

**CONFERINȚA ANUALĂ DE MATEMATICĂ
APLICATĂ
ȘI INFORMATICĂ**

GRAFURI MOLECULARE

2014

Prof.

**Rășinar Rodica – Liceul Tehnologic „Gheorghe Miron
Costin” Constanța**

Grasu Antoaneta – Liceul Toretic Negru Vodă

TEORIA GRAFURILOR

Aplicată în studiul structurilor moleculare se constituie ca ramură interdisciplinară numită **topologie moleculară**.

- Caracterizarea topologică a structurilor chimice permite ordonarea acestora după criterii de simetrie și similaritate.
- Din studii comparate, asupra unui set de molecule, topologia moleculară descifrează factorii structurali implicați în relația structură-proprietate/activitate biologică. Partiționarea unor proprietăți moleculare și recompunerea lor, prin modele aditive, cu ajutorul descriptorilor moleculari și a analizei statistice, ca și modelarea de noi structuri, cu proprietăți dorite, reprezintă unul din obiectivele topologiei moleculare. Înainte de detalierea problemelor specifice topologiei moleculare vor fi prezentate câteva definiții de bază ale teoriei grafurilor.

Grafuri neorientate

Definiție. Fie X o mulțime finită și nevidă și $U \subseteq \{\{x, y\} \mid x, y \in X\}$.
Numim graf (sau graf neorientat) perechea $G = (X, U)$.

Elemente din mulțimea X se numesc noduri sau vârfuri, iar elementele din mulțimea U poartă numele de muchii ale grafului.

Un graf se reprezintă grafic printr-o mulțime de puncte corespunzătoare vârfurilor grafului și o mulțime de segmente corespunzătoare muchiilor.

Dacă pentru un graf există o reprezentare în care să nu existe două segmente care să se intersecteze, atunci spunem că graful reprezentat este un graf planar.

Pornind de la definiție observăm că dacă X are n elemente, atunci U are cel mult $\binom{n}{2}$ elemente.

Dacă A este o mulțime, vom nota prin $|A|$ numărul de elemente ale mulțimii.

Definiție. Fie $G = (X, U)$ un graf și $x \in X$ un nod fixat. Prin gradul nodului x înțelegem numărul muchiilor incidente lui x și notăm acest număr cu $d(x)$. Dacă $d(x)=1$ spunem că x este nod terminal. Dacă $d(x)=0$ spunem că x este nod izolat.

Propoziția 1. Fie $G = (X, U)$ un graf în care $|U| = m$.

Propoziția 2. Pentru orice graf $G = (X, U)$, numărul vârfurilor de grad impar este par.

Definiție. Fie $G = (X, U)$ un graf. Numim graf parțial al lui G , graful $G' = (X, V)$, unde $V \subseteq U$.

Definiție. Fie $G = (X, U)$ un graf. Numim subgraf în G , graful G'' , $G'' = (Y, V)$, în care $Y \subseteq X$, iar $V = \{ \{x, y\} \in U \mid x, y \in Y \}$.

Definiție. Fie $G = (X, U)$ un graf cu n vârfuri ($|X| = n$). Spunem că G este un graf complet, dacă oricare ar fi $x, y \in X$, avem $\{x, y\} \in U$ (orice două vârfuri din G sunt conectate sau adiacente).

Graful complet cu n varfuri se notează prin K_n

Propoziția 3. Numărul muchiilor lui K_n este C_n^2 .

Definiție. Fie $G = (X, U)$ un graf. Dacă există X_1 și X_2 astfel încât $X_1 \cap X_2 = \emptyset$ și $X = X_1 \cup X_2$ (X admite o partiție din două blocuri X_1 și X_2) și orice muchie din G unește un nod din X_1 cu unul din X_2 (oricare ar fi $e = \{x, y\} \in U$, dacă $x \in X_1$, atunci $y \in X_2$) spunem că G este graf bipartit.

Într-un multigraf două puncte pot fi unite prin mai mult de o linie. Figura 1.1 arată cele trei tipuri de grafuri menționate mai înainte.

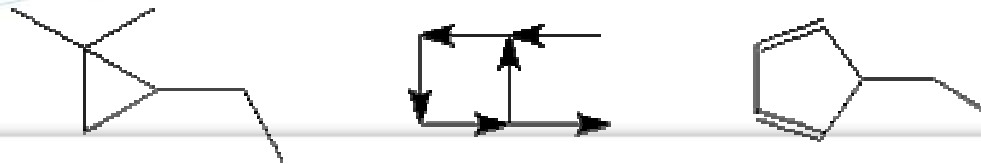


Figura 1.1

graf diagraf multigraf

O cale P (path) este un lanț (catenă) neramnificat(ă).

Un arbore T (tree), este o structură ramnificată.

O stea (star) este un set de vârfuri unite cu un vârf comun și se notează $S_{v'}$, cu $v' = v - 1$.

Un ciclu C este un lanț care pleacă și revine în unul și același vârf.

Figura 1.2



Cale

Arbora

Stea

Ciclu

Un graf complet, K_v , este un graf cu oricare două vârfuri adiacente. Numarul laturilor într-un graf complet este $v(v-1)/2$. In figura 1.3, sunt prezentate grafuri complete de la $v=1$ până la 5. Figura 1.3.




K_1 K_2 K_3 K_4 K_5

Un graf bipartit este graful al cărui set de vârfuri V poate fi împărțit în două subseturi: $V_1 \dot{\cup} V_2 = V$; $V_1 \cap V_2 = \emptyset$ astfel încât orice latură $(i,j) \in E(G)$ unește V_1 cu V_2 . Un graf este bipartit dacă și numai dacă toate ciclurile sale sunt pare. Dacă oricare vârf $i \in V_1$ este adiacent cu fiecare vârf $j \in V_2$, atunci G este un graf bipartit complet și este simbolizat $K_{m,n}$, cu $m = |V_1|$ și $n = |V_2|$. O stea este un graf bipartit complet $K_{1,n}$. Este evident că $K_{m,n}$ are $m \cdot n$ laturi. Figura 1.4 prezintă câteva grafuri bipartite.

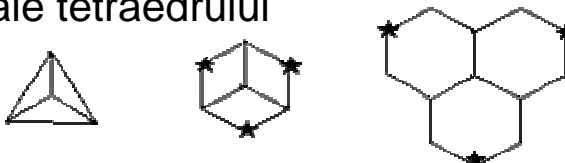


Figura 1.4. Graf bipartit $K_{1,3}$ $K_{2,3}$ $K_{3,3}$

Un graf marcat, este graful în care heteroatomul sau atomul de carbon cu un electron neîmperechiat este marcat.(Figura 1.5)

Figura 1.5. Grafuri marcate 

Un graf homeomorf este graful obținut prin inserarea unor vârfuri de valența 2 pe muchiile unui graf G. (Figura 1.6)

Figura 1.6. Grafuri homeomorfe ale tetraedrului 

Un subgraf al unui graf $G = (V, E)$ este un graf $H = (V_1, E_1)$, unde $V_1 \subseteq V$ și $E_1 \subseteq E$ reprezintă toate muchiile care au ambele extremități în V_1 . Altfel spus, un subgraf H al lui G se obține din G eliminând anumite vârfuri și păstrând doar acele muchii care au ambele extremități în mulțimea vârfurilor rămase. (Figura 1.7).

Figura 1.7. Un graf și un subgraf al său 

Un arbore de deschidere (spanning tree) este un subgraf conex $G_1=(V,E_1)$ conținând toate vârfurile lui G dar $E_1 \subsetneq E$ (Figura 1.8.).

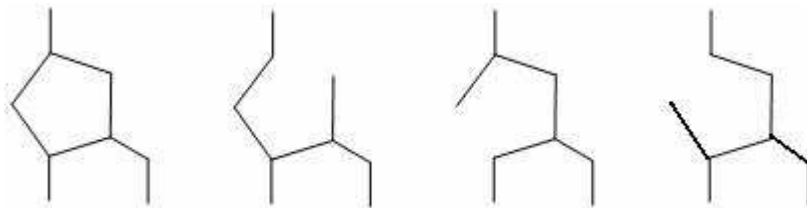
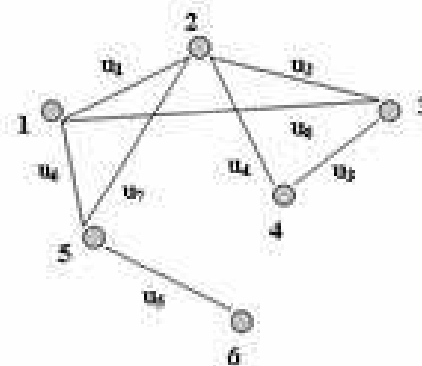


Figura 1.8. Un graf și câțiva dintre arborii lui de deschidere

Un graf G este conex dacă între oricare două vârfuri ale sale există o cale. (Figura 1.9).

Figura 1.9. Graf conex



Drumuri

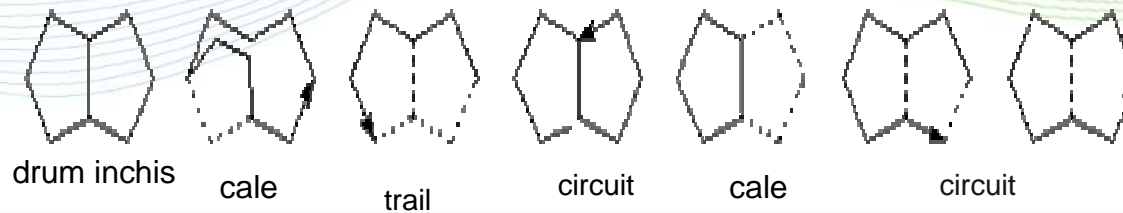


Figura 1.10 Diferite tipuri de drumuri

Distanța topologică $d(i,j)$: este lungimea căii celei mai scurte (dacă există) care unește vârfurile i și j : , altfel $d(i,j)=\text{∅}$. Distanța dintre două vârfuri adiacente este egală cu unitatea. Cea mai scurtă cale este numită și *geodezică*.

Excentricitatea unui vârf i , $ecc(i)$, într-un graf conex este distanța maximă dintre i și oricare vârf j al grafului: $ecc(i) = \max d(i,j)$.

Raza unui graf, $r(G)$, este excentricitatea minimă în G : $r(G) = \min ecc(i) = \min \max d(i,j)$. *Diametrul* unui graf, $d(G)$, este excentricitatea maximă în G : $d(G) = \max ecc(i) = \max \max d(i,j)$.

Mulțimea *tuturor distanțelor* (geodezicelor) în G se notează cu $D(G)$.

Detur, $d(i,j)$, reprezintă calea cea mai lungă (daca există) ce unește vârfurile i și j : , altfel $d(i,j)=\text{∅}$. Setul *tuturor detururilor* în G se notează cu $D(G)$.

- Detur, $d(i,j)$, reprezintă calea cea mai lungă (dacă există) ce unește vârfurile i și j : $d(i,j) = \max |P(i,j)_t|$, altfel $d(i,j) = \infty$. Setul tuturor deturilor în G se notează cu $D(G)$.

- Într-un graf conex, distanța și deturul reprezintă *metrici*, adică, pentru oricare vârfuri i, j, k :

$$m(i,j) \geq 0; m(i,j) = 0 \text{ dacă } i = j$$

$$m(i,j) = m(j,i)$$

$$m(i,j) + m(j,k) \geq m(i,k)$$

- Distanța metrică: este distanța, măsurată în metri sau submultipli ai acestuia (nm, Angstroms), dintre două vârfuri i și j . Tabelul 1.1 illustrează aceste noțiuni.

Graf	Distanța topologica	Distanța metrică (Å)
CH ₃ -CH ₃	1	1.54
CH ₂ =CH ₂	1	1.34
CH ₀ -CH	1	1.21

Valența sau gradul vârfului i , va_i : este numărul muchiilor incidente în vârful i . Dacă toate vârfurile au același grad, graful se numește regulat; altfel se numește neregulat (Figura 1.11).

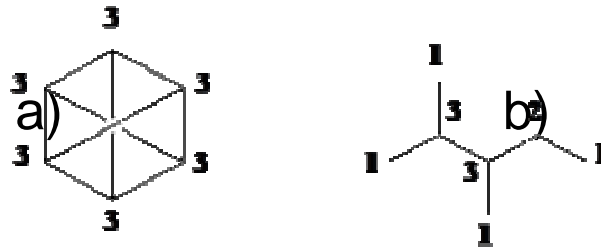


Figura 1.11. Grafuri ponderate cu valența vârfurilor: 3-regulat (a) și neregulat (b)

Invariant: este o proprietate a grafului teoretic, care se păstrează prin izomorfism. Cu alte cuvinte, valoarea numerică a proprietății rămâne neschimbată indiferent de numerotarea sau reprezentarea pictorială a grafului.

GRAFURI MOLECULARE

Un graf molecular este un model al unui sistem chimic, utilizat pentru caracterizarea interacțiunii componentelor sale: atomi, legături, grupuri de atomi sau molecule. Formula structurală a unei substanțe chimice poate fi reprezentată ca graf molecular, vârfurile lui fiind atomii, iar laturile corespunzând legăturilor covalente. În mod obișnuit, atomii de hidrogen se neglijează (Figura 1.12) și de asemenea unghiurile dintre legături. Atomii grei alții decât carbonul (heteroatomii) pot fi reprezentați ca în Figura 1.5. Similar, transformarea unei molecule (o reacție chimică) poate fi vizualizată printr-un graf de reacție, ale cărui vârfuri sunt specii chimice iar laturile căi de reacție.

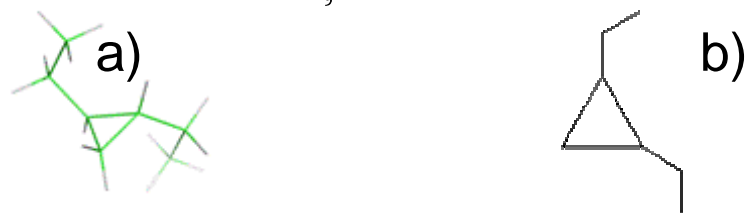
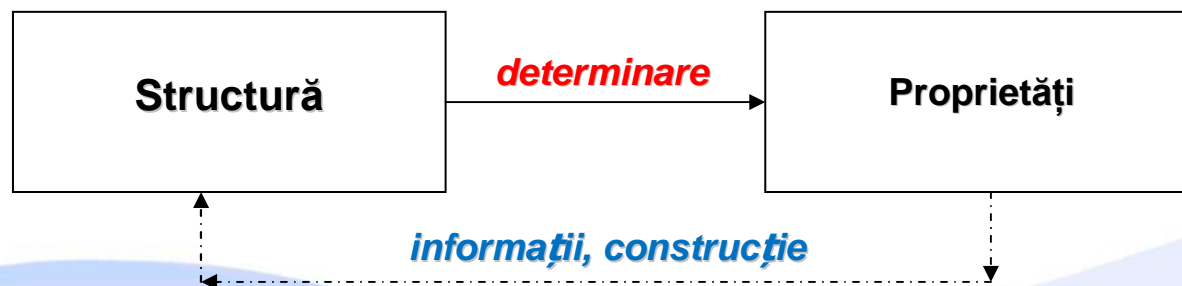


Figura. 1.16: O structură chimică (a) și graful molecular corespunzător acesteia (b).

APLICAȚIILE TEORIEI GRAFURILOR

Structura compușilor chimici este fundamentală pentru înțelegerea comportării lor. Între proprietățile compușilor chimici și structura lor există o relație foarte clară, în sensul că structura determină proprietățile, iar din evaluarea și interpretarea proprietăților se pot trage concluzii cu privire la anumite aspecte structurale.

Schematic, aceasta se poate reda astfel:



MODELARE

Proprietățile compușilor chimici sunt în general măsurabile (direct sau indirect) pot fi :

- proprietăți fizice prin valorile direct măsurate,
- proprietăți chimice prin viteza de reacție sau prin valori ale unor mărimi termodinamice, activitatea biologică prin expresiile $-\log C_i$).

Problema majoră care apare este cea legată de exprimarea structurii într-o formă numerică.

Dacă se reușește acest lucru, atunci corespondența structură – proprietăți de reduce (din punct de vedere matematic) la corespondența dintre două numere, ceea ce este mai ușor de realizat. Pentru a exprima numeric structura unui compus chimic (sau „cuantificarea structurii”) există mai multe modalități de lucru sau „modele”.

Un **model** reprezintă punerea în corespondență a unui sistem real cu un sistem material sau abstract, în general mai simplu, astfel încât urmărind comportarea celui mai simplu să putem trage concluzii asupra comportării originalului.

MODELE APLICATE ÎN CHIMIE

- **MODELUL TOPOLOGIC** în cadrul acestuia moleculele sunt considerate niște obiecte matematice numite **grafuri** și contează în principal relația de vecinătate între atomi (sau mai bine spus **starea legat – nelegat direct**). În cadrul acestui model, într-o primă aproximație nu contează nici natura atomilor constituenți, iar prezența atomilor de hidrogen se neglijează.
- **MODELUL MECANICII MOLECULARE** în cadrul acestuia moleculele sunt constituite din atomi care la rândul lor sunt sfere incompresibile (nu se face distincție între nucleu și electroni) legate între ele prin arcuri care sunt oscilatoare armonice. Totul se supune legilor fizicii clasice.
- **MODELUL MECANICII – CUANTICE** în cadrul acestuia moleculele sunt constituite din atomi care la rândul lor sunt sfere incompresibile (nu se face distincție între nucleu și electroni) legate între ele prin arcuri care sunt oscilatoare armonice. Totul se supune legilor fizicii clasice.

LFER În afara celor trei tipuri de modele care au fost prezentate și care se referă la întreaga structură moleculară în ansamblul ei, mai există și alte tipuri care iau în considerare modificările aduse unei structuri de bază prin substituenți și influența lor asupra proprietăților. Aceste modele se încadrează în domeniul relațiilor liniare de energie liberă

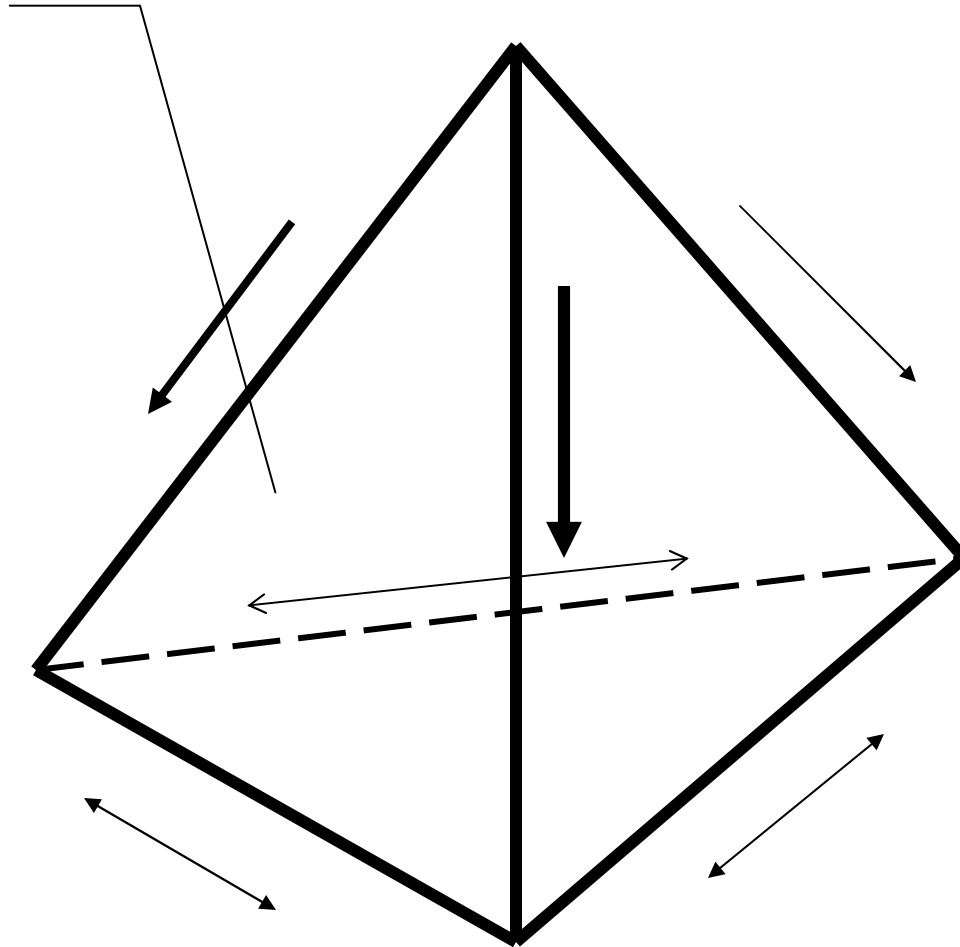
*Domenii de
interfață*

*Matematică +
calculatoare*

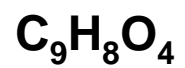
Chimia

Biologia

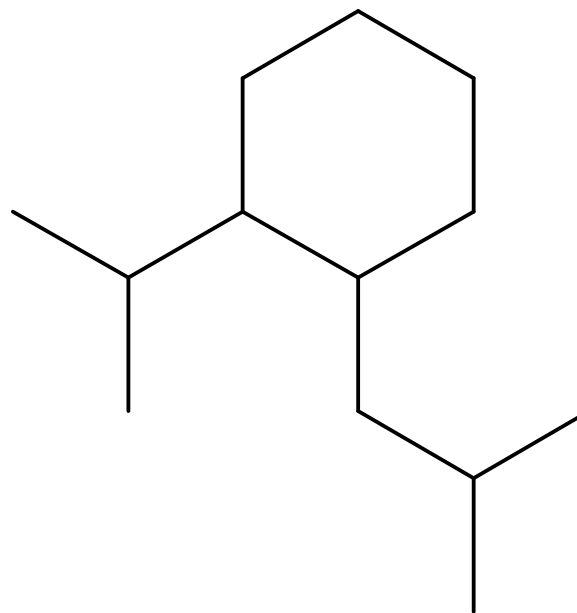
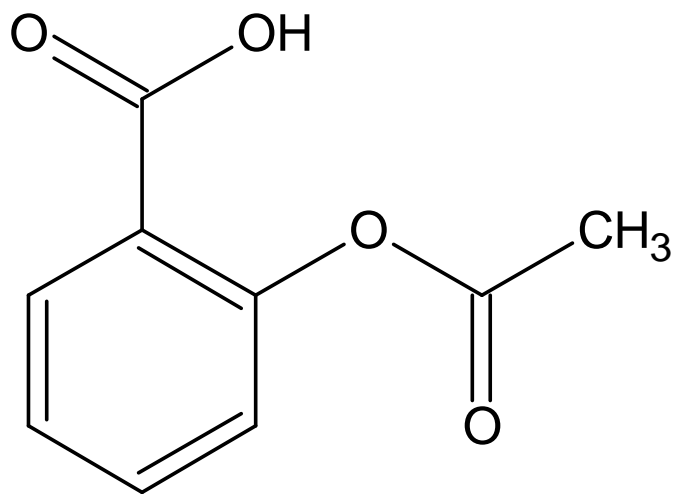
Fizica

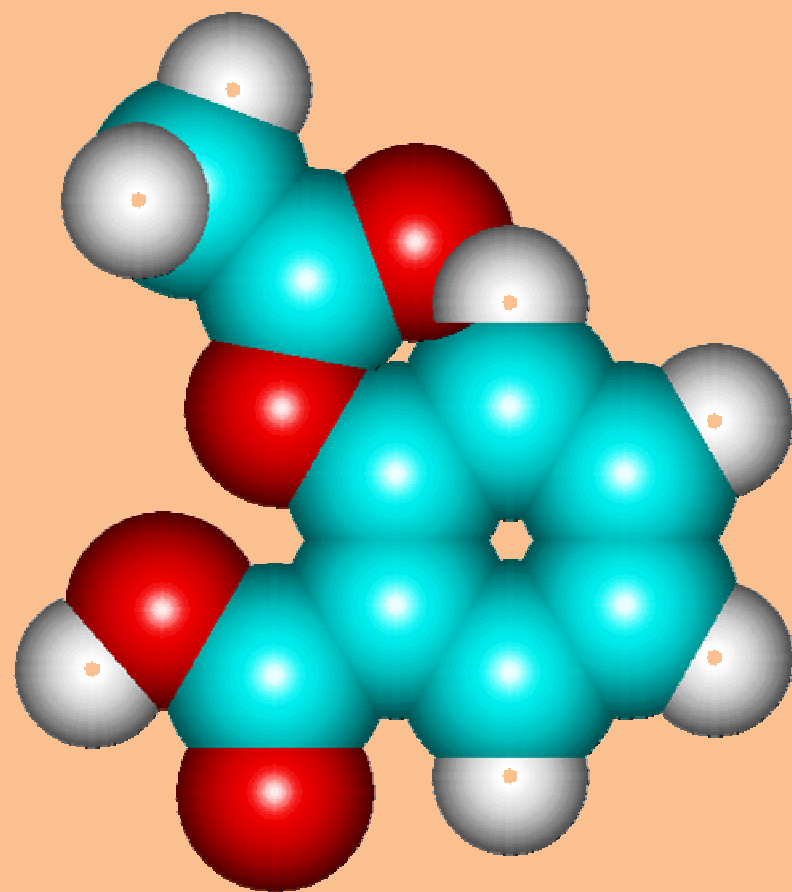
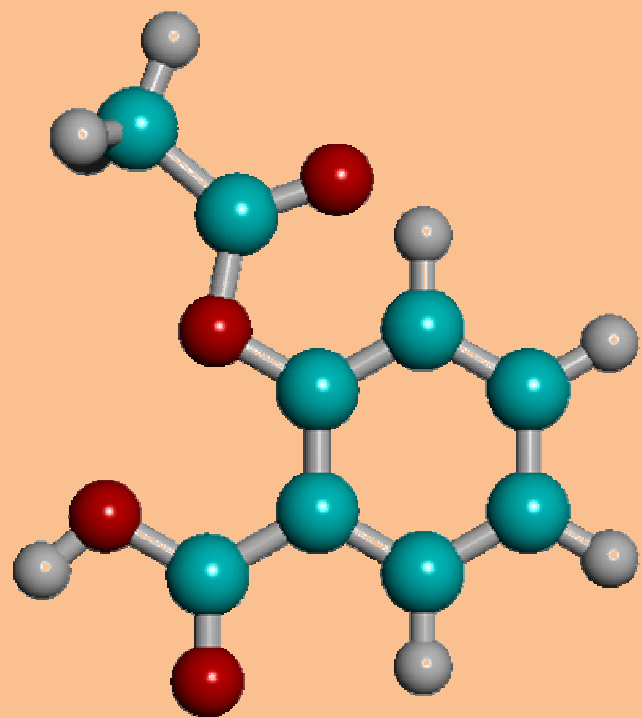


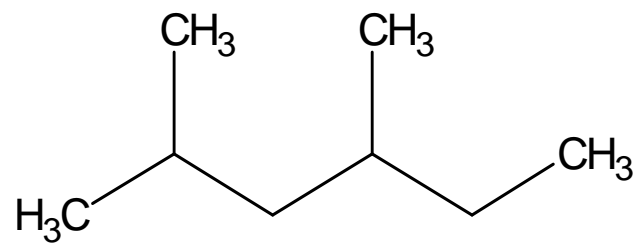
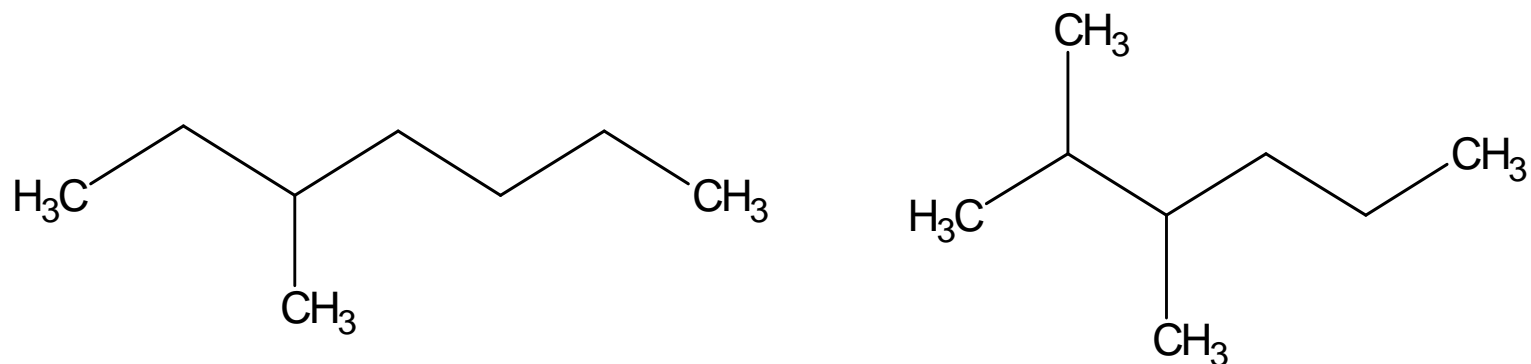
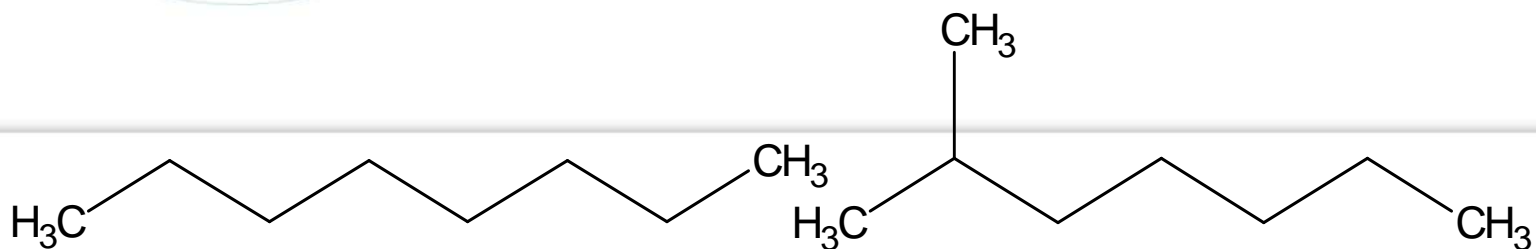
GRAFURI MOLECOLARE



= ASPIRINA

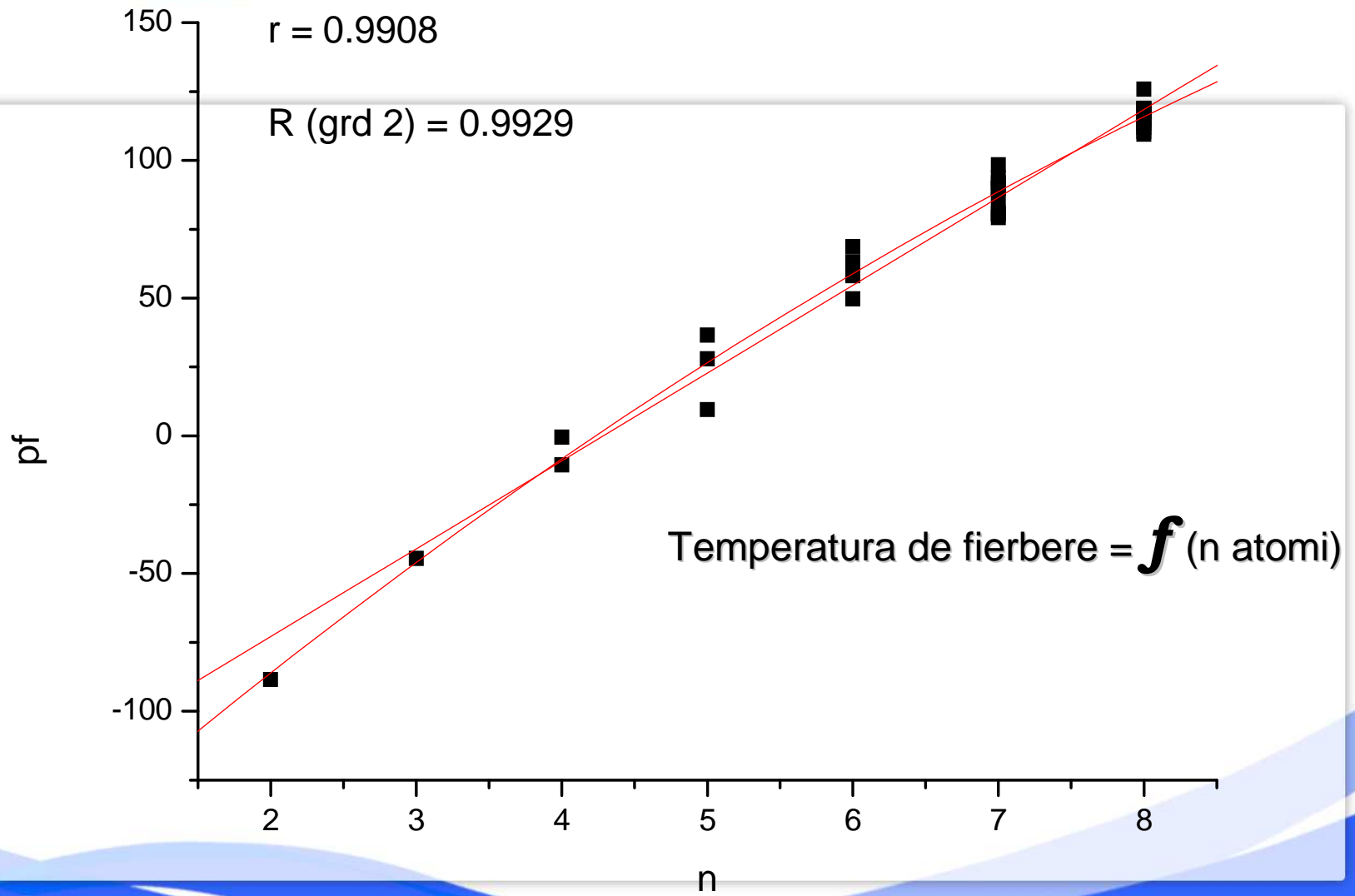






Ramificarea alcanilor (aprecieri calitative!)

Modelul topologic



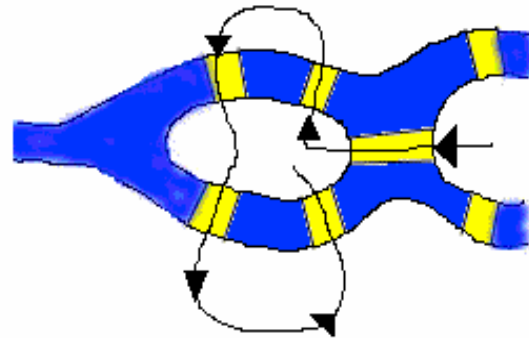
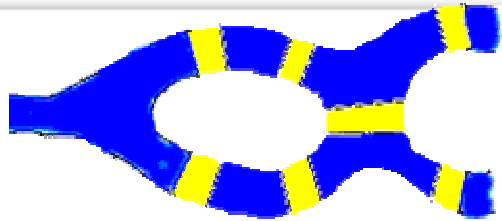
Dezavantajele reprezentării $pf = f(n)$:

- Gruparea datelor (nu se face distincție între alcanii cu același n)
- Creșterea gradului funcției nu este justificată de creșterea nesemnificativă a coeficientului de corelare
- Creșterea numărului de atomi nu aduce îmbunătățiri, iar ecuația nu poate fi folosită pentru predicții
- Nu avem informații legate de ramificarea alcanilor, factor care duce la scăderea temperaturilor de fierbere (date experimentale)

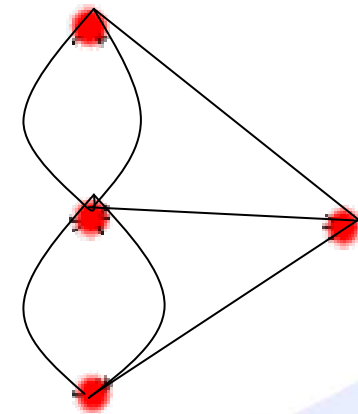
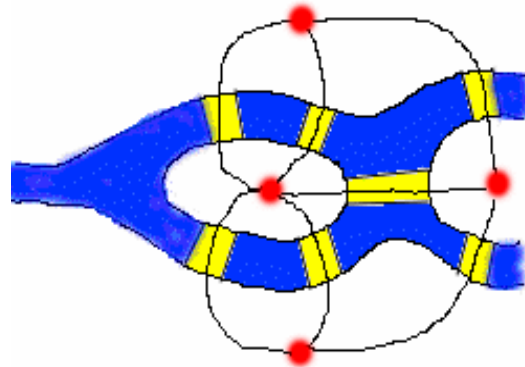
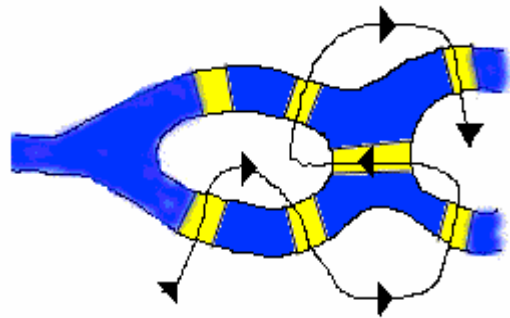
Există o metodă care scoată în evidență așezarea atomilor într-o moleculă (constituția ei), dar fără să ia în considerare natura atomilor și eventual să renunțe la hidrogeni?

În plus să se poată parcurge structura de la un capăt la celălalt, eventual pe drumul cel mai scurt, fără a se trece de două ori prin același “atom”?

Modelul topologic



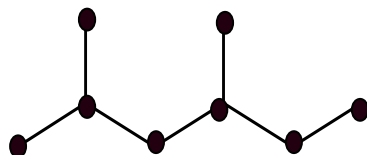
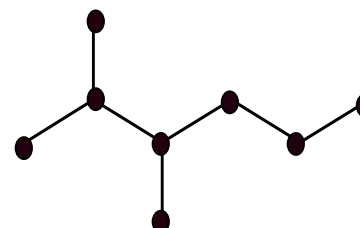
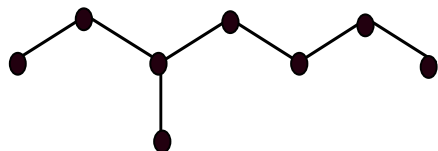
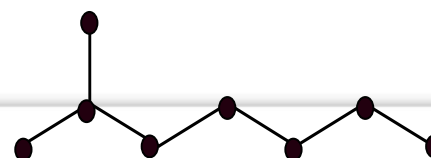
1735 → Rețele, topologie



Forma simplificată (vârfuri și arce) este aceeași indiferent de modificările pe care le aplicăm formei originale .

Graf

Modelul topologic



Structuri asemănătoare se obțin prin înlăturarea hidrogenilor, marcarea atomilor “grei” și legarea lor prin muchii (înlocuirea arcelor – situație de cele mai multe ori avantajoasă pentru chimiști)

Relații numerice în teoria grafurilor

Fie n = nr. de vârfuri într-un graf

q = nr. de muchii

p = nr. de componente (pentru grafurile chimice $p=1$)

Pentru un arbore (graf fără cicluri)

$$q = n - 1$$

Dacă există c cicluri (sau fețe)

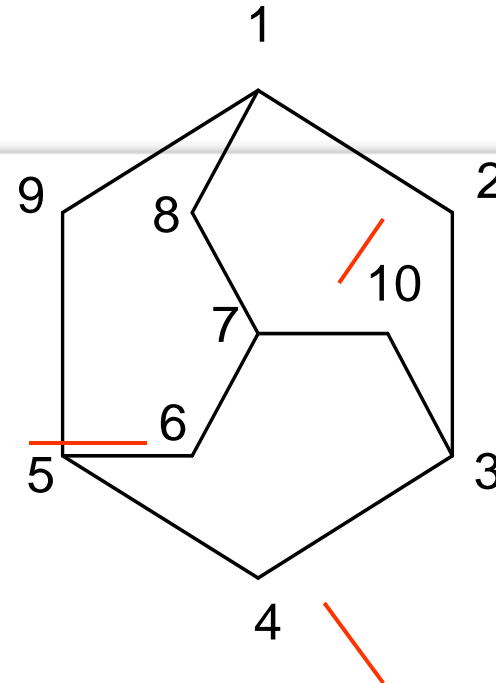
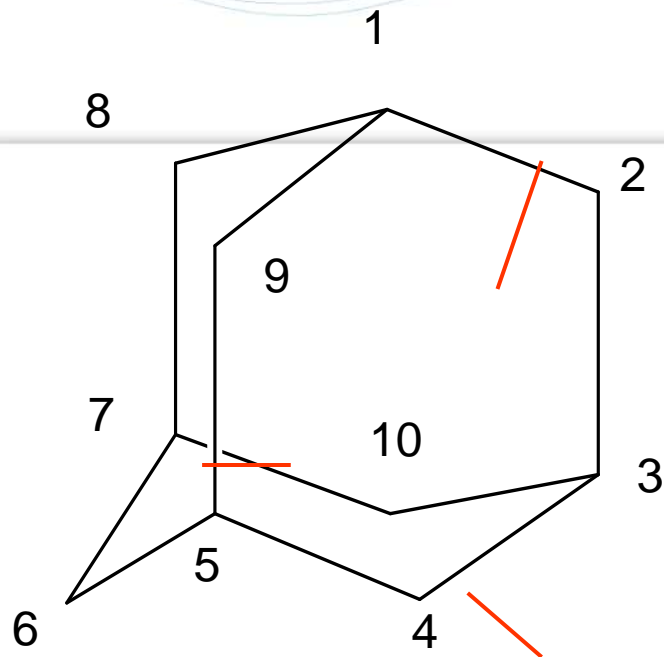
$$q = n + c - 1$$

$$v(G) = q - n + p$$

Numărul cicromatic este egal cu numărul maxim de cicluri liniar independente ale grafului

$$\sum_{i=1}^n \delta_i = 2 \cdot q$$

Modelul topologic



Adamantanul: $C_{10}H_{18}$
cicluri

$$n = 10 \quad q = 12 \quad p = 1 \quad v = 12 - 10 + 1 = 3$$

A =

Grafuri chimice - Posibilitatea de aplicare în chimie

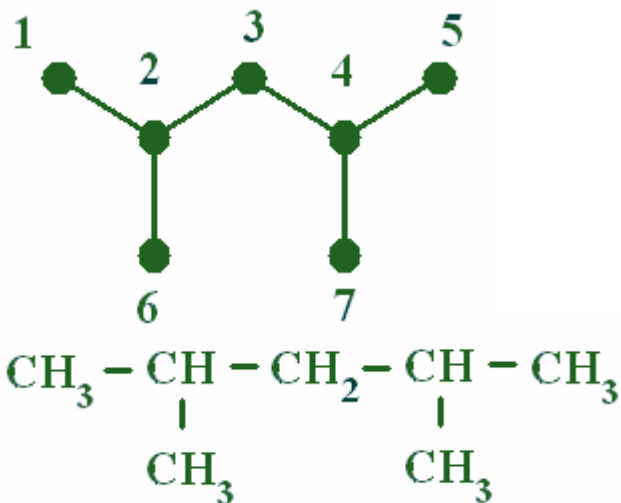
Matrici asociate

Matricea de conectivitate (adiacența) A

Matricea distanțelor D

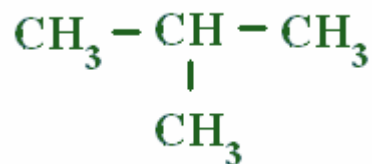
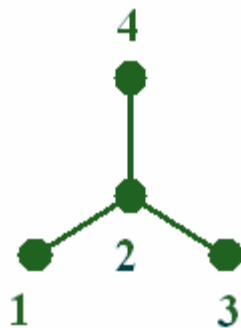
$$A = \begin{cases} 1 & i \neq j, i \text{ legat de } j \\ 0 & i = j, i \text{ nelegat de } j \end{cases}$$

$$D = \begin{cases} 1 & \exists (x_i = y_0, y_1, \dots, y_i = x_j) \text{ drum în } G \\ 0 & \text{altfel} \end{cases}$$



Matricile de adiacență și de distanță ale 2,4-dimetil-pentan

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 2 & 4 \\ 1 & 0 & 1 & 2 & 3 & 1 & 3 \\ 0 & 1 & 0 & 1 & 2 & 2 & 2 \\ 3 & 2 & 1 & 0 & 1 & 3 & 1 \\ 4 & 3 & 2 & 1 & 0 & 4 & 2 \\ 2 & 1 & 2 & 3 & 4 & 0 & 4 \\ 4 & 3 & 2 & 1 & 2 & 4 & 0 \end{bmatrix}$$



$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 1 & 2 & 2 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 2 \\ 2 & 1 & 2 & 0 \end{bmatrix}$$

**Matricile de adiacență și de distanță ale
2-metil-propanului**

Heteroatomi

Variantele Kier și Hall (Δ_i^v înlocuiesc invarianții locali)

$$\Delta_i^v = Z_i^v - H_i$$

unde Z_i^v reprezintă numărul electronilor de valență ai atomului i , iar H_i numărul de atomi de hidrogen atașați acestuia

$$\Delta_i^v = \frac{Z_i^v - H_i}{Z_i - Z_i^v - 1}$$

Z_i reprezintă numărul atomic al atomului i

Modificarea invarianților locali

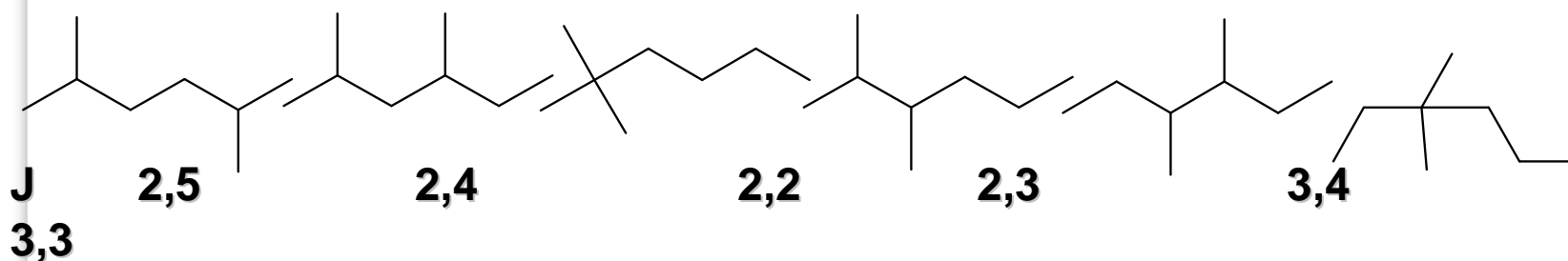
$$\delta_i^y = \delta_i \cdot y_i$$

y_i este o mărime de ponderare a invarianților locali și care depinde de natura atomilor: *raza covalentă, raza van der Waals, electronegativități*, etc.

$$s_i^y = s_i \cdot y_i$$

Aplicații ale indicilor topologici

Ordonarea alcanilor (gradul de ramificare)



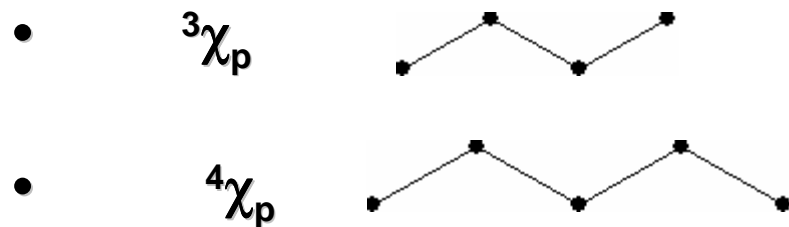
Solubilitatea alcoolilor în apă

$$-\log(S) = 11.267 \chi - 8.643 \chi^v - 9.417$$

N = 38 **R = 0.998** **s = 0.23**

Refracția molară = $f(^0\chi, ^1\chi, ^3\chi_p, ^4\chi_p)$

- **Pentru 55 alcani (pentani – nonani)**

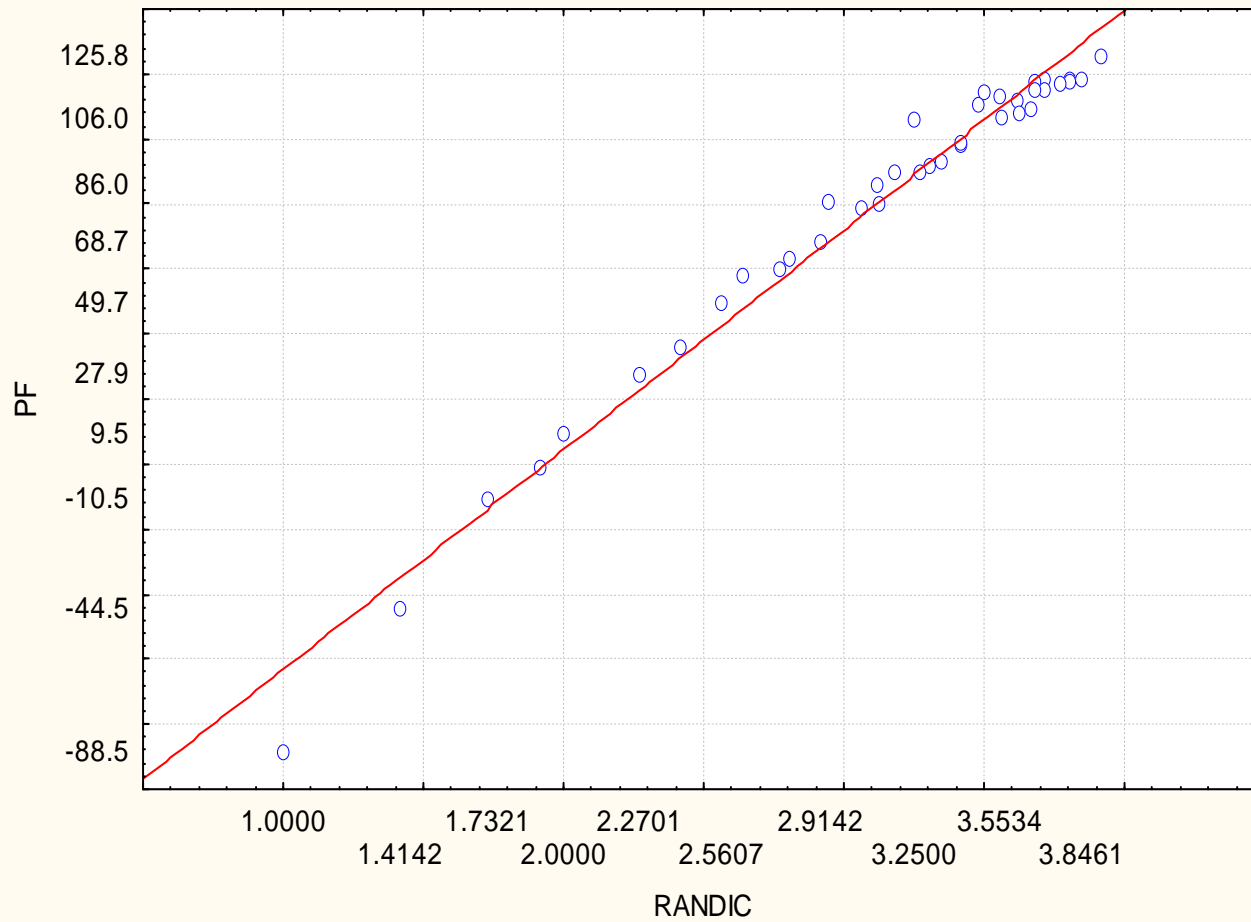


$$\text{MR} = 3.832^0\chi + 4.438^1\chi - 0.8727^3\chi_p - 0.4828^4\chi_p - 0.4558$$

$$R = 0.999999 \quad s = 0.043 \quad F = 194694 \quad n = 55$$

Scatterplot (ALCAN0 61v*39c)

$$PF = -130.8307 + 67.6087 * x$$



$$pf = f(\text{Randic}) \quad n=39 \quad r=0.988$$

Entalpii de vaporizare ΔH_v

$$\Delta {}^1\chi = {}^1\chi(\text{izomer normal}) - {}^1\chi(\text{izomer ramificat})$$

Alcoolii

$$\Delta H_v = 1.163n - 4.083 \Delta {}^1\chi + 6.641$$

$$N = 20$$

$$R = 0.993 \quad s = 0.36$$

Alcani

$$\Delta H_v = 1.191n - 2.384 \Delta {}^1\chi + 0.282$$

$$N = 44$$

$$R = 0.998 \quad s = 0.16$$

- Coeficienții $N \sim$ același în ambele cazuri \Rightarrow contribuția radicalilor alchil este aceeași
- Termenii liberi alcooli $6x >$ alcani $\Rightarrow \exists$ și alți factori (legăturile de H și dipolare)
- Coeficienții $\Delta^1\chi$ alcooli $>$ alcani \Rightarrow poziția grupei OH în moleculă este importantă (ramificarea datorată OH este mai importantă decât cea datorată CH_3).

Bibliografie

1. Silviu Barză, Luciana Maria Morogan - Algoritmica grafurilor, Ed. Fundației României de mâine Bucuresti,
2. Sylvester, J. J. On an application of the new atomic theory to the graphical representation of the invariants and covariants of binary quantics - with three appendices, Am. J. Math., 1874.
3. Trinajstić, N. Chemical Graph Theory , CRC Press, Inc., Boca Raton, Florida, 1983.
4. Kier, L. B.; Hall, L.H. Molecular Connectivity in Chemistry and Drug Research, Acad. Press, 1976.
5. Balaban, A. T.; Filip, P. Computer program for topological index J (average distance sum connectivity), MATCH, Commun. Math. Comput. Chem., 1984.
6. Ionescu, T. Grafuri-Aplicatii, Ed. Ped. Bucharest, 1973.
7. Oprescu, D.; Bejan, L.; Patrascu, V. Informatica, Ed. Niculescu, Bucuresti, 2001.
8. Berge, C. Teoria Grafurilor si Aplicatiile Ei, Ed. Tehnica , Bucuresti , 1969.
9. Kier, L. B. ; Hall, L. H. Molecular Connectivity in Chemistry and Drug Research, Acad. Press, 1976.
10. Diudea, M. V.; Gutman, I.; Jäntschi, L. Molecular Topology, NOVA, New York, 2002.
11. Diudea, M. V.; Florescu, M. S.; Khadikar, P. V. Molecular Topology and Its Applications, EFICON, Bucharest, 2006.