



Etude multi-objective de la conception des produits assistée par ordinateur : Formulation et applications aux molécules organiques.

Présentée par : GHOCHI Amira^a ; GHARBI Zineb^a

Encadrée par : SAIDAT Mustapha^a

^a Département de Chimie, Faculté des Mathématiques et des Sciences De La Matière, Université KASDI Merbah- Ouargla, BP 511 route de Ghardaia.30000 Ouargla, Algérie

Email : saidat.mustapha@gmail.com

Résumé:

Le problème de conception des molécules avec des propriétés désirées, est complexe. Ce dernier, est définie comme étant le chemin inverse de la prédiction des propriétés par les méthodes de contribution de groupes. Elle consiste à fixer dès le départ un ensemble de valeurs de propriétés dites cibles, puis de chercher par une combinaison de groupes moléculaires, des molécules ayant ces propriétés cibles. Dans le présent travail, une nouvelle technique de conception des molécules allant dans la première phase de l'analyse combinatoire de problème, qui suppose la disponibilité de la collection des groupes fonctionnels ainsi que les intervalles des valeurs des propriétés désirées. Les propriétés désirées constituent des contraintes sur les valeurs entières correspondant aux groupes fonctionnels. La région admissible définie par chaque propriété, est déterminé par un algorithme implique une stratégie de branchement. Cet algorithme, génère la collection des groupes fonctionnels qui constituent les molécules ou les composés admissibles satisfaisant les propriétés désirées. Dans la deuxième phase, et afin de réduire le nombre des molécules générés par la technique d'analyse combinatoire, nous avons utilisé la technique « PROPERTY CLUSTERING ». Cette dernière sert à représenter le problème à haute dimensions et le visualiser en deux ou trois dimensions. Pour le problème de conception, les valeurs des propriétés désirées sont représentés dans un diagramme ternaire afin de former une région admissible permet d'isoler les molécules performantes résultantes de la technique d'analyse combinatoire et donc choisir la bonne molécule correspondant aux propriétés désirées.

1- Introduction

La méthode traditionnelle d'essais et d'erreur, pour la conception de plusieurs molécules puis de les tester en laboratoire, est fastidieuse et souvent complexe. L'efficacité d'une telle approche est déterminée par la disponibilité des données et l'absence de biais en faveur à des solutions spécifiques. Le problème de conception moléculaire (CAMD) impliquent l'identification et la sélection de composés capables d'effectuer certains tâches ou possèdent certaines propriétés physiques ou fonctionnelles, comme la température d'ébullition, la chaleur de vaporisation, la solubilité, etc. Elle consiste à se fixer dès le départ un ensemble de valeurs de propriétés dites cibles, puis de chercher/trouver par une combinaison de groupes moléculaires, des molécules ayant ces propriétés cibles (Constantinou et al., 1996; Harper and Gani, 2000). Cette approche à fait l'objet de tentatives prometteuses dans plusieurs domaines de chimie de génie des procédés, à savoir le design des solvants (Gani & Brignole, 1983; Sinha & Achenie, 1998; Cismondi & Brignole, 2004; Chen et al., 2005), remplacements des réfrigérants (Churi & Achenie, 1997; Vaidyaraman & Maranas, 1999; Constantinou et al., 1996), design des polymères (Maranas, 1996; Maranas, 1997), conception des produits pharmaceutiques (Sachin Siddhaye et al., 2000; Siddhaye et al., 2004). L'objectif du présent travail est de modéliser le problème de conception des produits assistée par ordinateurs, par une technique multi-échelle allant de la représentation de la structure chimiques jusqu'à la conception de la molécule. Cette approche passe par plusieurs étapes à savoir l'estimation/classification des propriétés (objectives et subjectives), l'analyse combinatoire, la représentation des molécules, l'utilisation du concept Property Clustering et une méthode d'optimisation de la recherche des molécules.

3. Le cadre d'intégration de propriété :

L'introduction du cadre d'intégration des propriétés par Shelley et El-Halwagi (2000) permet de représenter les processus et les produits du point de vue des propriétés.

Fournir une méthodologie pour gérer la synthèse moléculaire / problèmes d'intelligence, l'infrastructure d'intégration de propriété est étendue aux méthodes de contribution de groupe (GCM), qui permette la prédiction de propriétés physiques à partir d'informations structurales. En combinant des techniques de regroupement de propriétés et des méthodes de contribution de groupe de premier ordre éprouvées, on obtient une méthodologie systématique qui fournit une plate-forme basée sur les propriétés facilitant la visualisation des exigences de performances des propriétés et la conception de nouvelles formulations.

3-1 prérequis :

Il est clair que l'opérateur les expressions seront invariablement des fragments moléculaires et des flux de processus différents, mais comme ils représentent la même propriété, il est possible de les visualiser de la même manière. En étendant cette technique pour inclure le GCM pour la conception moléculaire, introduit les opérateurs de propriété moléculaires. À l'instant, des opérateurs d'origine, leur formulation doit être telle qu'elle permette encore de simples règles d'addition linéaires.

3.2. Visualisation du problème

La conversion des données de propriété en valeurs de cluster dans la conception de processus a été développée par Eden et al. (2004). De même, la conversion des données de propriétés moléculaires en valeurs de cluster suit la procédure du tableau 2. L'outil de visualisation principal du cadre d'intégration de masse est la cartographie source – puits (El-Halwagi, 1997). Cet outil est également utilisé dans le cadre de la synthèse moléculaire. Dans la formulation de groupe initiale pour la conception de processus, le mélange de deux sources est une ligne droite, c'est-à-dire que l'opération de mélange peut être optimisée à l'aide de analyse par bras de levier. De manière analogue, en combinant ou en «mélangeant» deux molécules les fragments dans le domaine du cluster moléculaire suivent une ligne droite (un exemple illustratif est donné sur la figure 2 ci-dessous). Des règles de conception et d'optimisation ont été élaborées pour les problèmes de conception de processus basés sur les propriétés.

| step | Description |
|------|---|
| 1 | Calculer les opérateurs de propriétés moléculaires |
| 2 | Calculer les valeurs des propriétés moléculaires sans dimension |
| 3 | Calculer les indices de AUP moléculaires |
| 4 | Calculer les valeurs de cluster ternaires pour chaque formulation |

Tableau 2 Calcul des valeurs de clusterM à partir des données de propriétés prévues de GCM

3.3. Limites de la région de faisabilité

Dans le cadre du cluster de propriétés, les contraintes imposées à la fois les problèmes de processus et de conception moléculaire sont représentés comme une région de faisabilité sur le diagramme.

Maintenant que nous avons décrits en termes de grappes, il existe un cadre unificateur pour la résolution simultanée des problèmes de conception liés aux propriétés. En outre, la technique de regroupement réduit la dimensionnalité des deux problèmes. Il est donc possible d'identifier visuellement les solutions, ce qui constitue un avantage important de cette approche.

Conclusion

Une approche combinatoire pour générer molécules candidates avec des propriétés désirées basées sur la contribution du groupe

Le problème de conception des produits assistée par ordinateur est capable de composer les structures moléculaires de composés chimiques de manière optimale par la contribution du groupe. L'approche proposée sera un outil utile pour traiter rapidement sélectionner des composés ayant les propriétés souhaitées et pour identifier des produits chimiques à la lumière de chromatographie Les données. Il a été démontré que cette approche peut faciliter la première étape de l'approche quantitative de la conception de constituants de l'arôme sur la base de la structure-odeur relation. Une approche visuelle systématique de la conception moléculaire via la propriété groupes et méthodes de contribution de groupe résultant significatif de la méthodologie développée est que pour problèmes qui peuvent être décrits de manière satisfaisante par seulement trois propriétés, le problème de conception moléculaire est résolu visuellement sur une base ternaire diagramme, quel que soit le nombre de fragments sur les molécules dans l'espace de recherche. L'outil visuel donne au concepteur un guide sur les groupes à inclure dans la synthèse et lesquels ceux qui ne contribueront pas à satisfaire les exigences de performance cibles. Cependant, pour les cas nécessitant plus de trois propriétés, une approche de regroupement moléculaire algébrique a été développée qui permet d'abaisser la dimensionnalité de la conception problème

4. Synthèse moléculaire

Le problème est visualisé en convertissant les cibles de propriété en valeurs de cluster en suivant la méthodologie décrite précédemment dans le tableau 2.

Les contraintes de propriétés des molécules conçues sont représentées sous forme de région de faisabilité, qui est identifiée conformément aux règles de faisabilité décrites précédemment dans la section 3.3. Les valeurs AUP des six points qui délimitent la région de faisabilité définissent la plage AUP du puits, qui sera utilisée pour la synthèse post-moléculaire afin de satisfaire les conditions de faisabilité nécessaires. Le diagramme ternaire résultant est présenté à la figure 2, où les lignes pointillées représentent la région de faisabilité pour la conception du solvant.

Notez que même si certains des opérateurs de propriété formulés Auparavant, très complexes, la synthèse moléculaire sur le diagramme ternaire est encore simple car ces opérateurs sont obligés d'obéir à de simples règles additives linéaires. Il convient de noter que l'emplacement des molécules formulées est indépendant de la voie d'addition de groupe. L'emplacement de la formulation est unique et repose uniquement sur le numéro de chaque groupe dans la molécule. Par exemple, considérons le butyl méthyl éther (C₅H₁₂O), il est composé des groupes suivants: CH₃, CH₂ et CH₃O. Construire cette molécule sur le cluster ternaire diagramme, en utilisant les mêmes propriétés choisies plus tôt, peut être fait de différentes façons. Cependant, quelle que soit la séquence dans laquelle les groupes sont combinés, la molécule résultante (CH₃O-CH₂-CH₂-CH₂-CH₃) est située au même point unique. Pour vérifier que la règle 2 est satisfaite, comme indiqué à la section 3.2, le numéro de liaison libre (FBN) de chaque fragment moléculaire est placé entre parenthèses sur le diagramme ternaire.

En regardant la Fig. 3a, le point de départ est CH₃O puis ajout de trois molécules de CH₂ puis de CH₃. La figure 3b commence avec CH₃, puis ajoute trois molécules de CH₂ et enfin CH₃O. Les deux chemins indiqués, ainsi que de nombreux autres, aboutissent au même point, d'où le lieu de la formulation moléculaire est unique et indépendante de l'addition de groupe chemin.

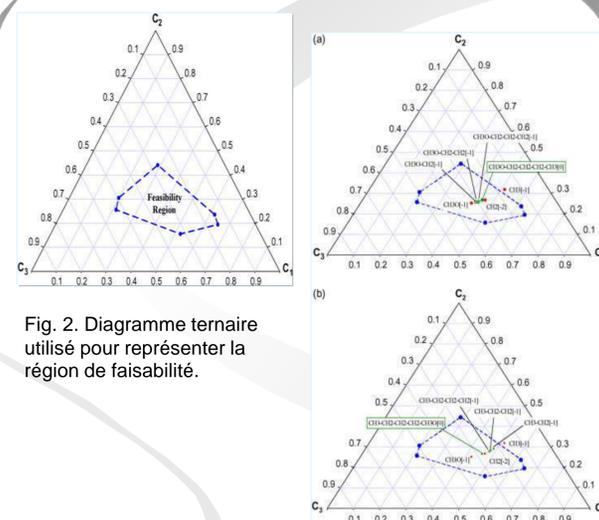


Fig. 2. Diagramme ternaire utilisé pour représenter la région de faisabilité.

Fig. 3. (a) Chemin A pour concevoir le butyl méthyl éther. (b) voie B pour concevoir le butylméthyléther.

2- Illustration :

Supposons que les quatre groupes fonctionnels >CH-, --CH₃, -F, and --CH₂- soient disponibles pour composer des molécules structures satisfaisant les contraintes sur l'ébullition (Tb) et points de fusion (Tm)

$$311 \text{ (K)} < T_b < 313 \text{ (K)}$$

$$116 \text{ (K)} < T_m < 119 \text{ (K)}$$

Les points d'ébullition normaux, T_b, et les points de fusion normaux, T_m, des composés sont calculés par la méthode de contribution de groupe de Joback respectivement, comme suit:

$$T_b(K) = 198.18 + \sum n_i \Delta T_{b,i}$$

$$T_m(K) = 122.5 + \sum n_i \Delta T_{m,i}$$

Le tableau 1 énumère les contributions du groupe; il est supposé que chaque groupe peut apparaître au plus deux fois. L'arbre de dénombrement est donné à la Fig. 1. La racine de l'arbre représente l'ensemble des contraintes données par les inégalités (6) à (10).

Maintenant, en étendant les valeurs de x₁, x₃ et x₄ des nombres entiers aux nombres réels, seul (P3) s'avère réalisable. Par conséquent, il est ramifié plus loin en considérant x₂ comme 0, 1 ou 2. La procédure est poursuivie jusqu'à ce que la recherche englobe tous les x_i. Chaque étape de poursuite est indiquée dans l'illustration de l'énumération complet présentée à la Fig. 1. Notez que ce problème simple a une seule partition possible avec deux groupes fonctionnels >CH-, deux groupes -CH₃, deux -F et un -CH₂-. Ces groupes donnent lieu à diverses structures moléculaires

| Groups | ΔT _b +(k) | ΔT _m ++(k) |
|---------------------|----------------------|-----------------------|
| >CH- | 23.58 | - 5.10 |
| --CH ₃ | 22.88 | 11.27 |
| -F | 21.74 | 12.64 |
| --CH ₂ - | - 0.03 | - 15.78 |

Tableau 1. Contributions de groupe pour T_b et T_m des groupes fonctionnels étudiés dans l'illustration.

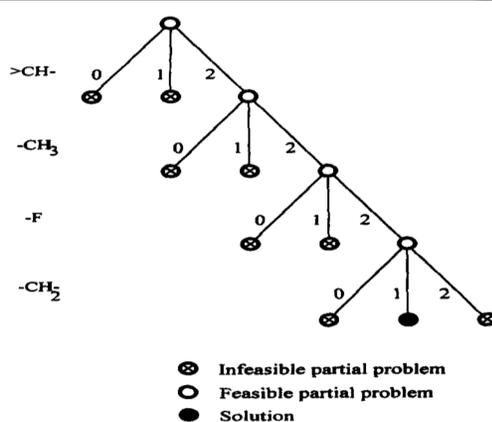


Fig. 1. Enumeration tree for the illustration

Références bibliographiques

- Barber K., J. Gmehling, U. Onken, 1979, Ber. Bunsenges Phys. Chem., 83, 1133–1136. Auswahl von Lösungsmitteln für die Extraktiv-Rektifikation mittels vorab berechneter Gleichgewichtsdaten
 Briggs, B. G., 1981, J. Agr. Food Chem., 29, 1050–1059
 Theoretical and Experimental Relationship Between Soil Adsorption, Octanol-water Partition Coefficients, Water Solubilities, Bioconcentration Factors and the Parachor
 Ador, A. F. (1985). *Handbook of Solubility Parameters and other Cohesion Parameters*. Boca Raton, FL: CRC Press.
 Brignole, E., & Cismondi, M. (2003). L. E. K. Achenie, R. Gani, & V. Venkatasubramanian (Eds.), *Molecular design—Generation and test methods. Computer aided molecular design: Theory and practice* (pp. 23–41). Elsevier.