



Design de Estruturas Bidimensionais de Carbono e Nitrogênio com o uso de Algoritmos Genéticos

O grafeno é um material bidimensional que tem chamado a atenção de todo o mundo por ser o material mais fino, forte e leve. Além disso é flexível, impermeável a moléculas e é ótimo condutor térmico, sendo usado para a fabricação de capacetes, tintas, preservativos, embalagens, entre outras aplicações da indústria. Porém, o grafeno puro não tem aplicações na indústria eletrônica por não ser capaz de conduzir eletricidade e, sendo assim, é necessário que as estruturas do grafeno sejam dopadas com átomos diferentes para que se torne capaz de conduzir eletricidade.

O programa proposto neste projeto pretende dopar as cadeias de carbono com átomos de nitrogênio a fim de manipular as propriedades eletrônicas do grafeno. A distribuição dos átomos de nitrogênio é fundamental para que algumas propriedades eletrônicas sejam garantidas, mas ao mesmo tempo, existem incontáveis possibilidades de *design*.

Este programa utilizará de métodos probabilísticos para encontrar uma boa dopagem para propriedades eletrônicas pré-definidas pelo usuário, a fim de reduzir o custo computacional envolvido no *design* de estruturas de carbono e nitrogênio.

Metodologia

O método probabilístico utilizado para a realização da busca e das otimizações será o Algoritmo Genético. Os parâmetros utilizados para verificar as características eletrônicas de cada solução são baseadas no Método de Hückel.

Algoritmo genético

O algoritmo genético é um tipo de algoritmo usado para busca e otimização. É usada uma população inicial de possíveis soluções, baseada em uma codificação do conjunto de parâmetros, ou seja, os parâmetros são escritos em códigos e a busca pela solução ótima é feita com o uso da probabilidade, gerando e testando populações de indivíduos.

Por mais que a busca seja feita de forma aleatória, ela é estruturada de modo a buscar pontos de alta aptidão, onde os valores da função a ser minimizada (ou maximizada)

sejam mínimos (ou máximos), encontrando assim uma solução que corresponda aos parâmetros desejados.

O algoritmo roda a partir de iterações, que são chamadas gerações, onde a população inicial é recombina formando novas populações baseadas nos princípios de seleção e reprodução pré-determinados. Os indivíduos com a maior índice de aptidão tem maior chance de se reproduzir, gerando populações cada vez mais aptas até que o índice de aptidão requerido seja alcançado.

Cada indivíduo tem seu fenótipo codificado como um vetor binário, onde 0 ou 1 representam a ausência ou presença de determinada característica (genótipo). O fenótipo da geração seguinte deve ser totalmente novo, mas preservando características das gerações anteriores. Para isso, é utilizado o cruzamento (preservação das características) e a mutação (diversidade genética) como operadores da iteração.

O procedimento do algoritmo genético inicia com a definição de uma população inicial (de indivíduos formados aleatoriamente) e um parâmetro temporal. A população é avaliada e é definido a condição de parada da iteração. Após isso é feita a seleção dos pais, recombinação, mutação e então a nova população é avaliada para a seleção dos indivíduos que sobrevivem para reiniciar o ciclo, como pode ser visto na Figura 1.

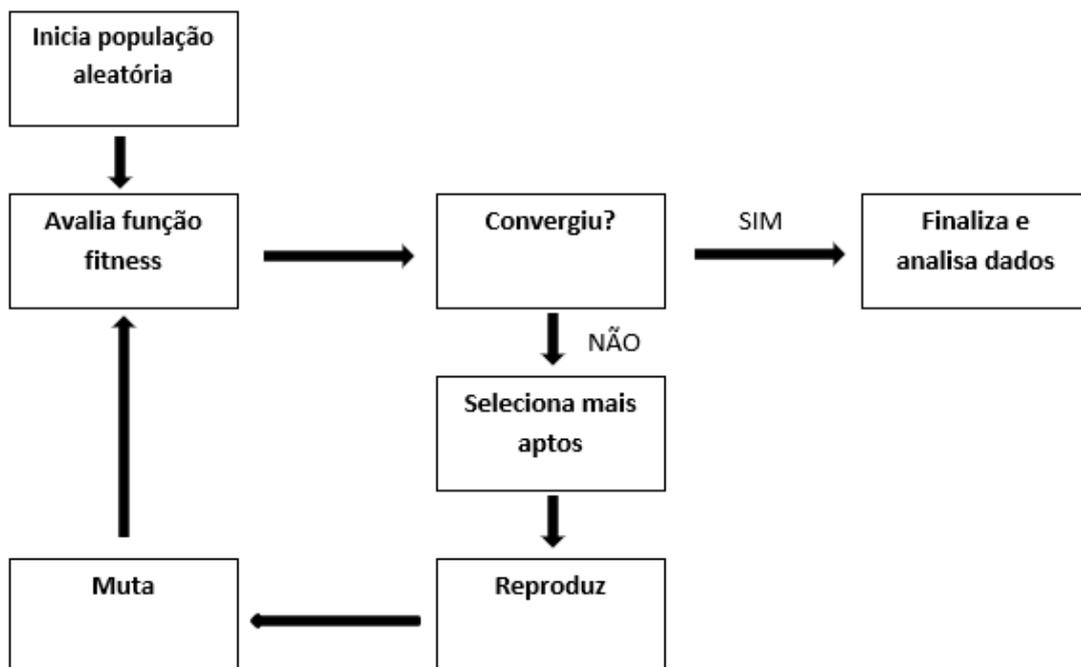


Figura 1: Esquema geral do algoritmo genético.

A população inicial é finita e gerada aleatoriamente, o que pode convergir para soluções não gerais, conduzindo as gerações para mínimos (ou máximos) locais. Para evitar isso,

a mutação introduz novos elementos na população a partir de uma taxa de mutação, que geralmente é baixa, por se tratar de um operador secundário.

Para a otimização em si usa-se a recombinação, considerado operador principal da iteração, dado pela taxa de *crossover*. A taxa deve ser relativamente alta em relação à taxa de mutação, podendo ser utilizada com um ponto de cruzamento, vários pontos de cruzamento ou até mesmo com um parâmetro probabilístico para cada gene.

O método de Hückel

O método de Hückel utiliza consecutivas aproximações para caracterizar os orbitais moleculares e suas energias para moléculas de hidrocarbonetos conjugados. As aproximações de Hückel permitem que se reduza os cálculos de orbitais moleculares à álgebra linear, fornecendo características eletrônicas da molécula de forma compacta, sem grande custo computacional.

Aproximações de Hückel

As aproximações de Hückel são simplificações que permitem:

- a) Visto que as moléculas têm ligações σ e π ortonormais, quando $i = j \mid s_{ij} \in S, s_{ij} = 1$ e quando $i \neq j \mid s_{ij} \in S, s_{ij} = 0$. Portanto, a sobreposição é simplificada numa matriz identidade ($S = I$);
- b) parametrizar H_{ij} com as integrais de Coulomb (α) e ressonância (β) que dizem o quão forte é a interação entre o elétron e o átomo em que está (Integral de Coulomb) ou com os demais átomos (Integral de ressonância).

i. quando $i = j$,

$$H_{ij} = \alpha,$$

ii. quando H_{ij} está nas vizinhas de α ,

$$H_{ij} = \beta;$$

- c) desconsiderar as ligações não vizinhas, tornando o resto da matriz que não é nem α , nem β , igual a zero.

GAP e IPN

As características macroscópicas da molécula estão intimamente relacionadas com o GAP de energia, que é dado pela subtração da energia do LUMO (menor orbital molecular desocupado) pela energia do HOMO (maior orbital molecular ocupado).

O cálculo da localização dos elétrons também é necessário para garantir a condutividade da estrutura desenhada. Para isso é usado o IPN (*Inverse Participation Number*) que varia de 0 (totalmente delocalizado) a 1 (totalmente localizado) em um orbital atômico.

O programa

Para a implementação do programa é necessário determinar os parâmetros do algoritmo genético como a representação dos genes e a função fitness, além de definir a Matriz de Hückel.

Representação dos genes

O benzeno é a unidade mínima, visto que ao anexar vários benzenos nós podemos dar origem ao grafeno. Utilizando-se o benzeno como unidade básica, limitamos os grafenos a terem $2(n+1)^2-2$ átomos e evita-se a presença de átomos nas pontas sem ligações. Isso garante a coerência química dos grafenos estudados e possibilita a utilização da álgebra linear (Método de Hückel) ao problema.

Os genes, como explicado na teoria, são descritos por vetores binários, onde 0 significa a presença de um átomo de carbono e 1 representa a presença de um átomo de nitrogênio, como ilustrado na Figura 2. A Figura 3, mostra os vetores descritos com base na posição dos átomos, partindo da esquerda para direita, de cima pra baixo, respectivamente. Nesta figura, os espaços foram incluídos para facilitar a visualização das linhas.

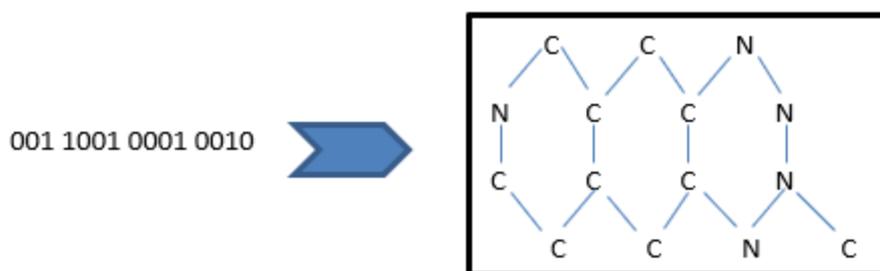


Figura 2. Exemplo da codificação da estrutura do benzeno com carbono e nitrogênio.

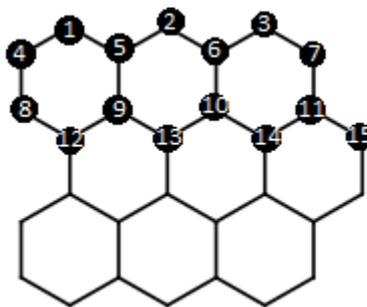


Figura 3. Numeração da cadeia de benzeno para a escrita dos vetores.

Função fitness

A função fitness se trata de uma função avaliadora, que quantifica a aptidão de cada indivíduo da população. Neste caso os parâmetros necessários são: o GAP relativo, o IPN e a presença de ligações quimicamente inviáveis, ou seja, nitrogênios com dois ou mais vizinhos.

A matriz de Hückel é capaz de nos fornecer o HOMO e o LUMO a partir dos autovalores e autovetores da matriz. Sabendo disso e contabilizando o número de elétrons excedentes é possível calcular o GAP eletrônico e relativizá-lo, para estabelecer uma relação entre o GAP atual e o pré-estipulado a partir da expressão abaixo

$$(GAP\ RELATIVO) = | (GAP\ ATUAL) - (GAP\ DESEJADO) | / (GAP\ DESEJADO) |$$

Esses valores são suficientes para se determinar o IPN.

Para cumprir os objetivos desse projeto o peso do GAP relativo deve ser maior que o peso dado ao IPN, assim como o número de ligações inviáveis deve ser minimizada, sem excluir possíveis soluções ótimas. Os valores utilizados são:

$$fitness = 0.7(GAP\ RELATIVO) + 0.2(IPN) + 0.2(LIGAÇÕES\ INVIÁVEIS)$$

A estrutura resultante, utilizando os valores acima, apresentou um elevado número de átomos de nitrogênio com dois ou mais vizinhos, o que representa ligações químicas inviáveis. Apesar disso, o custo computacional foi baixo (10 minutos), o que se deve ao fato de terem sido realizados testes em baixa escala (2-grafeno e 3-grafeno).