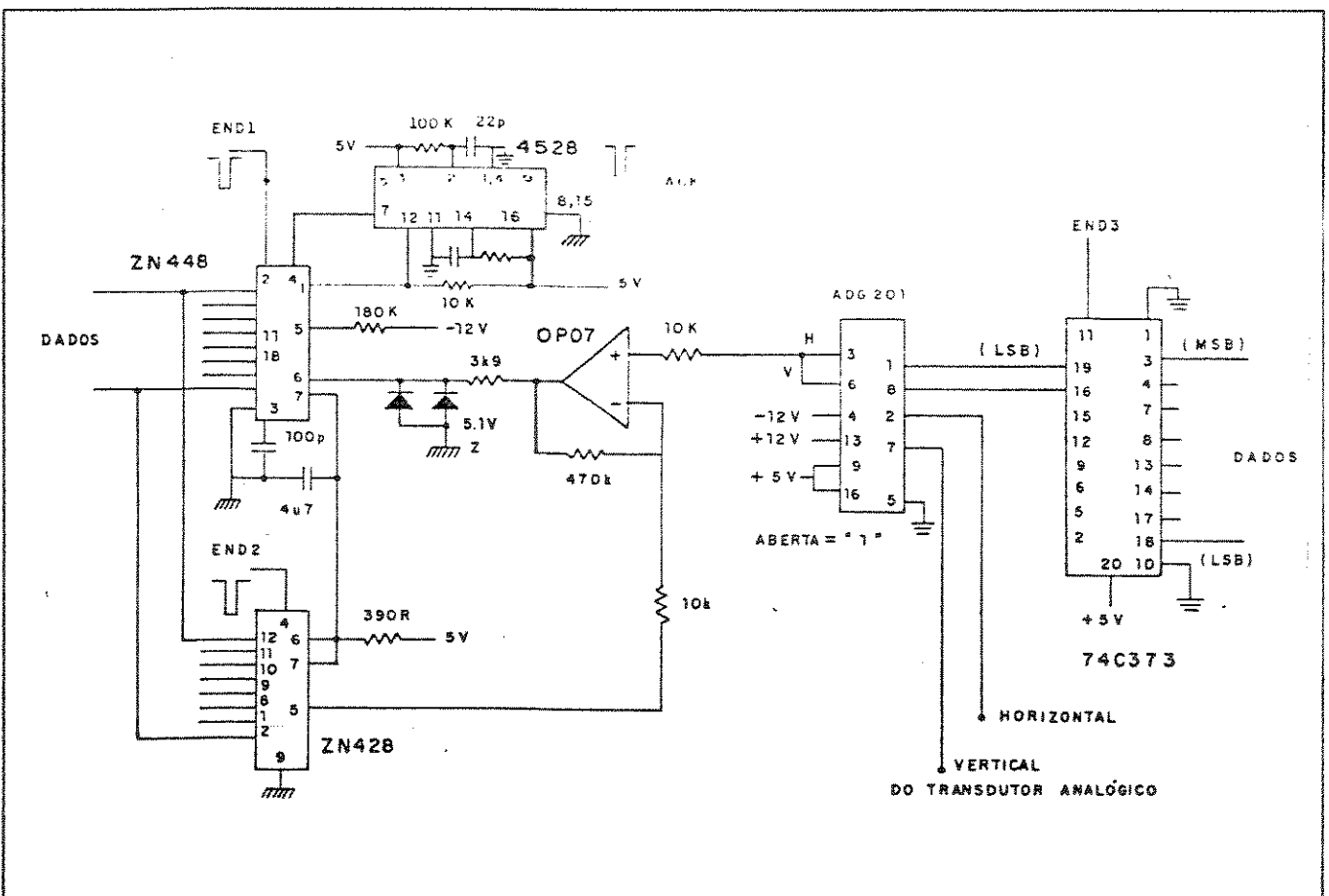


FIGURA A-5 - Circuito do conversor dinâmico analógico/digital com dois canais.

A.3 - RESULTADOS E DISCUSSÕES

Os primeiros testes com a mesa digitalizadora foram realizados, de forma a caracterizar a precisão, com a qual, as coordenadas poderiam ser obtidas. Em primeiro lugar, avaliou-se a linearidade dos transdutores analógicos de posição (resistores) movendo-se o cursor, ao longo de uma régua colocada sobre a mesa. Uma regressão linear de 20 e 30 pontos (potencial lido, posição em centímetro) coletados em diversas posições verticais e horizontais mostraram um coeficiente de correlação igual ou melhor que 0,9997. Desta forma, demonstrou-se que os transdutores são lineares em toda a extensão útil vertical e horizontal da mesa.

A seguir, vários pontos foram definidos em diversas posições da mesa, bem como, a leitura dos valores de potencial correspondentes as coordenadas vertical e horizontal. O cursor foi então deslocado aleatoriamente 10 vezes em ambas coordenadas, retornando a seguir, sobre o ponto de origem. Então, os valores de potencial eram novamente coletados. Os resultados deste experimento permitiram determinar a precisão, com a qual um ponto pode ser localizado, o que também, mostra qual é a resolução da mesa em unidades de deslocamento linear. Os desvios padrões absolutos médios foram iguais a 0,37 e 0,28 mm para as coordenadas horizontal e vertical, respectivamente.

Uma vez conhecidas as características da mesa digitalizadora, alguns espectros foram digitalizados para se observar o desempenho do programa de coleta de dados e para formar um pequeno arquivo utilizado no desenvolvimento e testes do programa especialista (EXPERTIV).

Com o auxílio do cursor gráfico procurou-se observar, as eventuais diferenças existentes entre o posicionamento de diversos picos no espectro original e no digitalizado. De uma forma geral, a posição é concordante com aquela do espectro original, dentro de um erro nunca superior a 5 cm^{-1} . Alguns espectros de tamanho muito reduzido, existentes

na literatura, devem ser ampliados em fotocópias e a seguir, digitalizados, a fim de se obter uma melhor concordância com os valores originais.

A principal vantagem da mesa digitalizadora desenvolvida, com relação aos equipamentos comerciais de função similar, existentes no mercado, está, sem dúvida, no seu baixo custo. No entanto, é importante ressaltar que o conhecimento detalhado da parte eletrônica e do protocolo de comunicação digital (dados nem sempre disponíveis, em se tratando de instrumentos comerciais) é de grande valia no desenvolvimento de aplicações específicas como no caso da digitalização de espectros.

Em uma avaliação final, pode-se afirmar que o digitalizador desenvolvido preserva essencialmente as informações contidas num espectro no I.V. original, podendo ser utilizado como interface para um programa especialista. Além disso, o seu uso pode ser estendido a elaboração de pequenas bibliotecas de espectros e, também, como ferramenta para aquisição de conjuntos de treinamento, para programas quimiométricos.

APÊNDICE B

- Base de Conhecimento "dba_links" do programa EXPERTIV -
- Arquivo DOS "dataiv.dba" -

faixa_absorcao("stretching_O-H",[3640,2500]).
faixa_absorcao("stretching_N-H",[3500,3000]).
faixa_absorcao("aromatic_stretching_C-H",[3100,2940]).
faixa_absorcao("unsaturated_stretching_C-H",[3090,2960]).
faixa_absorcao("aliphatic_stretching_C-H",[3080,2655]).
faixa_absorcao("stretching_CN",[2260,2050]).
faixa_absorcao("stretching_CC",[2260,2090]).
faixa_absorcao("stretching_C=O",[1900,1600]).
faixa_absorcao("unsaturated_stretching_C=C",[1790,1566]).
faixa_absorcao("aromatic_stretching_C=C",[1650,1400]).
faixa_absorcao("stretching_C=N",[1700,1500]).
faixa_absorcao("stretching_asymmetric_C-NO2",[1650,1500]).
faixa_absorcao("stretching_symmetric_C-NO2",[1380,1250]).
faixa_absorcao("stretching_C-O",[1440,1010]).
faixa_absorcao("stretching_C-N",[1550,1030]).
faixa_absorcao("stretching_C-C",[1300,1050]).
faixa_absorcao("stretching_C-F",[1400,1000]).
faixa_absorcao("stretching_C-Cl",[850,550]).
faixa_absorcao("stretching_C-Br",[690,500]).
faixa_absorcao("stretching_C-I",[600,500]).
faixa_absorcao("stretching_C=S",[1200,1050]).
faixa_absorcao("stretching_S-H",[2600,2550]).
faixa_absorcao("stretching_C-S",[800,600]).
faixa_absorcao("bending_C-H",[1460,960]).
faixa_absorcao("bending_C-H_in_plane",[1300,960]).
faixa_absorcao("bending_C-H_out_of_plane",[900,675]).
faixa_absorcao("overtone_of_bending_C-H",[2900,2655]).
faixa_absorcao("bending_O-H",[1420,1200]).
faixa_absorcao("bending_O-H_in_plane",[1500,1250]).
faixa_absorcao("bending_O-H_out_of_plane",[920,650]).
faixa_absorcao("bending_N-H",[1650,1490]).
faixa_absorcao("bending_N-H_in_plane",[1640,1560]).
faixa_absorcao("bending_N-H_out_of_plane",[900,650]).
faixa_absorcao("bending_C-N-O",[620,600]).
vibracoes_caracteristicas("acid_anhydride (-CO-O-CO)",["stretching_C=O","stretching_C=O","stretching_C-O"]).
vibracoes_caracteristicas("carboxylic_acid (-COOH)",["stretching_C=O","stretching_O-H","stretching_C-O"]).
vibracoes_caracteristicas("ester (-CO-O-)",["stretching_C=O","stretching_C-O"]).
vibracoes_caracteristicas("ketone (-CO-)",["stretching_C=O","stretching_C-C"]).
vibracoes_caracteristicas("aldehyde (-CHO)",["stretching_C=O","aliphatic_stretching_C-H","overtone_of_bending_C-H"]).
vibracoes_caracteristicas("primary_amide (-CO-NH2)",["stretching_C=O","stretching_N-H"]).
vibracoes_caracteristicas("secondary_amide (-CO-NHR-)",["stretching_C=O","stretching_N-H","stretching_C-N"]).
vibracoes_caracteristicas("acid_halide (-CO-X)",["stretching_C=O","stretching_C-X"]).
vibracoes_caracteristicas("alcohol (R-OH)",["stretching_O-H","stretching_C-O"]).

continua na próxima página

continuação

vibrações_características("phenol_(ph-OH)",["stretching_O-H","stretching_C-O"]);
vibrações_características("primary_amine_(R-NH2)",["stretching_N-H","stretching_N-H","stretching_C-N",
 "bending_N-H_in_plane","bending_N-H_out_of_plane"]);
vibrações_características("secondary_amine_(R-NH-R)",["stretching_N-H","stretching_C-N","bending_N-H"]);
vibrações_características("alkene_(CnH2n)",["unsaturated_stretching_C=C","unsaturated_stretching_C-H"]);
vibrações_características("aromatic",["aromatic_stretching_C-H","aromatic_stretching_C=C","bending_C-H_out_of_plane"]);
vibrações_características("acetylene_(CnH2n-2)",["stretching_CC"]);
vibrações_características("nitrile_(-CN)",["stretching_CN"]);
vibrações_características("alkane_(CnH2n+2)",["aliphatic_stretching_C-H","bending_C-H"]);
vibrações_características("ether_(R-O-R)",["stretching_C-O"]);
posicao_tipo_vibração("acid_anhydride_(-CO-O-CO)","stretching_C=O",[1830,1810],intensities(["strong",
 "very_strong"]),shape(["sharp","medium"]));
posicao_tipo_vibração("acid_anhydride_(-CO-O-CO)","stretching_C=O",[1770,1750],intensities(["medium",
 "strong","very_strong"]),shape(["sharp","medium"]));
posicao_tipo_vibração("acid_anhydride_(-CO-O-CO)","stretching_C-O",[1300,1045],intensities(["medium","strong",
 "very_strong"]),shape());
posicao_tipo_vibração("carboxylic_acid_(-COOH)","stretching_C=O",[1770,1700],intensities(["strong",
 "very_strong"]),shape());
posicao_tipo_vibração("carboxylic_acid_(-COOH)","stretching_O-H",[3550,2500],intensities(["medium","strong",
 "very_strong"]),shape(["sharp","broad"]));
posicao_tipo_vibração("carboxylic_acid_(-COOH)","stretching_C-O",[1440,1200],intensities(["medium","strong",
 "very_strong"]),shape());
posicao_tipo_vibração("ester_(-CO-O-)",["stretching_C=O",[1745,1725],intensities(["very_strong"]),
 shape(["sharp","medium"])]);
posicao_tipo_vibração("ester_(-CO-O-)",["stretching_C-O",[1300,1050],intensities(["strong","very_strong"]),
 shape()]);
posicao_tipo_vibração("ketone_(-CO-)",["stretching_C=O",[1725,1705],intensities(["strong","very_strong"]),
 shape(["sharp","medium"])]);
posicao_tipo_vibração("ketone_(-CO-)",["stretching_C-C",[1300,1100],intensities(["medium","strong"]),shape()]);
posicao_tipo_vibração("aldehyde_(-CHO)","stretching_C=O",[1740,1665],intensities(["medium","strong",
 "very_strong"]),shape(["sharp","medium"]));
posicao_tipo_vibração("aldehyde_(-CHO)","aliphatic_stretching_C-H",[2900,2695],intensities(["weak","medium",
 "strong","very_strong"]),shape());
posicao_tipo_vibração("aldehyde_(-CHO)","overtone_of_bending_C-H",[2900,2695],intensities(["weak","medium",
 "strong","very_strong"]),shape());
posicao_tipo_vibração("primary_amide_(-CO-NH2)","stretching_C=O",[1700,1640],intensities(["very_strong"]),
 shape(["medium"]));
posicao_tipo_vibração("primary_amide_(-CO-NH2)","stretching_N-H",[3500,3200],intensities(["strong",
 "very_strong"]),shape());
posicao_tipo_vibração("primary_amide_(-CO-NH2)","bending_N-H",[1650,1600],intensities(["strong",
 "very_strong"]),shape());
posicao_tipo_vibração("secondary_amide_(-CO-NHR-)",["stretching_C=O",[1690,1645],intensities(["very_strong"]),
 shape(["medium"])]);
posicao_tipo_vibração("secondary_amide_(-CO-NHR-)",["stretching_N-H",[3440,3300],intensities(["strong"]),
 shape()]);
posicao_tipo_vibração("secondary_amide_(-CO-NHR-)",["stretching_C-N",[1300,1260],intensities(),shape()]);
posicao_tipo_vibração("secondary_amide_(-CO-NHR-)",["bending_N-H",[1550,1530],intensities(),shape()]);
posicao_tipo_vibração("acid_halide_(-CO-X)","stretching_C=O",[1870,1750],intensities(["strong",
 "very_strong"]),shape(["sharp","medium"]));
posicao_tipo_vibração("acid_halide_(-CO-X)","stretching_C-F",[1400,1000],intensities(["very_strong"]),
 shape(["sharp"]));

continua na próxima página

continuação

posicao_tipo_vibracao("acid_halide_(-CO-X)", "stretching_C-Cl", [850,550], intensities(["strong"]), shape(()).
 posicao_tipo_vibracao("acid_halide_(-CO-X)", "stretching_C-Br", [600,500], intensities(["strong"]), shape(()).
 posicao_tipo_vibracao("acid_halide_(-CO-X)", "stretching_C-I", [600,500], intensities(["strong"]), shape(()).
 posicao_tipo_vibracao("alcohol_(R-OH)", "stretching_O-H", [3640,2500], intensities(["strong", "very_strong"]),
 shape(["sharp", "broad"])).
 posicao_tipo_vibracao("alcohol_(R-OH)", "stretching_C-O", [1280,1000], intensities(["strong", "very_strong"]),
 shape(()).
 posicao_tipo_vibracao("phenol_(ph-OH)", "stretching_O-H", [3640,3200], intensities(["medium", "strong",
 "very_strong"]), shape(["sharp", "broad"])).
 posicao_tipo_vibracao("phenol_(ph-OH)", "stretching_C-O", [1260,1000], intensities(["medium", "strong",
 "very_strong"]), shape(()).
 posicao_tipo_vibracao("primary_amine_(R-NH2)", "stretching_N-H", [3510,3490], intensities(["weak"]), shape(()).
 posicao_tipo_vibracao("primary_amine_(R-NH2)", "stretching_N-H", [3410,3390], intensities(["weak"]), shape(()).
 posicao_tipo_vibracao("primary_amine_(R-NH2)", "stretching_C-N", [1230,1030], intensities(["medium"]), shape(()).
 posicao_tipo_vibracao("primary_amine_(R-NH2)", "bending_N-H_in_plane", [1640,1560], intensities(["medium",
 "strong"]), shape(()).
 posicao_tipo_vibracao("primary_amine_(R-NH2)", "bending_N-H_out_of_plane", [900,650], intensities(["medium"]),
 shape(["broad"])).
 posicao_tipo_vibracao("secondary_amine_(R-NH-R)", "stretching_N-H", [3490,3300], intensities(["weak", "medium"]),
 shape(["sharp"])).
 posicao_tipo_vibracao("secondary_amine_(R-NH-R)", "stretching_C-N", [1360,1180], intensities(["medium",
 "strong"]), shape(()).
 posicao_tipo_vibracao("secondary_amine_(R-NH-R)", "bending_N-H", [1580,1490], intensities(["weak"]), shape(()).
 posicao_tipo_vibracao("alkene_(CnH2n)", "unsaturated_stretching_C=C", [1790,1566], intensities(["weak", "medium"]),
 shape(["sharp", "medium"])).
 posicao_tipo_vibracao("alkene_(CnH2n)", "unsaturated_stretching_C-H", [3090,2960], intensities(["medium", "strong"]),
 shape(["sharp", "medium"])).
 posicao_tipo_vibracao("acetylene_(CnH2n-2)", "stretching_CC", [2260,2090], intensities(["very_weak", "weak",
 "strong"]), shape(["sharp", "medium"])).
 posicao_tipo_vibracao("nitrile_(-CN)", "stretching_CN", [2260,2050], intensities(["weak", "medium", "strong",
 "very_strong"]), shape(()).
 posicao_tipo_vibracao("aromatic", "aromatic_stretching_C=C", [1650,1400], intensities(["medium",
 "strong"]), shape(["sharp", "medium"])).
 posicao_tipo_vibracao("aromatic", "aromatic_stretching_C-H", [3100,2940], intensities(["weak", "medium"]), shape(()).
 posicao_tipo_vibracao("aromatic", "bending_C-H_out_of_plane", [900,675], intensities(["medium", "strong",
 "very_strong"]), shape(()).
 posicao_tipo_vibracao("alkane_(CnH2n+2)", "aliphatic_stretching_C-H", [2970,2850], intensities(["medium", "strong",
 "very_strong"]), shape(()).
 posicao_tipo_vibracao("alkane_(CnH2n+2)", "bending_C-H", [1460,1355], intensities(["weak", "medium", "strong"]),
 shape(()).
 posicao_tipo_vibracao("ether_(R-O-R)", "stretching_C-O", [1275,1020], intensities(["strong", "very_strong"]),
 shape(()).
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1875,1855], "(5-membered)").
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1860,1840], "(maleic)").
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1860,1840], "(phthalic)").
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1810,1790], "(6-membered)").
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1800,1780], "(maleic)").
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1795,1775], "(acrylic_or_benzoic)").
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1795,1775], "(5-membered)").
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1790,1770], "(Alpha-Beta-unsatd_6-
 membered)").
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1780,1760], "(phthalic)").

continua na próxima página

continuação

posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1760,1740], "(6-membered)").
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1745,1725], "(Alpha-Beta-unsatd_6-membered)").
 posicao_tipo_especifico("acid_anhydride_(-CO-O-CO)", "stretching_C=O", [1735,1715], "(acrylic_or_benzoic)").
 posicao_tipo_especifico("carboxylic_acid_(-COOH)", "stretching_C=O", [1820,1800], "(Alpha_F)_monomer").
 posicao_tipo_especifico("carboxylic_acid_(-COOH)", "stretching_C=O", [1780,1760], "(Alpha_Br)_or_(Alpha_Cl)_monomer").
 posicao_tipo_especifico("carboxylic_acid_(-COOH)", "stretching_C=O", [1770,1750], "(Alpha_F)_dimer").
 posicao_tipo_especifico("carboxylic_acid_(-COOH)", "stretching_C=O", [1730,1700], "(Alpha_Br)_or_(Alpha_Cl)_dimer").
 posicao_tipo_especifico("carboxylic_acid_(-COOH)", "stretching_C=O", [1730,1710], "(C=C-COOH)_monomer").
 posicao_tipo_especifico("carboxylic_acid_(-COOH)", "stretching_C=O", [1700,1680], "(Ar-COOH)_dimer").
 posicao_tipo_especifico("carboxylic_acid_(-COOH)", "stretching_C=O", [1675,1655], "(salicylic_acid)").
 posicao_tipo_especifico("carboxylic_acid_(-COOH)", "stretching_O-H", [3000,2500], "(COOH)_dimer").
 posicao_tipo_especifico("carboxylic_acid_(-COOH)", "stretching_O-H", [3550,3000], "(COOH)_monomer").
 posicao_tipo_especifico("carboxylic_acid_(-COOH)", "stretching_C-O", [1440,1200], "(dimer)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1850,1830], "(lactone_4-membered)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1800,1780], "(Beta_Gama_lactone_5-membered)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1795,1775], "(Alpha_Beta_lactone_5-membered)_
(doublet_in_presence_Alpha_H)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1780,1760], "(O=C-O-CH2-C6H6)_polycyclic").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1775,1745], "(Alpha_halo)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1780,1760], "(lactone_5-membered)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1770,1750], "(CO-O-C=C)_or_(Gama_Delta_lactone_6-membered)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1765,1745], "(Alpha_Beta_lactone_5-membered)_
(doublet_in_presence_Alpha_H)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1760,1740], "(Alpha_Beta_lactone_5-membered)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1755,1735], "(Alpha_keto_ester)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1745,1725], "(Ar-CO-O-Ar)_or_(Beta_keto_ester_(keto))_or_(lactone_6-membered)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1740,1720], "(Alpha_pyrones)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1730,1710], "(C=C-CO-O-)_or_(Ar-CO-O-)_
or_(Alpha_Beta_lactone_6-membered_or_(Beta_keto_ester_(keto)))").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C=O", [1660,1640], "(Beta_keto_ester_(enol))").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C-O", [1300,1250], "(Ar-CO-O-)_or_(C=C-CO-O-)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C-O", [1250,1230], "(CH3-CO-OR)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C-O", [1220,1210], "(CO-O-C=C-)_in_(enol_and_phenol_acetates)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C-O", [1200,1050], "(C=C-CO-O-)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C-O", [1190], "(R-CO-OR)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C-O", [1190,1170], "(H-CO-OR)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C-O", [1180,1160], "(Ar-CO-O-)").
 posicao_tipo_especifico("ester_(-CO-O-)", "stretching_C-O", [1175,1155], "(R-CO-OCH3)").
 posicao_tipo_especifico("ketone_(-CO-)", "stretching_C=O", [1860,1840], "(3-membered)").
 posicao_tipo_especifico("ketone_(-CO-)", "stretching_C=O", [1790,1770], "(4-membered)").
 posicao_tipo_especifico("ketone_(-CO-)", "stretching_C=O", [1785,1765], "(CO-CO-)_ (5-membered_ring_in_s-cis_form)").
 posicao_tipo_especifico("ketone_(-CO-)", "stretching_C=O", [1770,1750], "(CO-CO-)_ (6-membered_ring_in_s-cis_form)_or_(CO-CO-)_ (5-membered_ring_in_s-cis_form)").
 posicao_tipo_especifico("ketone_(-CO-)", "stretching_C=O", [1760,1715], "(CX2-CO)_and_(CX-CO-CX)").
 posicao_tipo_especifico("ketone_(-CO-)", "stretching_C=O", [1755,1735], "(5-membered)").

continua na próxima página

continuação

posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1740,1715], "(Alpha-halo)_(Cl-C-C=O)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1740,1720], "(6-membered_ring_in_s-cis_form)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1735,1715], "(Alpha-halo)_(Br-C-C=O)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1730,1710], "(CO-CO-)_stretching_antysym_of_the_two_CO_or_(CO-CH-CO-)_occasionally_doublet").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1725,1715], "(Alpha-halo)_(I-C-C=O)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1725,1705], "(6-membered)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1715,1695], "(7-membered_and_larger)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1700,1680], "(Ar-CO-)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1695,1675], "(cyclopropyl-CO-R)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1685,1665], "(Alpha_Beta_unsatd)_or_(1,4_or_1,2_quinone)_or_(conj_with_cyclopropane)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1675,1655], "(Alpha_Beta_Gama_Delta_unsatd)_or_(Alpha_Beta_Alpha'_Beta'_unsatd)_or_(Ar-CO-Ar)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1660,1640], "(HO_or_NH2)-C=C-CO-)_free_or_(tropone)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1655,1635], "(extended_quinone)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1650,1630], "(RO-C=C-CO)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1640,1620], "(ortho-CO-C6H4-(OH_or_-NH2)").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1625,1605], "(HO-or_NH2)-C=C-CO)_when_intramol_H-bonded").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C=O", [1610,1590], "(tropolones)_intramolecularly_H-bonded").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C-C", [1300,1230], "stretching (-C(CO)-C-)_aromatic").
 posicao_tipo_especifico("ketone (-CO-)", "stretching C-C", [1230,1100], "stretching (-C(CO)-C-)_aliphatic").
 posicao_tipo_especifico("aldehyde (-CHO)", "stretching C=O", [1710,1690], "(Ar-CHO)").
 posicao_tipo_especifico("aldehyde (-CHO)", "stretching C=O", [1695,1675], "(Alpha_Beta_unsatd)").
 posicao_tipo_especifico("aldehyde (-CHO)", "stretching C=O", [1685,1665], "(Alpha_Beta_Gama_Delta_unsatd)").
 posicao_tipo_especifico("primary_amide (-CO-NH2)", "stretching C=O", [1700,1680], "(CONH2)_free_(Amide_I_band)").
 posicao_tipo_especifico("primary_amide (-CO-NH2)", "stretching C=O", [1660,1640], "(CONH2)_associated_(Amide_I_band)").
 posicao_tipo_especifico("primary_amide (-CO-NH2)", "stretching N-H", [3500,3400], "(CONH2)_free_two_bands").
 posicao_tipo_especifico("primary_amide (-CO-NH2)", "stretching N-H", [3350,3200], "(CONH2)_associated_several_bands").
 posicao_tipo_especifico("primary_amide (-CO-NH2)", "bending N-H", [1650,1640], "(CONH2)_associated_two_bands(I_band_is_stronger)").
 posicao_tipo_especifico("primary_amide (-CO-NH2)", "bending N-H", [1650,1630], "(CONH2)_associated_two_bands").
 posicao_tipo_especifico("primary_amide (-CO-NH2)", "bending N-H", [1610,1590], "(CONH2)_free_two_bands").
 posicao_tipo_especifico("secondary_amide (-CO-NHR-)", "stretching C=O", [1690,1670], "(CONHR)_free_(Amide_I_band)").
 posicao_tipo_especifico("secondary_amide (-CO-NHR-)", "stretching C=O", [1665,1645], "(CONHR)_associated_(Amide_I_band)").
 posicao_tipo_especifico("secondary_amide (-CO-NHR-)", "stretching N-H", [3450,3430], "(CONHR)_free").
 posicao_tipo_especifico("secondary_amide (-CO-NHR-)", "stretching N-H", [3310,3290], "(CONHR)_associated_(trans_assoc_form)").
 posicao_tipo_especifico("secondary_amide (-CO-NHR-)", "stretching C-N", [1310,1290], "(CONHR)_associated").
 posicao_tipo_especifico("secondary_amide (-CO-NHR-)", "stretching C-N", [1270,1250], "(CONHR)_free").

continua na próxima página

continuação

posicao_tipo_especifico("secondary_amide_(-CO-NHR-)", "bending_N-H", [1540,1520], "((-CONHR)_free)_or_(-CONHR)_associated").

posicao_tipo_especifico("acid_halide_(-CO-X)", "stretching_C=O", [1810,1790], "(-CO-Cl)_satd.").

posicao_tipo_especifico("acid_halide_(-CO-X)", "stretching_C=O", [1790,1740], "(-CO-X)_unsatd.").

posicao_tipo_especifico("acid_halide_(-CO-X)", "stretching_C=O", [1750,1735], "(-CO-X)_aryl_chloride").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_O-H", [3640,3610], "(free_monomeric)").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_O-H", [3650,3630], "(primary_alcohol)").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_O-H", [3640,3620], "(secondary_alcohol)").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_O-H", [3630,3610], "(tertiary_alcohol)").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_O-H", [3600,3500], "(dimeric)_intermol._H-bond_or_(polyvalent)_intramol._H-bond_or_(pi_H-bond)").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_O-H", [3560,3530], "(-O-O-H)_hydro-peroxide").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_O-H", [3400,3200], "(polymeric)").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_O-H", [3200,2500], "(chelation)_intramol._H-bond_with_(C=O, NO2, etc.)").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_C-O", [1150,1100], "(tertiary_alcohol)_ring_formation_between_(Alpha_and_Alpha)').

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_C-O", [1150,1060], "(tertiary_alcohol)_(Alpha_unsaturation_and_Alpha'_branching)_or_(Alpha_and_Alpha_unsaturation)").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_C-O", [1150,1010], "(tertiary_alcohol)_(Alpha,Alpha'_and_Alpha'_unsaturations)").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_C-O", [1100,1070], "(secondary_alcohol)_(Alpha_unsaturation)").

posicao_tipo_especifico("alcohol_(R-OH)", "stretching_C-O", [1050,1035], "(primary_alcohol)_(Alpha_branching)").

posicao_tipo_especifico("primary_amine", "stretching_N-H", [3510,3490], "(R-NH2)_free").

posicao_tipo_especifico("primary_amine", "stretching_N-H", [3410,3390], "(Ar-NH2)_free").

posicao_tipo_especifico("primary_amine", "stretching_C-N", [1230,1030], "(R-N-H2(R2))").

posicao_tipo_especifico("secondary_amine", "stretching_N-H", [3500,3480], "(pyrroles, indoles, etc.)_free").

posicao_tipo_especifico("secondary_amine", "stretching_N-H", [3460,3440], "(Ar-NH-R)_free").

posicao_tipo_especifico("secondary_amine", "stretching_N-H", [3400,3300], "(C=NH)_free_secondary").

posicao_tipo_especifico("secondary_amine", "stretching_N-H", [3350,3310], "(R-NH-R)_free").

posicao_tipo_especifico("secondary_amine", "stretching_C-N", [1360,1250], "(Ar-NH-R(R))_C_ arom.-NH_stretching").

posicao_tipo_especifico("secondary_amine", "stretching_C-N", [1280,1180], "(Ar-NHR(R))_C_aliph.-N_stretching").

posicao_tipo_especifico("secondary_amine", "stretching_C-N", [1690,1640], "(C=NH)_free_secondary").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C=C", [1685,1665], "(HRC=CR'H)_trans").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C=C", [1680,1660], "(RR'C=CR'H)_or_(RR'C=CR'R'')_tri_and_tetra-subst.").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C=C", [1670,1650], "(HRC=CR'H)_cis").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C=C", [1665,1645], "(R2C=CH2)_term._methylene").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C=C", [1650,1600], "(diene)_conjugated_double_bonds_or_(triene)_conjugated_double_bonds").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C=C", [1650,1580], "(polyene)_conjugated_double_bonds").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C=C", [1655,1635], "(R/CH=CH2)_term._vinyl").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C=C", [1576,1556], "(butene_cyclic)").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C-H", [3090,3070], "(=CH2)_stretching_antisymmetric_CH2").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C-H", [3030,3010], "(=CH-) _stretching_CH2").

posicao_tipo_especifico("alkene_(CnH2n)", "unsaturated_stretching_C-H", [2985,2965], "(=CH2)_stretching_symmetric_CH2").

continua na próxima página

continuação

posicao_tipo_especifico("acetylene (CnH2n-2)", "stretching_CC", [2140, 2100], "(stretching_tripleCH_terminal)").
 posicao_tipo_especifico("acetylene (CnH2n-2)", "stretching_CC", [2260, 2190], "(stretching_tripleCH_central)").
 posicao_tipo_especifico("nitrile (-CN)", "stretching_CN", [2260, 2210], "(-CN)").
 posicao_tipo_especifico("nitrile (-CN)", "stretching_CN", [2100, 2050], "(-CN)_alkali_metal_salts").
 posicao_tipo_especifico("aromatic", "aromatic_stretching_C-H", [3040, 3020], "(several_bands)").
 posicao_tipo_especifico("aromatic", "aromatic_stretching_C=C", [1600, 1580], "(phenyl_nucleus)_the_1500_cm-1_band_sometimes_stronger_than_C=C_band").
 posicao_tipo_especifico("aromatic", "aromatic_stretching_C=C", [1500, 1450], "(the_1600_cm-1_band_sometimes_stronger_than_C=C_band)").
 posicao_tipo_especifico("aromatic", "bending_C-H_in_plane", [1225, 1175], "(1:2- 1:4- 1:2:4- substitution)").
 posicao_tipo_especifico("aromatic", "bending_C-H_in_plane", [1175, 1125], "(1:2:4- substitution)").
 posicao_tipo_especifico("aromatic", "bending_C-H_in_plane", [1070, 1000], "(1:2- 1:4- 1:2:4- substitution (two_bands))").
 posicao_tipo_especifico("aromatic", "bending_C-H_in_plane", [1175, 1125], "(1- 1:3- 1:2:3- 1:3:5- substitution)").
 posicao_tipo_especifico("aromatic", "bending_C-H_in_plane", [1110, 1070], "(1- 1:3- 1:2:3- substitution)").
 posicao_tipo_especifico("aromatic", "bending_C-H_in_plane", [1070, 1000], "(1- 1:3- 1:2:3- 1:3:5- substitution)").
 posicao_tipo_especifico("aromatic", "bending_C-H_in_plane", [1000, 960], "(1:2- 1:2:3- 1:2:4- substitution)").
 posicao_tipo_especifico("alkane (CnH2n+2)", "aliphatic_stretching_C-H", [3080, 3060], "(>C-CH2-C<)_cyclopropane").
 posicao_tipo_especifico("alkane (CnH2n+2)", "aliphatic_stretching_C-H", [3050], "(stretching_antisymmetric_CH2_of_methylene)_or_(>C-NH-CH2_cyclo)_stretching_antisymmetric_CH2").
 posicao_tipo_especifico("alkane (CnH2n+2)", "aliphatic_stretching_C-H", [3010, 3000], "(>C-CRH-C<)_cyclopropane").
 posicao_tipo_especifico("alkane (CnH2n+2)", "aliphatic_stretching_C-H", [3000], "(>C-O-C<_epoxy)_strething_CH_of_methine").
 posicao_tipo_especifico("alkane (CnH2n+2)", "aliphatic_stretching_C-H", [2960], "(stretching_antisymmetric_CH3)").
 posicao_tipo_especifico("alkane (CnH2n+2)", "aliphatic_stretching_C-H", [2925], "(stretching_antisymmetric_CH2)").
 posicao_tipo_especifico("alkane (CnH2n+2)", "aliphatic_stretching_C-H", [2870], "(stretching_symmetric_CH3)").
 posicao_tipo_especifico("alkane (CnH2n+2)", "aliphatic_stretching_C-H", [2850], "(stretching_symmetric_CH2)").
 posicao_tipo_especifico("ether (R-O-R)", "stretching_C-O", [1275, 1200], "(=C-O-C)_antisymmetric_stretching_aromatic_and_vinyl").
 posicao_tipo_especifico("ether (R-O-R)", "stretching_C-O", [1260, 1240], "(>COC<)_epoxy").
 posicao_tipo_especifico("ether (R-O-R)", "stretching_C-O", [1250, 1040], "(methylenedioxy)").
 posicao_tipo_especifico("ether (R-O-R)", "stretching_C-O", [1200, 1030], "(C-O-C-O-C_and_cyclic)_Ketals_and_acetals").
 posicao_tipo_especifico("ether (R-O-R)", "stretching_C-O", [1150, 1070], "(C-O-C)_antisym.stretching").
 posicao_tipo_especifico("ether (R-O-R)", "stretching_C-O", [1075, 1020], "(=C-O-C)_symmetric_stretching_aromatic_and_vinyl").
 outro_grupo_ausente("acid_halide (-CO-X)", "halogen (C-X)").
 outro_grupo_ausente("alkene (CnH2n)", "double_bound (C=C)").
 outro_grupo_ausente("aromatic", "aromatic").
 outro_grupo_ausente("phenol (ph-OH)", "aromatic").
 outro_grupo_ausente("acetylene (CnH2n-2)", "acetylene (CnH2n-2)").
 outro_grupo_ausente("nitrile (-CN)", "nitrile (-CN)").
 vibracoes_caracteristicas_outro_grupo("amine (R-NH2)", ["stretching_N-H", "stretching_C-N"]).
 vibracoes_caracteristicas_outro_grupo("amine (R-NH-R)", ["stretching_N-H", "stretching_C-N"]).
 vibracoes_caracteristicas_outro_grupo("double_bound (C=C)", ["unsaturated_stretching_C=C", "unsaturated_stretching_C-H"]).
 vibracoes_caracteristicas_outro_grupo("aromatic", ["aromatic_stretching_C=C", "aromatic_stretching_C-H", "bending_C-H_out_of_plane"]).

continua na próxima página

continuação

vibracoes_caracteristicas_outro_grupo("nitro_(C-NO2)",["stretching_asymmetric_C-NO2","stretching_symmetric_C-NO2"]).

vibracoes_caracteristicas_outro_grupo("halogen_(C-X)",["stretching_C-X"]).

vibracoes_caracteristicas_outro_grupo("acetylene_(CnH2n-2)",["stretching_CC"]).

vibracoes_caracteristicas_outro_grupo("nitrile_(-CN)",["stretching_CN"]).

vibracoes_caracteristicas_outro_grupo("tio_grupo",["stretching_C=S"]).

vibracoes_caracteristicas_outro_grupo("tio_grupo",["stretching_S-H"]).

vibracoes_caracteristicas_outro_grupo("tio_grupo",["stretching_C-S"]).

tipo_vibracao_outro_grupo("amine_(R-NH2)", "stretching_N-H",[3500,3400],intensities(["weak"]),shape(1)).

tipo_vibracao_outro_grupo("amine_(R-NH2)", "stretching_C-N",[1250,1020],intensities(["medium"]),shape(1)).

tipo_vibracao_outro_grupo("amine_(R-NH2)", "bending_N-H_in_plane",[1650,1560],intensities(["medium","strong","very_strong"]),shape(1)).

tipo_vibracao_outro_grupo("amine_(R-NH-R)", "stretching_N-H",[3350,3310],intensities(["weak"]),shape(1)).

tipo_vibracao_outro_grupo("amine_(R-NH-R)", "stretching_C-N",[1360,1180],intensities(["weak","medium","strong"]),shape(1)).

tipo_vibracao_outro_grupo("double_bound_(C=C)", "unsaturated_stretching_C=C",[1675,1640],intensities(["weak","medium"]),shape(1)).

tipo_vibracao_outro_grupo("double_bound_(C=C)", "unsaturated_stretching_C-H",[3090,2960],intensities(["medium","strong"]),shape(["sharp","medium"])).

tipo_vibracao_outro_grupo("aromatic", "aromatic_stretching_C=C",[1650,1400],intensities(["medium","strong"]),shape(["sharp","medium"])).

tipo_vibracao_outro_grupo("aromatic", "aromatic_stretching_C-H",[3100,2940],intensities(["weak","medium"]),shape(1)).

tipo_vibracao_outro_grupo("aromatic", "bending_C-H_out_of_plane",[900,675],intensities(["medium","strong","very_strong"]),shape(1)).

tipo_vibracao_outro_grupo("halogen_(C-X)", "stretching_C-F",[1400,1000],intensities(["very_strong"]),shape(["sharp"])).

tipo_vibracao_outro_grupo("halogen_(C-X)", "stretching_C-Cl",[850,550],intensities(["strong","very_strong"]),shape(1)).

tipo_vibracao_outro_grupo("halogen_(C-X)", "stretching_C-Br",[690,500],intensities(["strong","very_strong"]),shape(1)).

tipo_vibracao_outro_grupo("halogen_(C-X)", "stretching_C-I",[600,500],intensities(["strong"]),shape(1)).

tipo_vibracao_outro_grupo("nitro_(C-NO2)", "stretching_asymmetric_C-NO2",[1650,1500],intensities(["strong","very_strong"]),shape(1)).

tipo_vibracao_outro_grupo("nitro_(C-NO2)", "stretching_symmetric_C-NO2",[1370,1250],intensities(1),shape(1)).

tipo_vibracao_outro_grupo("acetylene_(CnH2n-2)", "stretching_CC",[2260,2090],intensities(["very_weak","weak","strong"]),shape(["sharp","medium"])).

tipo_vibracao_outro_grupo("nitrile_(-CN)", "stretching_CN",[2260,2050],intensities(["weak","medium","strong","very_strong"]),shape(1)).

tipo_vibracao_outro_grupo("tio_grupo", "stretching_C=S",[1200,1050],intensities(["strong"]),shape(1)).

tipo_vibracao_outro_grupo("tio_grupo", "stretching_S-H",[2660,2550],intensities(["weak"]),shape(1)).

tipo_vibracao_outro_grupo("tio_grupo", "stretching_C-S",[800,600],intensities(["weak"]),shape(1)).

APÊNDICE C

Relação dos 350 compostos orgânicos interpretados pelo programa EXPERTIV, através de seus espectros de I.V..

ANIDRIDOS DE ÁCIDOS

1. ¹³C-ACETIC ANHYDRIDE (CH₃-COOCO-CH₃)¹³
2. ¹³C-BUTYRIC ANHYDRIDE (CH₃CH₂CH₂-COOCO-CH₂CH₂CH₃)¹³
3. ¹³C-HEXANOIC ANHYDRIDE (C₁₂H₂₂O₃)¹³
4. ¹³C-ACETIC ANHYDRIDE-d₆ (CD₃-COOCO-CD₃)¹³
5. ¹³C-PROPIONIC ANHYDRIDE (CH₃CH₂COOCOCH₂CH₃)¹³
6. ¹³C-PIVALIC ANHYDRIDE (C₁₀H₁₈O₃)¹³
7. ¹³C-cis-1-2-CYCLOBUTANEDICARBOXYLIC ANHYDRIDE¹³
8. ¹³C-4-METHYL-1-2-CYCLOHEXANEDICARBOXYLIC ANHYDRID¹³
9. ¹³C-CINNAMIC ANHYDRIDE (C₈H₁₀O₃)¹³
10. ¹³C-CROTONIC ANHYDRIDE (CH₃CH=CHCOOCOCH=CHCH₃)¹³
11. ¹³C-GLYCINE ANHYDRIDE (H₂NCH₂COOCOCH₂NH₂)¹³
12. ¹³C-BENZOIC ANHYDRIDE (C₁₄H₁₀O₃)¹³
13. ¹³C-TETRACHLOROPHTHALIC ANHYDRIDE (C₈Cl₄O₃)¹³
14. ¹³C-TETRABROMOPHTHALIC ANHYDRIDE (C₈Br₄O₃)¹³
15. ¹³C-SUCCINIC ANHYDRIDE (C₄H₄O₃)¹³
16. ¹³C-DIMETHYLMALEIC ANHYDRIDE (C₆H₆O₃)¹³
17. ¹³C-DICHLOROMALEIC ANHYDRIDE (C₄Cl₂O₃)¹³
18. ¹³C-(+)-CAMPHORIC ANHIDRIDE (C₁₀H₁₄O₃)¹³

ÁCIDOS CARBOXÍLICOS

1. ¹³C-3-BROMOPROPIONIC ACID (Br-CH₂-CH₂-COOH)¹³
2. ¹³C-PIVALIC ACID (PROPIONIC ACID_2-2-DIMETHYL)¹³
3. ¹³C-ISOBUTYRIC ACID ((CH₃)₂-CH-COOH)¹³
4. ¹³C-DL-2-METHYLBUTYRIC ACID (CH₃CH₂CHCH₃COOH)¹³
5. ¹³C-ISOVALERIC ACID (BUTYRIC ACID - 3-METHYL)¹³
6. ¹³C-2-BROMO-3-METHYLBUTYRIC ACID (C₅H₉BrO₂)¹³
7. ¹³C-2-BROMOHEXANOIC ACID (CH₃CH₂CH₂CH₂CHBrCOOH)¹³
8. ¹³C-2-BROMOHEPTANOIC ACID (C₇H₁₃BrO₂)¹³
9. ¹³C-2-ETHYL-2-HEXENOIC ACID (C₈H₁₄O₂)¹³
10. ¹³C-2-HEPTENOIC ACID (CH₃CH₂CH₂CH₂-CH=CHCOOH)¹³
11. ¹³C-3-HEPTENOIC ACID (CH₃CH₂CH₂-CH=CH-CH₂COOH)¹³
12. ¹³C-4-PHENYLBUTYRIC ACID (Ph-CH₂CH₂CH₂COOH)¹³
13. ¹³C-TRIFLUOROACETIC ACID (CF₃COOH)¹³
14. ¹³C-TRICHLOROACETIC ACID (Cl₃CCOOH)¹³
15. ¹³C-PHENYLACETIC ACID (Ph-CH₂COOH)¹³
16. ¹³C-DIPHENYLACETIC ACID ((Ph)₂CHCOOH)¹³
17. ¹³C-p-TOLYLACETIC ACID (CH₃-Ph-CH₂-COOH)¹³
18. ¹³C-7-AMINOHEPTANOIC ACID (H₂N-CH₂(CH₂)₄CH₂-COOH)¹³
19. ¹³C-ETHOXYACETIC ACID (CH₃CH₂-O-CH₂COOH)¹³
20. ¹³C-p-IODOBENZOIC ACID (I-Ph-COOH)¹³
21. ¹³C-3-5-DIBROMOBENZOIC ACID (Br₂-Ph-COOH)¹³
22. ¹³C-SUCCINIC ACID (HOOC-CH₂CH₂-COOH)¹³
23. ¹³C-ISOPHTHALIC ACID (HOOC-Ph-COOH)¹³
24. ¹³C-TETRAFLUOROPHTHALIC ACID (F₄-Ph(COOH)₂)¹³
25. ¹³C-4-METHYLVALERIC ACID ((CH₃)₂CH-CH₂CH₂-COOH)¹³
26. ¹³C-OLEIC ACID (C₁₈H₃₄O₂)¹³
27. ¹³C-2-CHLOROPROPIONIC ACID (CH₃CHClCOOH)¹³
28. ¹³C-DIBROMOACETIC ACID (Br₂CHCOOH)¹³

ÉSTERES

1. "CHLOROFORMIC ACID - METHYL ESTER (C₂H₃ClO₂)"
2. "FORMIC ACID - ETHYL ESTER (CH₃CH₂-O-COH)"
3. "FORMIC ACID - PROPYL ESTER (CH₃CH₂CH₂-O-COH)"
4. "CHLOROFORMIC ACID - 3-CHLOROPROPYL ESTER"
5. "FORMIC ACID - BUTYL ESTER (CH₃(CH₂)₃-O-COH)"
6. "FORMIC ACID - ISOBUTYL ESTER (C₅H₁₀O₂)"
7. "FORMIC ACID - HEXYL ESTER (CH₃(CH₂)₄CH₂OCOH)"
8. "FORMIC ACID - ALLYL ESTER (H₂C=CHCH₂OCOH)"
9. "DIFLUOROACETIC ACID - ETHYL ESTER (C₄H₆F₂O₂)"
10. "DICHLOROACETIC ACID - ETHYL ESTER (C₄H₆Cl₂O₂)"
11. "DIBROMOACETIC ACID - ETHYL ESTER (C₄H₆Br₂O₂)"
12. "ACETIC ACID - ISOPENTYL ESTER (C₇H₁₄O₂)"
13. "PROPIONIC ACID - METHYL ESTER (C₄H₈O₂)"
14. "3-BROMOPROPIONIC ACID - METHYL ESTER"
15. "2-BROMOPROPIONIC ACID - ETHYL ESTER"
16. "PROPIONIC ACID - ISOBUTYL ESTER (C₇H₁₄O₂)"
17. "PROPIONIC ACID - ISOPENTYL ESTER (C₈H₁₆O₂)"
18. "PROPIONIC ACID - ALLYL ESTER (C₆H₁₀O₂)"
19. "BUTYRIC ACID - METHYL ESTER (C₅H₁₀O₂)"
20. "ISOBUTYRIC ACID - METHYL ESTER"
21. "2-BROMOBUTYRIC ACID - ETHYL ESTER"
22. "2-ETHYLHEXANOIC ACID - METHYL ESTER (C₉H₁₈O₂)"
23. "2-BROMOHEXANOIC ACID - ETHYL ESTER"
24. "6-BROMOHEXANOIC ACID - ISOPROPYL ESTER"
25. "HEXANOIC ACID - PENTYL ESTER (C₁₁H₂₂O₂)"
26. "HEPTANOIC ACID - ETHYL ESTER (C₉H₁₈O₂)"
27. "2-HEPTENOIC ACID - METHYL ESTER (C₈H₁₄O₂)"
28. "OCTANOIC ACID - ISOPROPYL ESTER (C₁₁H₂₂O₂)"
29. "OCTANOIC ACID - OCTYL ESTER (C₁₆H₃₂O₂)"
30. "o-BROMOBENZOIC ACID - METHYL ESTER (C₈H₇O₂)"
31. "3-5-DIBROMOBENZOIC ACID - ETHYL ESTER"
32. "2-BROMO-3-METHYLBUTYRIC ACID - ETHYL ESTER"
33. "3-5-DICHLORO-4-NITROBENZOIC ACID - ETHYL ESTER"
34. "p-NITROBENZOIC ACID - ETHYL ESTER (C₉H₉O₄)"
35. "PHTHALIC ACID - MONOMETHYL ESTER (C₉H₈O₄)"
36. "PHTHALIC ACID - DIPROPYL ESTER (C₁₄H₁₈O₄)"
37. "PHTHALIC ACID - MONOBENZYL ESTER (C₁₅H₁₂O₄)"
38. "PALMITIC ACID - ETHYL ESTER (C₁₈H₃₂O₂)"
39. "PALMITIC ACID - PROPYL ESTER (C₁₉H₃₈O₂)"
40. "PALMITIC ACID - BUTYL ESTER (C₂₀H₄₀O₂)"
41. "PALMITIC ACID - DECYL ESTER (C₂₆H₅₂O₂)"
42. "LAURIC ACID - BUTYL ESTER (C₁₆H₃₂O₂)"
43. "FUMARIC ACID - DIETHYL ESTER (C₈H₁₂O₄)"
44. "(+)-TARTARIC ACID - DIETHYL ESTER (C₈H₁₄O₆)"

CETONAS

1. "ACETONE (CH₃COCH₃)"
2. "2-BUTANONE (CH₃COCH₂CH₃)"
3. "3-3-DIMETHYL-2-BUTANONE (C₆H₁₂O)"
4. "CYCLOPENTANONE (C₅H₈O)"
5. "2-ETHYLCYCLOHEXANONE (C₈H₁₄O)"
6. "2-BUTYLCYCLOHEXANONE (C₁₀H₁₈O)"
7. "1-CYCLOHEXYLCYCLOHEXANONE (C₁₂H₂₀O)"
8. "2-DECANONE (C₁₀H₂₀O)"
9. "2-DODECANONE (C₁₂H₂₄O)"
10. "3-DODECANONE (C₁₂H₂₄O)"

CETONAS

11. ¹³5-DODECANONE (C₁₂H₂₄O)¹³
12. ¹³4-ETHYLCYCLOHEXANONE (C₈H₁₄O)¹³
13. ¹³CYCLOHEXANONE (C₆H₁₀O)¹³
14. ¹³2-5-DIMETHYLCYCLOPENTANONE (C₇H₁₂O)¹³
15. ¹³ACETONE-d₆ (C₃D₆O)¹³
16. ¹³3-BUTEN-2-ONE (CH₂=CH(CO)-CH₃)¹³
17. ¹³2-OCTANONE (C₈H₁₆O)¹³
18. ¹³CYCLOPROPYL METHYL KETONE (C₅H₈O)¹³
19. ¹³CYCLOPROPYL KETONE (C₇H₁₀O)¹³
20. ¹³1-CHLORO-2-PROPANONE (CH₃-(CO)-CH₂-Cl)¹³
21. ¹³1-3-DIFLUORO-2-PROPANONE (F-CH₂-(C=O)-CH₂-F)¹³
22. ¹³1-3-DICHLORO-2-PROPANONE (C₃H₄Cl₂O)¹³
23. ¹³1-CYCLOPENTHYL-2-PROPANONE (C₈H₁₄O)¹³
24. ¹³1-3-DIPHENYL-2-PROPANONE (C₁₅H₁₄O)¹³

HALETOS DE ACILA

1. ¹³ACETYL-d₃ CHLORIDE (CD₃COCl)¹³
2. ¹³3-BROMOPROPIONYL CHLORIDE (Br-CH₂CH₂COCl)¹³
3. ¹³PIVALOYL CHLORIDE ((t-BUT)COCl)¹³
4. ¹³UNDECANOYL CHLORIDE (CH₃CH₂(CH₂)₇CH₂COCl)¹³
5. ¹³METHYLENESUCCINYL CHLORIDE (C₅H₄Cl₂O₂)¹³
6. ¹³GLYTARYL CHLORIDE (ClCOCH₂CH₂CH₂COCl)¹³
7. ¹³PIMELOYL CHLORIDE (ClCOCH₂(CH₂)₃CH₂COCl)¹³
8. ¹³SUBEROYL CHLORIDE (ClCOCH₂(CH₂)₄CH₂COCl)¹³
9. ¹³4-(CHLOROFORMYL)BUTYRIC ACID - METHYL ESTER¹³
10. ¹³CHLOROFORMIC ACID -2-2-2-TRICHLOROETHYL ESTER¹³
11. ¹³CHLOROFORMIC ACID -DOCECYL ESTER (C₁₃H₂₅ClO₂)¹³
12. ¹³CHLOROFORMIC ACID - p-NITROBENZYL ESTER¹³
13. ¹³CHLOROFORMIC ACID - PHENYL ESTER (C₇H₅ClO₂)¹³
14. ¹³CHLOROGLYOXYLIC ACID - ETHYL ESTER (C₄H₅ClO₃)¹³
15. ¹³BENZOYL CHLORIDE (C₇H₅ClO)¹³
16. ¹³BENZOYL BROMIDE (C₇H₅BrO)¹³
17. ¹³o-NITROBENZOYL CHLORIDE (C₇H₄ClNO₃)¹³
18. ¹³o-TOLUOYL CHLORIDE (C₈H₇ClO)¹³

ALDEÍDOS

1. ¹³DIPHENYLACETALDEHYDE (C₁₄H₁₂O)¹³
2. ¹³PROPIONALDEHYDE (CH₃CH₂COH)¹³
3. ¹³CROTONALDEHYDE (CH₃CH=CHCOH)¹³
4. ¹³CINNAMALDEHYDE (Ph-CH=CHCOH)¹³
5. ¹³2-ETHYL-3-METHYLBUTYRALDEHYDE (C₇H₁₄O)¹³
6. ¹³A-BUTYLCINNAMALDEHYDE (C₁₃H₁₆O)¹³
7. ¹³2-METHYLVALERALDEHYDE (C₆H₁₂O)¹³
8. ¹³3-METHYLVALERALDEHYDE (C₆H₁₂O)¹³
9. ¹³HEXANAL (CH₃CH₂(CH₂)₂CH₂COH)¹³
10. ¹³2-ETHYL-2-HEXENAL (C₈H₁₄O)¹³
11. ¹³trans-3-7-DIMETHYL-6-OCTENAL (C₁₀H₁₈O)¹³
12. ¹³3-7-DIMETHYL-2-6-OCTADIENAL (C₁₀H₁₆O)¹³
13. ¹³NONANAL (CH₃(CH₂)₆CH₂COH)¹³
14. ¹³DECENAL (CH₃CH₂(CH₂)₆CH₂COH)¹³
15. ¹³10-UNDECENAL (CH₂=CHCH₂(CH₂)₆CH₂COH)¹³
16. ¹³MYRISTALDEHYDE (CH₃CH₂(CH₂)₁₀CH₂COH)¹³
17. ¹³BENZALDEHYDE (PhCOH)¹³
18. ¹³2-5-DIMETHYLBENZALDEHYDE (C₉H₁₀O)¹³
19. ¹³p-ISOPROPYLBENZALDEHYDE (C₁₀H₁₂O)¹³
20. ¹³o-NITROBENZALDEHYDE (C₇H₅NO₃)¹³

ALDEÍDOS

21. ^m-NITROBENZALDEHYDE (C₇H₅NO₃)^a
22. ^p-NITROBENZALDEHYDE (C₇H₅NO₃)^a
23. ^o-TOLUALDEHYDE (C₈H₈O)^a
24. ^m-TOLUALDEHYDE (C₈H₈O)^a
25. ^p-TOLUALDEHYDE (C₈H₈O)^a
26. ^hHEPTANAL (CH₃CH₂(CH₂)₃CH₂COH)^a
27. ⁱISOBUTYRALDEHYDE (C₄H₈O)^a
28. ²-ETHYLHEXANAL^a

ÁLCOOIS

1. ^p-FLUOROBENZYL ALCOHOL (C₇H₇FO)^a
2. ²-CHLOROETHANOL (Cl-CH₂-CH₂-OH)^a
3. ³-BROMO-1-PROPANOL (Br-CH₂CH₂CH₂-OH)^a
4. ¹⁻³-EPOXY-2-PROPANOL (C₃H₆O₂)^a
5. ¹⁻³-DIAMINO-2-PROPANOL (NH₂-CH₂-CH(OH)CH₂-NH₂)^a
6. ²-AMINO-1-BUTANOL (CH₃-CH₂-CH(NH₂)CH₂-OH)^a
7. ²-PENTANOL (CH₃CH₂CH₂-CH(OH)CH₃)^a
8. ³-PENTANOL (CH₃CH₂-CH(OH)CH₂CH₃)^a
9. ²⁻³-DIMETHYL-3-PENTANOL (C₇H₁₆O)^a
10. ³-ETHYL-3-PENTANOL (C₇H₁₆O)^a
11. ^hHEXYL ALCOHOL (C₆H₁₄O)^a
12. ³-OCTANOL (C₈H₁₈O)^a
13. ²-METHYL-2-PROPEN-1-OL (CH₂=C(CH₃)CH₂-OH)^a
14. ²-BUTEN-1-OL (CH₃-CH=CH-CH₂-OH)^a
15. ^{cis-3}-HEXEN-1-OL (C₆H₁₂O)^a
16. ¹-OCTEN-3-OL (C₈H₁₆O)^a
17. ²-PENTYN-1-OL (CH₃CH₂-CC-CH₂-OH)^a
18. ¹⁻³-PROPANEDIOL-DIBUTYRATE (C₁₁H₂₀O₄)^a
19. ²-AMINO-2-ETHYL-1-3-PROPANEDIOL (C₅H₁₃NO₂)^a
20. ⁴-CHLOROCYCLOHEXANOL (C₆H₁₁ClO)^a
21. ^o-HYDROXYBENZYL ALCOHOL (C₇H₈O₂)^a
22. ^m-HYDROXYBENZYL ALCOHOL (C₇H₈O₂)^a
23. ^o-FLUOROBENZYL ALCOHOL (C₇H₇FO)^a
24. ^p-FLUOROBENZYL ALCOHOL (C₇H₇FO)^a
25. ^o-METHOXY-^a-METHYLBENZYL ALCOHOL (C₉H₁₂O₂)^a

FENÓIS

1. ^hPHENOL (Ph-OH)^a
2. ²⁻³⁻⁵-TRIMETHYLPHENOL (C₉H₁₂O)^a
3. ²⁻⁴⁻⁵-TRIMETHYLPHENOL (C₉H₁₂O)^a
4. ^p-ETHYLPHENOL (CH₃CH₂-Ph-OH)^a
5. ^p-ISOPROPYLPHENOL (C₉H₁₂O)^a
6. ^m-BUTYLPHENOL (CH₃CH₂CH₂CH₂-Ph-OH)^a
7. ^p-BUTYLPHENOL (CH₃CH₂CH₂CH₂-Ph-OH)^a
8. ^{o-sec}-BUTYLPHENOL (C₁₀H₁₄O)^a
9. ^{6-tert}-BUTYL-^m-CRESOL (C₁₁H₁₆O)^a
10. ^{p-tert}-BUTYLPHENOL (C₁₀H₁₄O)^a
11. ^{p-tert}-PENTYLPHENOL (C₁₁H₁₆O)^a
12. ^{o-tert}-PENTYLPHENOL (C₁₁H₁₆O)^a
13. ^m-DODECYLPHENOL (CH₃(CH₂)₁₆CH₂-Ph-OH)^a
14. ^m-PENTADECYLPHENOL (CH₃(CH₂)₁₃CH₂-Ph-OH)^a
15. ^o-CRESOL (H₃C-Ph-OH)^a
16. ^m-CRESOL (CH₃-Ph-OH)^a
17. ^p-CRESOL (CH₃-Ph-OH)^a
18. ²⁻⁵-XYLENOL (C₈H₁₀O)^a
19. ³⁻⁵-XYLENOL (C₈H₁₀O)^a

FENÓIS

20. ^o-ETHYLPHENOL (C₈H₁₀O)^a
21. ^m-ETHYLPHENOL (C₈H₁₀O)^a
22. ²⁻⁶-DIBROMOPHENOL (C₆H₄Br₂O)^a
23. ²-CHLORO-4-FLUOROPHENOL (C₆H₄ClFO)^a
24. ^m-NITROPHENOL (C₆H₅NO₃)^a
25. ^o-ETHOXYPHENOL (C₈H₁₀O₂)^a
26. ^m-ETHOXYPHENOL (C₈H₁₀O₂)^a

ALCENOS

1. ^{TRICHLOROETHYLENE} (Cl₂C=CHCl)^a
2. ²-METHYLPROPENE (C₄H₈)^a
3. ³-METHYL-1-BUTENE (C₅H₁₀)^a
4. ²-METHYL-2-BUTENE (C₅H₁₀)^a
5. ²⁻⁴-DIMETHYL-1-PENTENE (C₇H₁₄)^a
6. ⁴-METHYL-2-PENTENE (C₆H₁₂)^a
7. ³-ETHYL-1-PENTENE (C₇H₁₄)^a
8. ²⁻⁴-DIMETHYL-2-PENTENE (C₇H₁₄)^a
9. ²-METHYL-1-HEXENE (C₇H₁₄)^a
10. ¹-ETHYLCYCLOHEXENE (C₈H₁₄)^a
11. ²-HEXENE (CH₃CH₂CH₂CH=CHCH₃)^a
12. ^{trans-2-5}-DIMETHYL-3-HEXENE (C₈H₁₆)^a
13. ^{trans-3}-HEPTENE (C₇H₁₄)^a
14. ¹-OCTENE (CH₂=CHCH₂(CH₂)₃CH₂CH₃)^a
15. ²-METHYL-1-OCTENE (C₉H₁₈)^a
16. ²-OCTENE (CH₃(CH₂)₄CH=CHCH₃)^a
17. ³-CHLOROPROPENE (ClCH₂CH=CH₂)^a
18. ¹⁻¹-DICHLOROETHYLENE (CH₂=CCl₂)^a
19. ³-CHLOROPROPENE (Cl-CH₂-CH=CH₂)^a
20. ³-BROMOPROPENE (BR-CH₂-CH=CH₂)^a
21. ³-CHLORO-2-METHYLPROPENE (Cl-CH₂-C-(CH₃)=CH₂)^a
22. ²-METHYL-1-PENTENE (CH₃-CH₂-CH₂-C(CH₃)=CH₂)^a
23. ^{CYCLOHEXENE} (C₆H₁₀)^a
24. ¹-METHYLCYCLOHEXENE (C₇H₁₂)^a
25. ⁴-METHYLCYCLOHEXENE (C₇H₁₂)^a
26. ³-BROMOCYCLOHEXENE (C₆H₉Br)^a

AROMÁTICOS

1. ^{ETHYLBENZENE} (Ph-CH₂CH₃)^a
2. ^m-DIISOPROPYLBENZENE (C₁₂H₁₈)^a
3. ^p-DIISOPROPYLBENZENE (C₁₂H₁₈)^a
4. ^{BUTYLBENZENE} (Ph-CH₂CH₂CH₂CH₃)^a
5. ^{sec}-BUTYLBENZENE (C₁₀H₁₄)^a
6. ^{ISOBUTYLBENZENE} (C₁₀H₁₄)^a
7. ^{tert}-BUTYLBENZENE (C₁₀H₁₄)^a
8. ^p-DI-^{tert}-BUTYLBENZENE (C₁₄H₂₂)^a
9. ^{TOLUENE} (Ph-CH₃)^a
10. ^o-ETHYLTOLUENE (CH₃-Ph-CH₂CH₃)^a
11. ^m-ETHYLTOLUENE (CH₃-Ph-CH₂CH₃)^a
12. ^p-ETHYLTOLUENE (CH₃-Ph-CH₂CH₃)^a
13. ^o-XYLENE (CH₃-Ph-CH₃)^a
14. ^m-XYLENE (CH₃-Ph-CH₃)^a
15. ²-ETHYL-^m-XYLENE (C₁₀H₁₄)^a
16. ^{5-tert}-BUTYL-^m-XYLENE (C₁₂H₁₈)^a
17. ^p-XYLENE (CH₃-Ph-CH₃)^a
18. ¹-PHENYLDECANE (Ph-CH₂(CH₂)₈CH₃)^a
19. ^{CUMENE} (C₉H₁₂)^a

AROMÁTICOS

20. ³3-5-DIMETHYLCUMENE (C₁₁H₁₆)³
21. ⁴4-FLUORO-*o*-XYLENE (C₈H₉F)⁴
22. ⁴4-CHLORO-*o*-XYLENE (C₈H₉Cl)⁴
23. ²2-4-DIMETHOXYTOLUENE (C₉H₁₂O₂)²
24. ^o*o*-FLUOROTOLUENE (C₇H₇F)^o
25. ^m*m*-FLUOROTOLUENE^m
26. ^p*p*-FLUOROTOLUENE (C₇H₇F)^p

ALCINOS

1. ³3-3-DIMETHYL-1-BUTYNE (C₆H₁₀)³
2. ²2-BUTYNE (CH₃CCCH₃)²
3. ¹1-PENTYNE (CH₃CH₂CH₂CCH)¹
4. ⁵5-PHENYL-1-PENTYNE (C₁₁H₁₂)⁵
5. ⁴4-4-DIMETHYL-2-PENTYNE (C₇H₁₂)⁴
6. ¹1-HEXYNE (CH₃(CH₂)₂CH₂CCH)¹
7. ⁶6-PHENYL-2-HEXYNE (Ph-CH₂CH₂CH₂CCCH₃)⁶
8. ¹1-NONYNE (CH₃(CH₂)₅CH₂CCH)¹
9. ¹1-DECYNE (CH₃CH₂(CH₂)₅CH₂CCH)¹
10. ¹1-7-OCTADIYNE (HCCCH₂(CH₂)₂CH₂CCH)¹
11. ¹1-8-NONADIYNE (HCCCH₂(CH₂)₃CH₂CCH)¹
12. ⁹9-9-BIS(2-PROPYNYL)FLUORENE (C₉H₁₄)⁹
13. ⁵(5-HEXYNYL)FLUORENE (C₁₉H₁₈)⁵
14. ¹1-4-DICHLORO-2-BUTYNE (C₄H₄Cl₂)¹
15. ^DDIPHENYLACETYLENE (C₁₄H₁₀)^D
16. ^EETHYLNILBENZENE (C₈H₆)^E
17. ¹(1-PROPYNYL)BENZENE (C₉H₈)¹

NITRILAS

1. ^TTRICHLOROACETONITRILE (CCl₃-CN)^T
2. ^DDIBROMOACETONITRILE (Br₂CHCN)^D
3. ¹1-CYCLOHEXENE-1-ACETONITRILE (Ph-CH₂CN)¹
4. ^PPHENYLACETONITRILE (Ph-CH₂CN)^P
5. ^m(*m*-CHLOROPHENYL)ACETONITRILE (Cl-Ph-CH₂CN)^m
6. ²2-ETHYL-2-PHENYLBUTYRONITRILE (C₁₂H₁₅N)²
7. ^VVALERONITRILE (CH₃CH₂CH₂CH₂CN)^V
8. ^IISOVALERONITRILE (C₅H₉N)^I
9. ⁷7-BROMOHEPTANONITRILE (Br-CH₂(CH₂)₄CH₂CN)⁷
10. ^UUNDECANENITRILE (CH₃(CH₂)₉CN)^U
11. ^BBENZONITRILE (Ph-CN)^B
12. ²2-3-DIMETHYLBENZONITRILE (C₉H₉N)²
13. ²2-4-DIMETHYLBENZONITRILE (C₉H₉N)²
14. ^p*p*-BROMOBENZONITRILE (C₇H₄BrN)^p
15. ⁴4-FLUOROBENZONITRILE (F-Ph-CN)⁴
16. ^p*p*-NITROBENZONITRILE (O₂N-Ph-CN)^p
17. ^o*o*-TOLUNITRILE (CH₃-Ph-CN)^o
18. ^A*A*-BROMO-*o*-TOLUNITRILE (BrCH₂-Ph-CN)^A
19. ^A*A*-BROMO-*p*-TOLUNITRILE (BrCH₂-Ph-CN)^A
20. ²2-NORBORNANECARBONITRILE (C₈H₁₁N)²
21. ³3-BROMOPROPIONITRILE (Br-CH₂-CH₂-CN)³
22. ^BBUTYRONITRILE (CH₃-CH₂-CH₂-CN)^B
23. ²(2-6-DICHLOROPHENYL)ACETONITRILE (C₈H₅Cl₂N)²
24. ^BBENZOYLACETONITRILE (C₉H₇NO)^B
25. ^m*m*-NITROBENZONITRILE (C₇H₄N₂O₂)^m

ALCANOS

1. "2-METHYLBUTANE (C₅H₁₂)"
2. "2-2-DIMETHYLBUTANE (C₆H₁₄)"
3. "2-3-DIMETHYLBUTANE (C₆H₁₄)"
4. "PENTANE (CH₃CH₂CH₂CH₂CH₃)"
5. "2-2-DIMETHYLPENTANE (C₇H₁₆)"
6. "2-4-DIMETHYLPENTANE (C₇H₁₆)"
7. "2-2-4-TRIMETHYLPENTANE (C₈H₁₈)"
8. "2-3-4-TRIMETHYLPENTANE (C₈H₁₈)"
9. "3-ETHYLPENTANE (C₇H₁₆)"
10. "HEXANE (C₆H₁₄)"
11. "2-5-DIMETHYLHEXANE (C₈H₁₈)"
12. "2-2-5-TRIMETHYLHEXANE (C₉H₂₀)"
13. "HEPTANE (C₇H₁₆)"
14. "2-METHYLHEPTANE (C₈H₁₈)"
15. "4-METHYLHEPTANE (C₈H₁₈)"
16. "OCTANE (C₈H₁₈)"
17. "2-METHYLOCTANE (C₉H₂₀)"
18. "4-METHYLOCTANE (C₉H₂₀)"
19. "NONANE (C₉H₂₀)"
20. "3-METHYLNONANE (C₁₀H₂₂)"
21. "4-METHYLNONANE (C₁₀H₂₂)"
22. "5-METHYLNONANE (C₁₀H₂₂)"
23. "DECANE (C₁₀H₂₂)"
24. "HEPTADECANE (C₁₇H₃₆)"

ÉTER

1. "METHYL ETHER (CH₃-O-CH₃)"
2. "ETHYL ETHER (CH₃CH₂-O-CH₂CH₃)"
3. "PROPYL ETHER (CH₃CH₂CH₂-O-CH₂CH₂CH₃)"
4. "BUTYL ETHER (CH₃CH₂CH₂CH₂-O-CH₂CH₂CH₂CH₃)"
5. "HEXYL ETHER (C₁₂H₂₆O)"
6. "DODECYL ETHER (C₂₄H₅₀O)"
7. "HEXADECYL ETHER (C₃₂H₆₆O)"
8. "ETHYL PENTYL ETHER (C₇H₁₆O)"
9. "ETHYL VINYL ETHER (CH₃CH₂-O-CH=CH₂)"
10. "HEXADECYL VINYL ETHER (CH₂=CH₂OCH₂(CH₂)₁₄CH₃)"
11. "ALLYL PHENYL ETHER (Ph-O-CH₂CH=CH₂)"
12. "ALLYL o-TOLYL ETHER (C₁₀H₁₂O)"
13. "ALLYL p-TOLYL ETHER (C₁₀H₁₂O)"
14. "PHENYL PROPYL ETHER (Ph-O-CH₂CH₂CH₃)"
15. "BUTYL PHENYL ETHER (Ph-O-CH₂CH₂CH₂CH₃)"
16. "p-tert-BUTYLPHENYL DODECYL ETHER (C₂₂H₃₈O)"
17. "ISOPROPYL ETHER (C₆H₁₄O)"
18. "PENTHYL ETHER (C₁₀H₂₂O)"
19. "ISOPENTHYL ETHER (C₁₀H₂₂O)"
20. "ALLYL ETHER (CH₂=CH-CH₂-O-CH₂-CH=CH₂)"
21. "PHENYL ETHER (C₁₂H₁₀O)"