

# 大量の化学構造データの解析を高速化，効率化するアプリケーションの開発

○三瓶 匡史

E-Mail: masashi\_sampe@frontier.hokudai.ac.jp

**【序論】** 近年，理論化学の発達およびコンピュータの計算速度の向上により，計算化学は大きな発展を遂げている。特にGRRMプログラムによる反応経路自動探索により，膨大な数の化学構造データが得られる。しかし，それらを目視で分析するのは多大な労力を必要とする。そこで大量の化学構造データ（XYZ座標）の解析の高速化，効率化を目指してWindowsアプリケーション「CoordMatch」を開発した。このアプリケーションはXYZ座標が記載されたログファイルを用意すれば，マウス操作のみで容易に利用できる。本発表では，CoordMatchの機能を説明し，実演する。

**【動作環境】** Window7, Windows10にて確認。

**【機能】** CoordMatchは下記の流れで解析を効率化する。Fig. に具体例として，ベンゼンと塩化メチルを反応物，塩化アルミニウムを触媒としてトルエンを生成する反応の解析イメージを示す。

## 1. XYZ座標データの分類

XYZ座標中に存在する分子を結合様式から判断し，座標の分類を行う。分子の構造異性体や光学異性体の判断も可能である。光学異性体については，不斉炭素原子の検出とS体・R体の判別を行う。

## 2. データベースからの分子情報の取得

分類した座標に含まれる分子の化合物名などの情報を化学分子データベースから取得する。また，データベースに存在しない分子を新たに登録することができる。

## 3. 目視での確認

構成分子が明らかとなった座標グループの分子情報を利用者が目視で確認し，反応物，生成物，中間体などを判断する。

**【性能】** 1つの座標につき34原子を持つXYZ座標306個を分類するのに必要な時間は1秒程度であった。各座標が持つ原子数が多いほど，座標の数が多いほど分類に時間を要する。また，不斉炭素原子を持つ場合，その判別にも多少の時間を要する。

**【展望】** 現在実現できていない下記の機能を今後実装する予定である。

1. XYZ座標だけでなく，内部座標への拡張。
2. MacやLinuxへの対応。
3. 結合角や二面体角を用いた座標の分類。
4. 分子モデリングソフトとの連携。

## XYZ座標データ

座標 1			
C	4.044261880	-0.639383136	0.522232281
C	4.625545514	-1.291139001	-0.573545330
C	3.433836008	-1.386535725	1.537590944
H	5.101431731	-0.712078429	-1.358917029
H	2.986009771	-0.882140085	2.386517862
(略)			
座標 2			
C	4.044261880	-0.639383136	0.522232281
C	4.625545514	-1.291139001	-0.573545330
C	3.433836008	-1.386535725	1.537590944
H	5.101431731	-0.712078429	-1.358917029
H	2.986009771	-0.882140085	2.386517862
(略)			
座標 3			
(略)			

## 1. XYZ座標データの分類

**座標グループA** = 座標 1, 座標 2, ...  
 $C_6H_6 + CH_3Cl + Al_2Cl_6$   
**座標グループB** = 座標 3, 座標 5, ...  
 $C_7H_9Al_2Cl_7$   
**座標グループC** = 座標 4, 座標 6, ...  
 $C_7H_8 + HCl + Al_2Cl_6$

## 2. 分子情報の取得

**座標グループA** = 座標 1, 座標 2, ...  
ベンゼン + 塩化メチル + 塩化アルミニウム  
**座標グループB** = 座標 3, 座標 5, ...  
 $C_7H_9Al_2Cl_7$  (情報なし)  
**座標グループC** = 座標 4, 座標 6, ...  
トルエン + 塩化水素 + 塩化アルミニウム

## 3. 目視での確認

**座標グループA** = 反応物  
ベンゼン + 塩化メチル + 塩化アルミニウム  
**座標グループB** = 中間体  
 $C_7H_9Al_2Cl_7$  (情報なし)  
**座標グループC** = 生成物  
トルエン + 塩化水素 + 塩化アルミニウム

Fig. 解析の流れ