Complexidade Computacional do PGDM

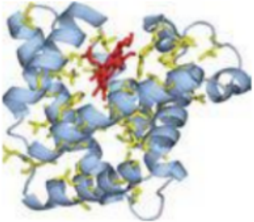
Guilherme Philippi

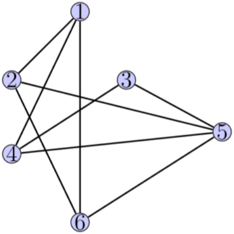
Acadêmico de Engenharia de Controle e Automação
Campus Blumenau
Universidade Federal de Santa Catarina
UFSC
Orientado por Felipe Delfini Caetano Fidalgo

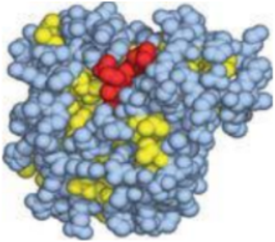
SIC 2018
Seminário de Iniciação Científica
Blumenau - Santa Catarina - Brasil

Preliminares

Problema de Geometria de Distâncias Moleculares







Preliminares

Teoria de Complexidade Computacional

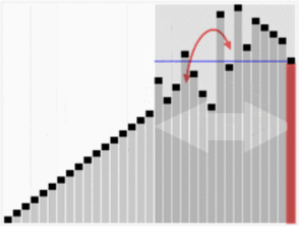
Trata-se da *contagem de operações* necessárias para solucionar um problema, que podem variar abruptamente de caso a caso.

Nos preocupamos em estudar os seguintes:

- Melhor Caso
- Pior Caso
- Caso Médio

Notação **Big O**

$$T(n) = n^3 + 7n^2 + 40\ln(n)^3 \Rightarrow T(n) = O(n^3)$$



Modelando o Problema

Quais são as Entradas?

- Quantidade n de átomos;
- Sequência teórica de átomos;
- Distância teórica entre átomos;
- Distância real entre átomos;

Quais são as Saídas?

- Posições $x_1, ..., x_n \in \mathbb{R}^3$ dos n átomos da molécula.

Perceba, as distâncias reais contém incertezas associadas, logo, a RMN trás *distâncias intervalares* da forma: $[\underline{d_{i,j}}, \overline{d_{i,j}}]$, ou seja

$$0 \leq \underline{d_{i,j}} \leq d_{i,j} \leq \overline{d_{i,j}}$$

Analisando um Caso Particular

Considere um **PGDM plano**, onde temos um grafo $V(G) = \{u, v, r, s\}$ e $E(G) = \{\{u, v\}, \{u, r\}, \{v, r\}, \{v, s\}, \{r, s\}, \{u, s\}\}$. **Fixando** u, v, r, ou seja, **determinando** $x_u, x_v, x_r \in \mathbb{R}^2$ de tal modo que $\|x_u - x_v\| = d_{uv}$, $\|x_u - x_r\| = d_{ur}$ e $\|x_v - x_r\| = d_{vr}$, podemos montar o seguinte sistema quadrático:

$$\|x_s - x_u\| = d_{us}$$

$$\|x_s - x_v\| = d_{vs}$$

$$\|x_s - x_r\| = d_{rs}$$

Analisando um Caso Particular

Elevando os termos ao quadrado,

$$\|x_s\|^2 - 2(x_s.x_u) + \|x_u\|^2 = d_{us}^2$$
$$\|x_s\|^2 - 2(x_s.x_v) + \|x_v\|^2 = d_{vs}^2$$
$$\|x_s\|^2 - 2(x_s.x_r) + \|x_r\|^2 = d_{rs}^2$$

subtraindo a primeira equação das outras duas, temos:

$$2(x_s.x_v) - 2(x_s.x_u) = \|x_v\|^2 - \|x_u\|^2 + d_{us}^2 - d_{vs}^2$$

$$2(x_s.x_r) - 2(x_s.x_u) = \|x_r\|^2 - \|x_u\|^2 + d_{us}^2 - d_{rs}^2$$

Colocando x_s em evidência para ficar claro as incógnitas

$$2(x_v - x_u)x_s = \|x_v\|^2 - \|x_u\|^2 + d_{us}^2 - d_{vs}^2$$

$$2(x_r - x_u)x_s = \|x_r\|^2 - \|x_u\|^2 + d_{us}^2 - d_{rs}^2$$

Por tanto, temos um sistema linear $Ax = b$

Analisando um Caso Particular

$$A = 2\begin{bmatrix} x_{v1} - x_{u1} & x_{v2} - x_{u2} \\ x_{r1} - x_{u1} & x_{r2} - x_{u2} \end{bmatrix}$$

$$b = \begin{bmatrix} \|x_v\|^2 - \|x_u\|^2 + d_{us}^2 - d_{vs}^2 \\ \|x_r\|^2 - \|x_u\|^2 + d_{us}^2 - d_{rs}^2 \end{bmatrix}$$

$$x = \begin{bmatrix} x_{s1} \\ x_{s2} \end{bmatrix}$$

Perceba que se A for inversível, temos uma única solução $x^* = A^{-1}b$

Analisando um Caso Particular

Conclusão: Se o grafo G de um **PGDM for completo**, então podemos resolver o problema através da resolução de **N sistemas lineares**, sendo N proporcional ao numero de vértices.

Analisando um Caso Particular

Conclusão: Se o grafo G de um **PGDM for completo**, então podemos resolver o problema através da resolução de **N sistemas lineares**, sendo N proporcional ao numero de vértices.

Dizemos então que o problema *pode* ser resolvido em **tempo linear**

$O(n)$

Analisando um Caso Particular

Conclusão: Se o grafo G de um **PGDM for completo**, então podemos resolver o problema através da resolução de **N sistemas lineares**, sendo N proporcional ao numero de vértices.

Dizemos então que o problema *pode* ser resolvido em **tempo linear**

O(n)

Infelizmente isso não acontece. O grafo molecular não é completo.

Analisando o caso Geral

Para verificar a complexidade do PGDM nos resta olhar para a **otimização global** de

$$f(x_1, ..., x_n) = \sum_{(i,j)\in E} (\|x_i - x_j\| - d_{ij})^2$$

por tanto:

$$\min_{x_t \in \mathbb{R}^n} f(x_1, ..., x_n)$$

Analisando o caso Geral

Para verificar a complexidade do PGDM nos resta olhar para a **otimização global** de

$$f(x_1, ..., x_n) = \sum_{(i,j)\in E} (\|x_i - x_j\| - d_{ij})^2$$

por tanto:

$$\min_{x_t \in \mathbb{R}^n} f(x_1, ..., x_n)$$

Infelizmente **tal minimização é inviável**, a quantidade de mínimos locais cresce **exponencialmente** com n.

Analisando o caso Geral

Conclusão: O custo computacional para resolver um PGDM que, suponha, todas as distâncias da RMN são precisas e que de fato representam as distâncias entre os átomos pode ser proporcional a $2^{\|v\|}$.

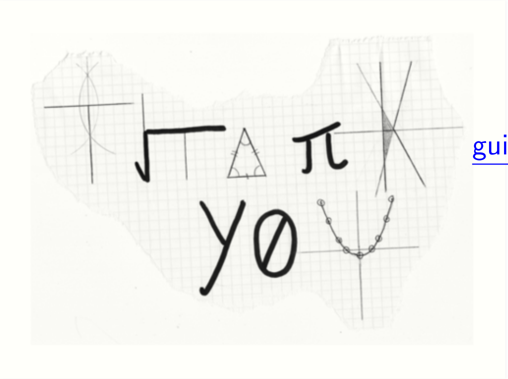
Analisando o caso Geral

Conclusão: O custo computacional para resolver um PGDM que, suponha, todas as distâncias da RMN são precisas e que de fato representam as distâncias entre os átomos pode ser proporcional a $2^{\|v\|}$.

Logo,

Teorema

O PGDM é um problema NP-Difícil.



guilherme.philippi@hotmail.com

UFSC - Blumenau