

Automatização dos Procedimentos Computacionais que possam ser aplicados no Espalhamento Elástico de Elétrons por Moléculas

Automation of Computational Procedures that Can Be Applied to Elastic Electron Spreading by Molecules

RESUMO

Jordanis Kyriazidis
jordanisk@alunos.utfpr.edu.br
Universidade Tecnológica Federal do Paraná, Ponta grossa, Paraná, Brasil

Abel Dionizio Azeredo
aazeredo@gmail.com
Universidade Tecnológica Federal do Paraná, Ponta grossa, Paraná, Brasil

Para a realização do estudo do Espalhamento Elástico de Elétrons por Moléculas são necessários diversos dados vindos de bancos de dados que são alimentados por pesquisadores. Porém também são necessários softwares como o GAMESS para uma otimização inicial dos dados, e o ALCHEMY para a realização dos cálculos necessários para a pesquisa como por exemplo o cálculo de orbitais virtuais, e também certo conhecimento de linguagem de programação FORTRAN devido a estes softwares utilizarem a linguagem FORTRAN em seu código base. Devido a necessidade da realização de diversos passos repetitivos para cada molécula o pesquisador acaba tendo que estudar os softwares em si, o que não é o foco da pesquisa, visto que é possível automatizar este processo, foi desenvolvido um software na linguagem C capaz de gerar um destes arquivos de entrada visto que os dados estão todos no próprio ALCHEMY após um primeiro processamento, sendo então possível encontrar e organizar estes dados de forma que o pesquisador não precise realizar um estudo aprofundado do código do ALCHEMY.

PALAVRAS-CHAVE: Espalhamento. ALCHEMY. Programa.

Recebido: 19 ago. 2019.

Aprovado: 01 out. 2019.

Direito autoral: Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.



ABSTRACT

To carry out the study of elastic electron spreading by molecules, several data from databases that are fed by researchers are required. However, software such as GAMESS is also required for initial data optimization, and alchemy for performing the necessary research calculations such as virtual orbital calculations, as well as some knowledge of FORTRAN programming language due to the fact that such software uses the FORTRAN language in its base code. Due to the need to perform several repetitive steps for each molecule the researcher ends up having to study the software itself, which is not the focus of the research, since it is possible to automate this process, was developed a software in C language able to generate one of these input files since the data is all in alchemy itself after first processing, so it is possible to find and organize this data so that the researcher does not need to do a thorough study of the ALCHEMY code.

KEYWORDS: Spread. ALCHEMY. Program.

INTRODUÇÃO

Para a realização do estudo do Espalhamento Elástico de Elétrons por Moléculas são utilizados dados de moléculas providos por bancos de dados, por exemplo o CCCBDB (Computational Chemistry Comparison and Benchmark Database) onde pesquisadores inserem diversos dados sobre as mais diversas moléculas, estes dados são otimizados pelo GAMESS (General Atomic and Molecular Electronic Structure System), um software não-livre mas que pode ser obtido para a realização de pesquisas pelo site do próprio GAMESS <http://www.msg.ameslab.gov/gamess/download.html>.

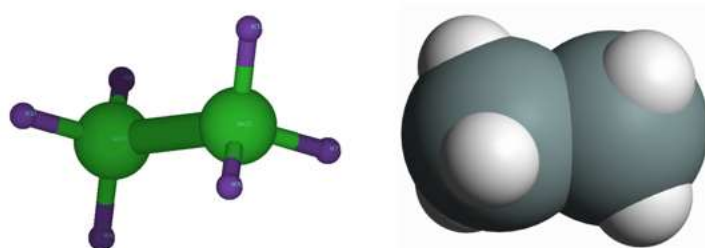
Com os dados já otimizados é realizado o cálculo dos orbitais moleculares, chamados também de orbitais virtuais que são orbitais possíveis de serem ocupados por um elétron do contínuo porém no estado fundamental da molécula encontra-se desocupado. Estes orbitais são calculados utilizando os dados obtidos do banco de dados já otimizados pelo GAMESS e organizados no formato de entrada do ALCHEMY, o ALCHEMY é um software colaborativo existindo diversas versões existentes e nenhuma página para seu download, foi iniciado por M. A. P. Lima na Unicamp e seus alunos de doutorado e pós-doutorado continuaram seu desenvolvimento assim como diversos outros pesquisadores que acabaram por utilizar do software e acabaram realizando alguma alteração para uso próprio não necessariamente compartilhando da sua versão.

Todo o processo da obtenção dos dados até inserir e organizar os dados em arquivos de entrada para a leitura do ALCHEMY é feito pelo usuário, tudo em linguagem FORTRAN, segundo Helder (2003, P. 4) “A formatação dos códigos em FORTRAN, principalmente em formato fixo deve seguir um estilo diferente dos usados na maioria das linguagens de programação. Estes conceitos iniciais podem não ficar claro para os iniciantes na linguagem”, exigindo que o usuário conheça a linguagem, onde podem detalhes como a linguagem poder ser escrita em formulário livre ou fixo acabam sendo cruciais para que não ocorram erros. Automatizar estes procedimentos foi o objetivo proposto para esta etapa do projeto.

METODOLOGIA

O GAMESS além da otimização da geometria molecular também é capaz de gerar uma visualização 3-D da molécula e por consequência determinar seu grupo de simetria.

Figura 1 – Representação artística da molécula de Ge₂H₆, obtida via Gamess.



Fonte: Produção própria.

Além de ser possível gerar estas imagens, o GAMESS acaba também gerando o input inicial do ALCHEMY para então ser possível dar início ao cálculo dos orbitais moleculares, este arquivo de saída do GAMESS contem diversos dados necessários para o estudo como por exemplo nome, simetria, posição espacial, que unidades de medida devem ser adotadas, etc, todos sendo importantes para a construção do arquivo de entrada do ALCHEMY.

Para a organização e construção dos arquivos de entrada do ALCHEMY são utilizados os dados obtidos após execuções do ALCHEMY ou do GAMESS, todos são construídos em FORTRAN e tem suas próprias extensões para que o ALCHEMY os reconheça. Visto que estes arquivos de entrada são arquivos de texto com dados que estão contidos no próprio ALCHEMY foi visto a possibilidade de criar um software na linguagem C[3], Renato A. (2006, p. 4) afirma que “As variáveis e constantes são os tipos de dados básicos manipulados num programa.”, sabendo disso a linguagem foi utilizada para que organize estes dados e construa os arquivos além de executar os comandos do ALCHEMY, ou seja, fazer a transição de dados de saída de um software para dados de entrada de outro e em seguida executar um comando para que o software realize os cálculos e assim seguir as etapas seguintes necessárias.

Figura 2 – Arquivo de entrada(intpp.lis) ainda não organizado.

```

1007      160      F2      NUB      YY      1 160      0.14441855
1008      161      F2      NUB      YY      1 161      0.14441855
1009      162      F2      NUB      YZ      1 162      0.14441855
1010      163      F2      NUB      YZ      1 163      0.14441855
1011      164      F2      NUB      ZZ      1 164      0.14441855
1012      165      F2      NUB      ZZ      1 165      0.14441855
1013      1
1014
1015
1016      | SYMMETRY TRANSFORMATION INFO
1017      | 165 SYMMETRY ADAPTED BASIS FUNCTIONS
1018
1019      | IRREDUCIBLE REPRESENTATION NUMBER 1
1020
1021      1      R1      S      1 1      1.00000000
1022      2      R1      S      1 2      1.00000000
1023      3      R1      S      1 3      1.00000000
1024      4      R1      S      1 4      1.00000000
1025      5      R1      S      1 5      1.00000000

```

Fonte: Produção própria.

Figura 3 – Arquivo de entrada (btran.inp) organizado.

```

1 |s
2 R1 1 3 6 5 2
3 C 1 3 6 5 2
4 F1 1 3 6 5 2
5 F2 2 3 6 5 2
6 2
7 105 60
8
9
10
11
12 1 R1 S 1 1 1.00000000
13 2 R1 S 1 2 1.00000000
14 3 R1 S 1 3 1.00000000
15 4 R1 S 1 4 1.00000000
16 5 R1 S 1 5 1.00000000
17 6 R1 S 1 6 1.00000000
18 7 R1 Y 1 8 1.00000000
19 8 R1 Z 1 9 1.00000000
20 9 R1 Y 1 11 1.00000000
21 10 R1 Z 1 12 1.00000000
22 11 R1 Y 1 14 1.00000000
23 12 R1 Z 1 15 1.00000000
  
```

Fonte: Produção própria.

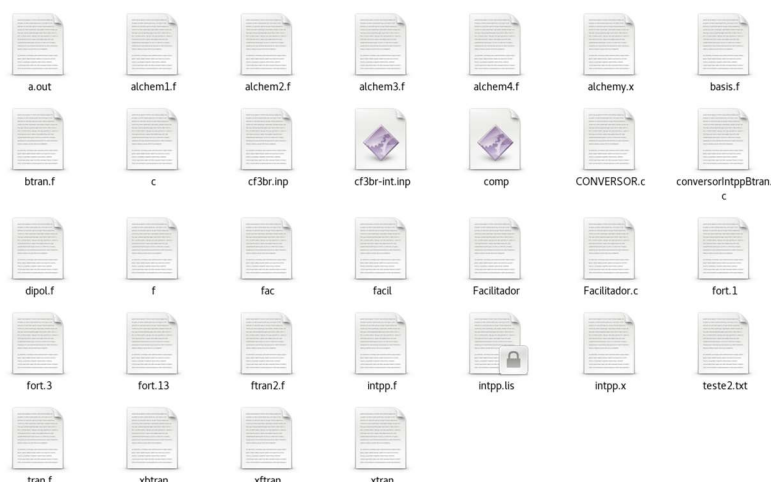
Como pode ser visto na figura 2 os dados necessários para a montagem do arquivo para a entrada seguinte esta somente após a linha 1019, então foi necessário uma extensa análise dos arquivos de dados e como estão organizados até mesmo os dados que estão implícitos dentro dos diversos arquivos que são gerados para então criar um programa que encontre estes dados e construa o próximo arquivo de entrada como pode ser visto na figura 3, onde estão contidos dados de diferentes arquivos gerados pelo ALCHEMY porem organizados como uma nova entrada que pode ser executada logo a seguir.

RESULTADOS E DISCUSSÃO

Um software foi desenvolvido para automatização do processo, que é lento e demorado devido a necessidade de ser construído manualmente cada arquivo de entrada de cada etapa, sendo preciso encontrar dados em arquivos extensos com muitos dados e sem muita descrição para somente então avançar para a próxima etapa do processamento computacional da pesquisa.

Como por exemplo do processo de construção do arquivo BTRAN.inp por meio do software que coleta os dados tanto do arquivo de saída do GAMESS quanto do arquivo INTPP.lis que é um arquivo de saída do ALCHEMY logo após a primeira etapa do processamento. No INTPP.lis estão contidos dados sobre os orbitais virtuais e também dos dados que compõem a molécula, a distribuição de elétrons, o numero de funções s p e d assim com o número de funções por representação irreduzível do grupo de simetria da molécula.

Figura 4 – Arquivos do ALCHEMY com entradas e saídas já inclusas.



Fonte: Produção própria.

A construção é feita com os diversos arquivos de entrada e saída que o ALCHEMY gera como pode ser visto na figura 4.

Figura 5 – Parte do código do conversor.

```

59     if(strstr(linha,"NUCLEAR REPULSION ENERGY")){
60         total_dados = cont_copia;
61         cont_copia = 0;
62     }
63     //Começa a filtrar
64     if (cont_copia > 0){
65         //Conta o numero de representações
66         if(strlen(linha) > 2 && !strstr(linha,"IRREDUCIBLE"))
67             dado[cont_copia].quantidade++;
68     }
69     //Encontrando as moléculas
70     if (!strstr(linha, "IRREDUCIBLE")){
71         //Encontra os dados da molécula e coloca na auxiliar
72         aux[0] = linha[10];
73         aux[1] = linha[11];
74         strcat(aux, "\0");
75         //Verifica se a molécula já foi adicionada ao conjunto
76         controle_molecula = 0;
77         for (int i = 0; i < 10; ++i){
78             if(strcmp(aux, molecula[i]) == 0 || strcmp(aux, "IR") == 0)
79                 controle_molecula = 1;
80         }
81         if (controle_molecula == 0){
82             if(strcmp(aux, "IR") != 0){
83                 strcpy(molecula[j], aux);
84                 //printf("%s\n", molecula[j]);
85                 j++;
86             }
87         }
88     }
89 }
90
91 //fim das moléculas
92
93 //Copia as linhas da representação correspondente
94 strcat(dado[cont_copia].conteudo, linha);
95

```

Fonte: Produção própria.

Como pode ser visto na figura 5 parte do código do conversor que faz esta busca e organização dos arquivos.

CONCLUSÃO

Foi possível desenvolver um programa que consegue organizar uma boa parte do processo da rotina de montagem de dados mas devido a falta de uma maior quantidade de dados brutos, apenas foram testadas moléculas pequenas, ainda não é possível determinar o limite de dados que podem ser processados, também devido a alta complexidade e a falta de compreensão em relação aos dados que são necessários e onde podem ser obtidos entre os arquivos, para que servem e

como são utilizados nos processos, sendo somente assim possível desenvolver um software que consiga construir novas entradas de forma eficiente e capaz de processar uma grande quantidade de dados como foi feito para esta etapa.

Devido o curto tempo de desenvolvimento não foi possível desenvolver todo o processo da rotina, saindo de um arquivo bruto de dados para os dados necessários para estudo.

AGRADECIMENTOS

Agradeço a UTPR pela possibilidade do desenvolvimento do projeto e ao CNPQ pelo incentivo.

REFERÊNCIAS

SCHMIDT, M. W.; BALDRIDGE, K. K.; BOATZ, J. A.; ELBERT, S. T.; GORDON, M. S.; JENSEN, J. H.; KOESKI, S.; MATSUNAGA, N.; NGUYEN, K. A.; SU, S. J.; WINDUS, T. L.; DUPUIS M.; MONTGOMERY, J. A. GAMESS: General Atomic and Molecular Eletronic Structures Systems.

CRISTO, Helder Pereira; HORIZONTE, Belo. Programação em Linguagem FORTRAN. Arquivo Livre: Fortran. pdf, 2003.

SILVA, Renato A.; PROGRAMANDO MICROCONTROLADOR, P. I. C. Linguagem C. São Paulo: Ensino Profissional, 2006.