

Biokémiai reakcióutak szintézise és analóg feladatok megoldása

Rövid összefoglalás

A kutatás témája - a reakcióutak szintézise - a kémiai és biokémiai reakciók mechanizmusának megértését segíti. A feladat megoldása, komplexitása miatt, hatékony informatikai támogatás nélkül elképzelhetetlen. Mindemellett, több olyan mérnöki feladat is található, mely matematikai szempontból analóg feladatra vezet.

Az első év során a feladat szakirodalomban rendelkezésre álló modelljeit és megoldó módszereit gyűjtöttük és hasonlítottuk össze. A kritikus futási idő miatt vizsgáltuk a módszerek implementációs kérdéseit is.

A második évben azt vizsgáltuk, hogy egyrészt a különböző módszerekben használt modellek és perprezentációk által hordozott információ hogyan segít a reakcióutak meghatározását. Másrészt, lehetséges-e többféle reprezentációban megjelenő információt párhuzamosan hasznosítani. Harmadrészt, sok megengedett reakcióút esetén hogyan lehet a dominánsakat kiválasztani.

A harmadik évben implementáltuk a különböző matematikai modellek összefüggéseit is kihasználó általunk kidolgozott integrált megoldó algoritmusokat. Keretrendszer fejlesztettünk a megoldások elemzésére különböző reprezentációk szerint. Kidolgoztunk egy olyan technikát, mely biztosítja, hogy kezelhetetlenül sok megoldás esetén csak a leginkább valószínűket generáljuk, de nem a nagyszámú megoldás szűrésével, hanem úgy, hogy a generált megoldások számával arányban a futási idő is csökkenjen. Mindemellett kidolgoztunk egy analóg feladat megoldásához szükséges adaptációt.

Az egyszemélyes ifjúsági OTKA pályázatban 3 évre összesen 1 550 000 forint támogatást igényeltem egy számítógép vásárlására és konferencia részvételek támogatására. A projekt során 18 publikáció született, köztük 3 magas impakt faktorú tudományos folyóiratban, 9 előadás rangos nemzetközi konferenciákon, 1 magyar nyelvű tudományos konferencia előadás. Továbbá 4 diplomadolgozat és 1 OTDK dolgozat készült a témában témavezetésemmel.

Az elért eredmények részletes ismertetése

A reakcióút azonosításra ismert szisztematikus módszerek gráfelméleti, lineáris algebrai vagy konvex elemzésre épülnek. A lineáris algebrai és a konvex elemző módszerek matematikai modellje és megoldása nagyon hasonló, különbségük az, hogy a konvex elemzés a reakcióutak elemi lépéseinek nem megfordíthatóságát is képes figyelembe venni. A gráfelméleti és lineáris algebrai megközelítésre épülő matematikai modelleket és a feladatmegoldás lépéseit elemeztük, kapcsolatukat feltártuk [1, 3].

Vizsgáltuk a megoldó módszerek alkalmazhatóságát és eredményességét gyakorlati feladatokon [2-5, 8]. A különböző megoldó módszerek által szolgáltatott eredmények közti különbségeket fontos gyakorlati példán mutattuk meg [2, 8]. Felhívtuk a szakmai közösség figyelmét a nem megfelelően választott modell alkalmazásának veszélyeire [8].

A módszerek összevetése során megfogalmaztuk a különböző publikációkban más-más szóhasználattal leírt, de ekvivalens feladatok halmazát [1]. A feladat exponenciális jellege miatt a megoldó algoritmusok implementációjának hatékonysága sem elhanyagolható. A matematikai modellek elemzése mellett azok szoftverimplementációban megjelenő

lehetséges informatikai modelljeit (adatszerkezetek, objektum-orientált modellek) is elemeztük [7]. Generikus (sablon alapú) programozás felhasználásával javaslatot tettünk a reakcióút szintézis feladat olyan megfogalmazására C++ nyelven, mely a strukturális vizsgálatok során használt leképezések szemantikai ellenőrzését fordítási időben teszi lehetővé, így segítve a megoldó szoftverek helyességének ellenőrzését. Az informatikai modelleket nyílt forrású szoftverek részeként kívánunk publikálni a közeljövőben. Megmutattuk, hogy a lehetséges megoldások kereséséhez használt gráfelméleti módszer eredményessége a keresési tér szűkítésében tovább javítható a reakcióutak elindíthatóságának elemzésére történő kiterjesztésével [6].

A reakcióút generálás kimerítő leszámolására épülő algoritmikus módszerek elvárt kimenetét sokan sok szóhasználattal írták le a szakirodalomban. Továbbá, ha feladatot matematikai modell szintjén írjuk le, akkor számos analóg feladatot és megoldó módszert is találhatunk. Sajnos az irodalomban találtunk példát arra is, hogy a sokféle megfogalmazás a feladat megértését és megoldását is félreviheti. Erre vonatkozó észrevételünket ugyanabban a folyóiratban publikáltuk [8]. Ebből kiindulva két dolgot tettünk. Egyrészt szinonima szótárt készítettünk, hogy a különböző iskolák képviselői megértsék egymás [10], másrészt matematikai modell szintjén pontosítottuk a megoldandó feladatot [9].

A megfogalmazott matematikai modell korábbi megoldásai során használt kétféle alapvetően különböző reprezentáció (páros gráf, lineáris egyenletrendszer) együttes alkalmazásával új algoritmust és szoftvert hoztunk létre [11,13]. Az algoritmus kidolgozása több lépésben történt, mert vizsgálataink során a modellekben egyre több olyan tulajdonságot sikerült feltárnunk, melyek együttes figyelembevétele jelentősen támogatja a feladat megoldásának eredményességét és számítási idejének csökkentését. Az szoftver fejlesztése közben számos implementációs kérdést is tárgyaltunk [12]. Ezen kérdések megválaszolása nagyban hozzájárul ahhoz, hogy az integrált megoldók kidolgozása során nemcsak hatékony, de mások által könnyen továbbfejleszhető eszközöket adhassunk közre.

Korábbi munkánk során már bizonyítottuk, hogy heurisztikus eljárások általában nem adnak teljes megoldáshalmazt. Kimenetük kisebb, könnyebben kezelhető, de nagy a kockázata, hogy fontos megoldásokat is elveszítenek. Természetes azonban az igény, hogy a kimerítő szisztematikus módszerekkel generált nagyszámú reakcióút halmazt algoritmikusan szűkítsük. Egy reakció mechanizmusában általában több reakcióút szerepel egy időben. Gyakorlatban fontos, hogy mégis melyek a dominánsak. A domináns reakcióutak meghatározására két megközelítéssel éltünk. Egyik esetben azt a kumulált aktiválási energiát néztük, ami ahhoz szükséges, hogy a reakcióút minden lépése végbemenjen [15]. Ahol ez az energia legkisebb, ott a legnagyobb annak esélye, hogy a reakcióút végbemegy. Másik esetben termodinamikai számítások alapján pontosítottuk, hogy adott körülmények között (például hőmérséklet) melyek lehetnek a domináns reakcióutak [17].

Az energetikailag legkedvezőbb reakcióutak generálását kezdetben a teljes megoldáshalmaz szűkítésével végeztük, majd meghatároztunk egy korlátozás és szétválasztás technikára alapuló eljárást, mely közvetlenül csak a legkedvezőbb reakcióutakat eredményezi teljes aktiválási energiájuk növekvő sorrendjében [14]. Gyakorlati tapasztalatok szerint a futási idő arányosan csökken a generálni kívánt reakcióutak számával, így olyan nagy méretű feladatok energetikailag legkedvezőbb reakcióútjai is generálhatóvá váltak, melyekhez a teljes megoldáshalmaz előállítás a gyakorlatban elfogadhatatlanul időigényes lenne.

A különböző matematikai modellek nem csak a reakcióút generálás folyamatát támogatják, de az eredményül kapott reakcióutak megértését is. Ennek fényében megvalósítottunk egy olyan keretrendszert, ahol a többféle modellre épülő algoritmusok eredményei többféle reprezentációban is megjeleníthetőek [16, 18].

A kész módszert és szoftverrendszert elkezdjük adaptálni analóg feladatok megoldására. Ennek első eredményeként sikerrel oldottunk meg segítségével reakció hálózat

szintézis feladatokat. A reakcióhálózat elemei a reakcióutakkal ellentétben több külön közegben zajlanak, ezért sokkal nagyobb a beavatkozás lehetősége. Ennek következtében, a reakcióút azonosításhoz mérten a reakcióhálózat szintézis egy alulhatározott feladat megoldását követeli meg. Ez az alulhatározottság némely matematikai modellben nem leírható, de keretrendszerben felhalmozott eszköztárral sikeresen kezelhető. Az adaptáció eredményeként született reakcióút szintézis módszert a közeljövőben publikáljuk. Kísérletekkel bizonyítottuk továbbá, hogy egy másik alkalmazási terület lehet a szétválasztási hálózatok tervezéséhez alternatív struktúrák generálása.

A támogatott kutatás eredményeként előállt algoritmusok átfogóan támogatják a reakcióút szintézist és analóg feladatok megoldását. Eszközt adnak a reakcióutak teljes halmazának generálására többféle matematikai modell együttes használatával, vagy nagyméretű feladatok esetén az energetikailag leginkább valószínű kis számú reakcióút gyors meghatározására. Az eredmények megjeleníthetők többféle (például gráfos és mátrixos) reprezentációban. A szoftverek kizárólag nyílt forrásban terjeszthető komponensekre épülnek, így a megfelelő licencekkel ellátva szabadon terjeszthetők, elősegítve azok további adaptációit hasonló modellre vezető feladatok megoldására.

Kapcsolódó publikációk

- [1] Bertok B; Kalauz K; Fan LT: ***Exhaustively Generating Basis Solutions for a System of Stoichiometrically Balanced Elementary Reactions by Resorting to the P-graph-based Method***, Veszprém Optimization Conference: Advanced Algorithms (VOCAL), Veszprem, Hungary, December 13-15, 2006
- [2] Fan LT; Lin YC; Shafie S; Hohn KL; Bertók B; Friedler F: ***Microkinetic Model for the Water-gas Shift Reaction on Supported Copper Catalyst***, Proceedings of the AIChE 2006 Annual Meeting, San Francisco, CA, November 12 – 17, 2006
- [3] Kalauz K; Bertok B; Friedler F; Fan LT: ***Reaction pathway identification: relations of methods based on graph theory and linear algebra***, CHISA 2006, Praha, Czech Republic, August 27-31, 2006
- [4] Lin YC; Fan LT; Shafie S; Hohn KL; Bertók B; Friedler F: ***Exhaustive Identification of Stoichiometrically Feasible Pathways of Partial Oxidation of Methanol on Copper-Zinc Catalysts***, Proceedings of the AIChE 2006 Annual Meeting, San Francisco, CA, November 12 – 17, 2006
- [5] Nagy Z; Bertok B; Friedler F; Lee DY; Fan LT; Shafie S: ***Exploring Mechanisms Leading to Robust Biochemical Systems***, PRES 2006, Praha, Czech Republic, August 27-31, 2006
- [6] Nagy Zs (Témavezető: Bertók B): ***Reakcióutak strukturális elindíthatóságának algoritmikus vizsgálata***, diplomadolgozat, Pannon Egyetem, 2006
- [7] Nyitrai Sz (Témavezető: Bertók B): ***Szoftverfejlesztés UML alapú CASE eszköz segítségével***, diplomadolgozat, Pannon Egyetem, 2006
- [8] Fan LT; Lin YC; Shafie S; Hohn KL; Bertók B; Friedler F: ***Comment on: An improved microkinetic model for the water-gas shift reaction on copper [Surf. Sci. 541 (2003) 21-30]***, Surface Science, 601, 2401-2405, 2007
- [9] Bertok B; Kalauz K; Friedler F, Fan LT: ***Sztöchiometriai egyenletrendszerek minimális számú aktív változót tartalmazó megoldásainak generálása a P-gráf módszertan alkalmazásával***, XXVII. Magyar Operációkutatási Konferencia, Balatonöszöd, Június 7-9, 2007

- [10] Bertok B; Kalauz K; Friedler F, Fan LT: ***Systematic Methods for Reaction-Pathway Identification***, XIX Polish Conference of Chemical and Process Engineering, Rzeszow, Poland, September 3-7, 2007
- [11] Kalauz K (Témavezető: Bertók B): ***Reakcióút azonosítás: Gráf elméleti és lineáris algebrai módszerek integrálása***, OTDK dolgozat, Pannon Egyetem, 2007
- [12] Kalauz K (Témavezető: Bertók B): ***Reakcióút szintézis algoritmus és szoftver kidolgozása***, diplomadolgozat, Pannon Egyetem, 2007
- [13] Bertok B; Kalauz K; Friedler F; Fan LT: ***Reaction-Pathway Identification: Models and Algorithms***, VOCAL 2008 (Veszpém Optimization Conference: Advanced Algorithms), Veszprém, Hungary, December 15-17, 2008
- [14] Bertok B; Kalauz K; Friedler F; Lin YC; Fan LT; S. Shafie: ***Combinatorially accelerated branch-and-bound method for identifying potentially dominant catalytic pathways***, PRES 2008 (11th Conference on Process Integration, Modelling and Optimisation for Energy Saving and Pollution Reduction), Praha, Czech Republic, August 24-28, 2008
- [15] Fan LT; Lin YC; Shafie S, Hohn KL, Bertok B; Friedler F: ***Graph-theoretic and energetic exploration of catalytic pathways of the water-gas shift reaction***, Chin. Inst. Chem. Eng., 39, 467-473, 2008
- [16] Kalauz K; Bertok B; Fan LT; Friedler F: ***Reaction-pathway identification: methods and software***, PRES 2008 (11th Conference on Process Integration, Modelling and Optimisation for Energy Saving and Pollution Reduction), Praha, Czech Republic, August 24-28, 2008
- [17] Lin YC; Fan LT; Shafie S, Hohn KL, Bertok B; Friedler F: ***Catalytic Pathways Identification for Partial Oxidation of Methanol on Copper-Zinc Catalysts: $CH_3OH + 1/2 O_2 \leftrightarrow CO_2 + 2 H_2$*** , Ind. Eng. Chem. Res., 47, 2523-2527, 2008
- [18] Szlama A (Témavezető: Bertók B): ***Keretrendszer reakcióút-azonosításban használt szoftverek integrált alkalmazására és összehasonlító elemzésére***, diplomadolgozat, Pannon Egyetem, 2008