Bridging The Gap Between Chemistry And Chemical Engineering

A critical bottleneck in the computational science grand challenge problem of bridging the gap between space and time scales, from individual reactions (length scale of nanometers and time scale of femtoseconds) to reacting flow (length scale of meters and time scale of minutes to hours).

Bridging The Gap Between Chemistry And Chemical Engineering

A major problem in constructing a chemical kinetic model is the very large number of possible reactions, products, and reactions intermediates involved.

Examples:

Combustion of n-heptane(1) : 3662 reactions
Involving 470 species
Pyrolysis of tetradecane(2) : 479206 reactions and 19052 species

1. P.A. Glaude, F. Battin-Leclerc, R. Fournet, V. Warth, G.M. Come, G. Scacchi, *Combustion And Flame* 122, **2000,** pp. 451-462
2. M.J. De Witt, D.J. Dooling, L.J. Broadbelt, *Ind. Eng. Res.* 39, **2000**, pp. 2228-2237

Bridging The Gap Between Chemistry And Chemical Engineering

NetGen (L. Broadbelt et al.):
Matrix representation based system.
Single molecule is represented as a square matrix whereas chemical reaction as matrix transformation.
Example: ethane fission

$$C\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}+\begin{bmatrix}1 & -1\\ -1 & -1\end{bmatrix}=C\cdot\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

$$CH_3-CH_3 \longrightarrow \cdot CH_3+\cdot CH_3$$

Literature:

1. L.J. Broadbelt, S.M. Stark, M.T. Klein, *Ind. Eng. Chem. Res.*, **1994,** 33, pp. 790-799
2. R.G. Susnow, A.M. Dean,, W.H. Green, L.J. Broadbelt, *J. Phys. Chem. A* 101, **1997**, pp. 3731-3740

Bridging The Gap Between Chemistry And Chemical Engineering

NetGen (L. Broadbelt et al.):

Advantages:

✓ Simplicity – coding chemical reactions is easy and generator is very fast

Disadvantages:

✓ Only limited set of reaction and molecular species can be coded using this approach, no cyclic compounds handled
✓ The notation is unclear for a-lay audience
✓ Reaction types are hard-coded

Linear notation approach

EXGAS (Come et al.)

Linear chemical notation based system. Each specie is externally represented by a 1D-string. This notation is non-ambiguous, but also non canonical. That means that species are not represented uniquely, an complicated algorithm is required to handle them.

$$CH3/CH3 \rightarrow CH3 + CH3$$

Literature:
1.V. Warth, F. Battin-Leclerc, R. Fournet, P.A. Glaude, G.M. Come, G. Scacchi, *Computers and Chemistry* 24, **2000**, pp. 541-560
2. P.A. Glaude, F. Battin-Leclerc, R. Fournet, V. Warth, G.M. Come, G. Scacchi, *Combustion And Flame* 122, **2000,** pp. 451-462

Linear notation approach

EXGAS (Come et al.)

Advantages:

✓ Wider (then in matrix approach) spectrum of reaction classes available
✓ Chemical database associated with the program, group additivity approach implemented

Disadvantages:

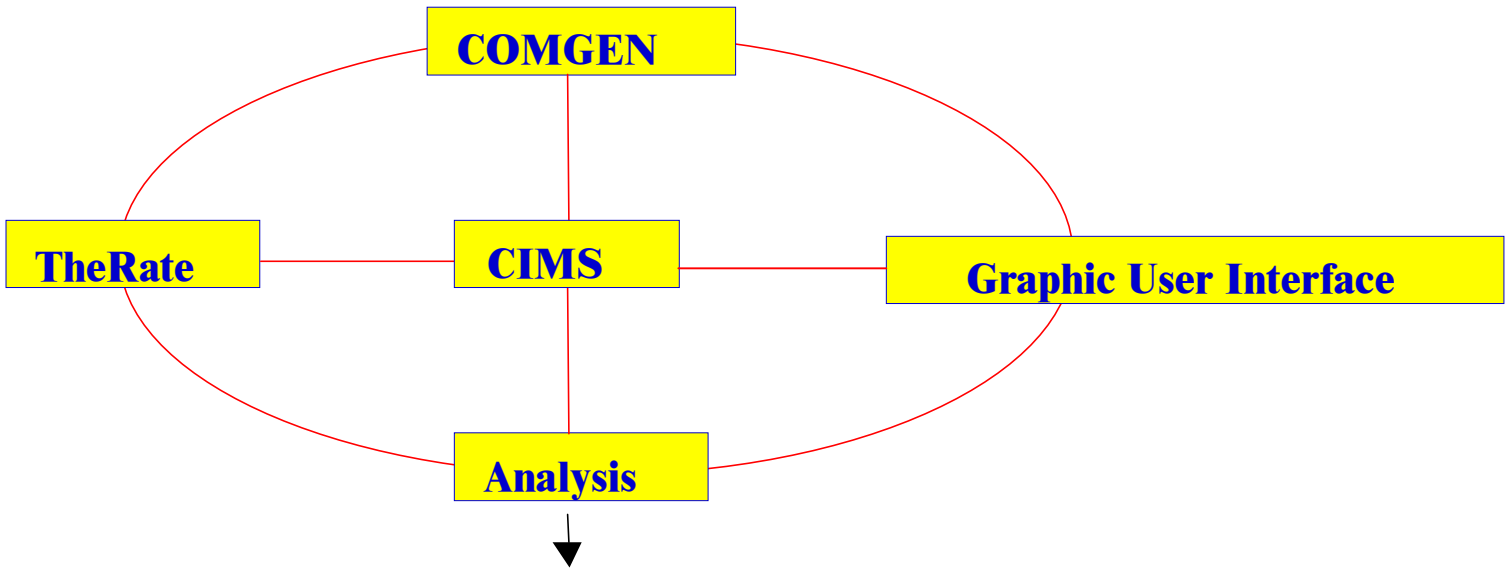
✓ The canonicity algorithm and the manner of handling species is very complicated and depends on the class of specie (for example different canonicity algorithms are implemented to cyclic and acyclic species)
✓ Reaction types are hard-coded

COMGEN – A New Automatic Mechanism Generator

A new graph theory based automatic reaction mechanism generator **COMGEN** is presented. **COMGEN** will be a part of integrated web-based problem solving environment for research in reaction kinetics and engineering.

COMGEN is based on chemical graph theory. Species are naturally represented as molecular graphs. The mechanism is generated using a set of reaction patterns (sub-graphs). They are internal representation for a given class of reactions thus eliminating the possibility of generating unimportant species a priori. Every specie is canonically represented as a set of real numbers (topological indices)

Reaction Design Solving Environment

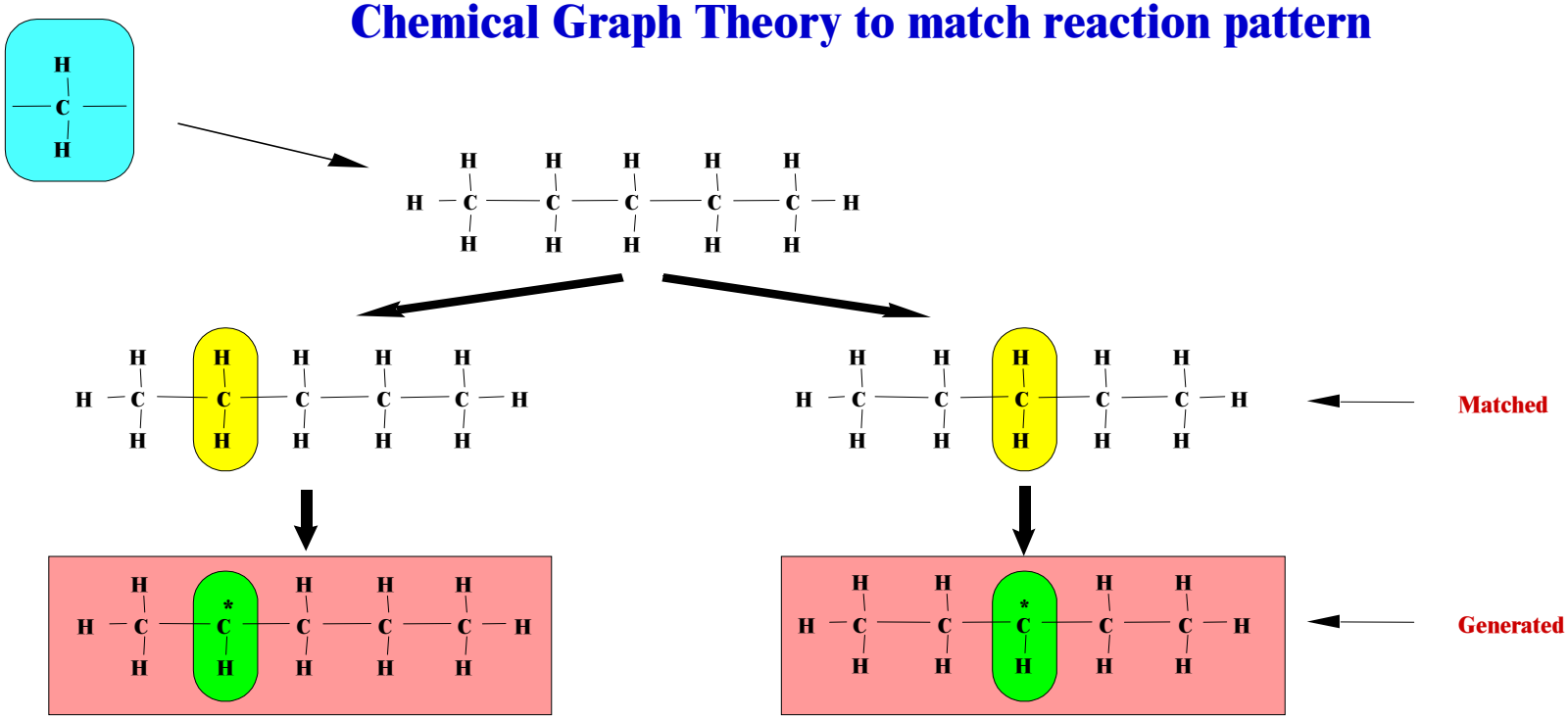


Computational Reaction Engineering

http://vklab.hec.utah.edu/

COMGEN Procedure

Chemical Graph Theory to match reaction pattern

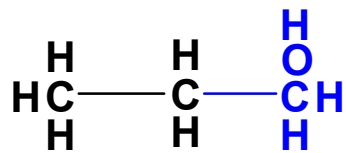


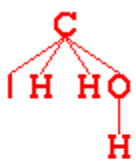
Pattern Recognition

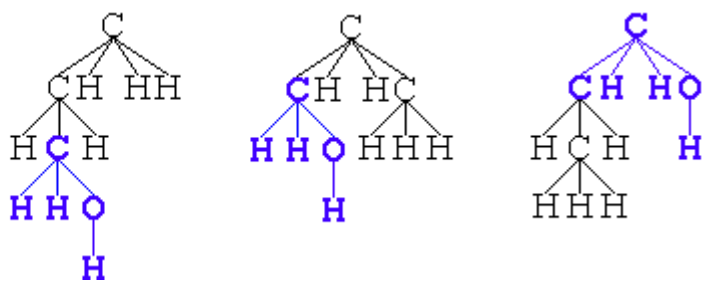
Pattern recognition is accomplished by comparing the tree-like structures starting from C atom in the pattern and every C atom in the molecule:

Pattern

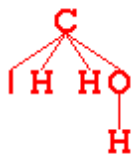
Molecule

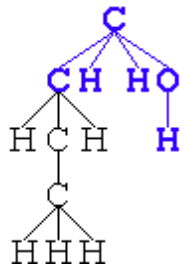


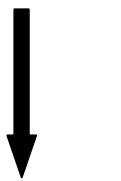




COMGEN Pattern Recognition







C C C C
| H H O
H
1 2 3 4

C C C C
C H H O
H
a b c d

(1→a 2→b 3→c 4→d)

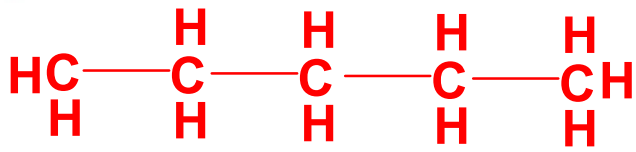
COMGEN Molecular Isomorphism

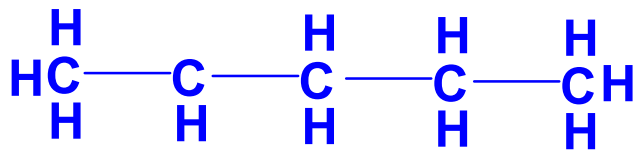
Molecular isomorphism is a procedure that determines whether newly generated molecule has been already generated or not.
A **molecular ID** concept has been implemented – each species (graph) is represented by a set of real numbers (topological indices).

Indices calculated:

1. Randic indices of all possible orders (from 0^{th} to $(n-1)^{th}$) when n is a number of heavy atoms in molecule
2. Wiener indices
3. Schulz indices (MTI)
4. determinant of (distance + adjacency) matrix (TI)
5. WID index.

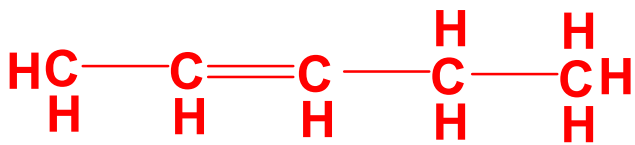
COMGEN Molecular Isomorphism

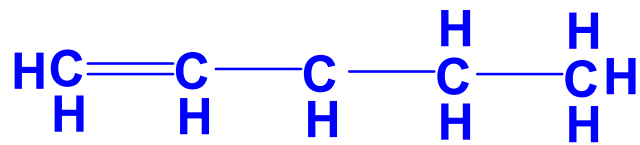




χ_0	3.8284271	3.9915638
χ_1	2.2071068	2.1927053
χ_2	1.2071068	1.1969234
χ_3	0.6035534	0.5773503
χ_4	0.2500000	0.2886751
χ_5	0.0000000	0.0000000
Wiener	20.000000	20.0000000
MTI	74.000000	74.0000000
WID	5.0581531	5.0581531

COMGEN Molecular Isomorphism



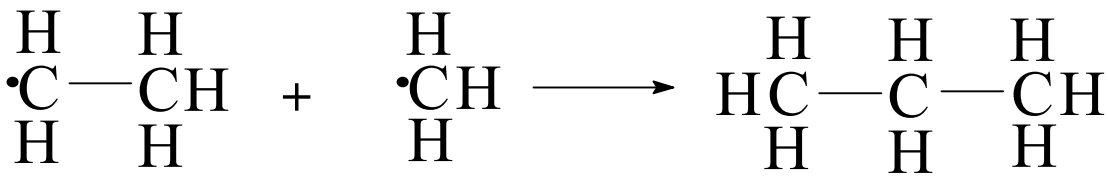


χ_0	3.5689141	3.5689141
χ_1	1.8569367	1.8569367
χ_2	0.8796528	0.8796528
χ_3	0.4023689	0.4023689
χ_4	0.1666667	0.1666667
χ_5	0.0000000	0.0000000
Wiener	17.0000000	18.0000000
MTI	62.5000000	67.0000000
WID	5.0743406	5.0656449

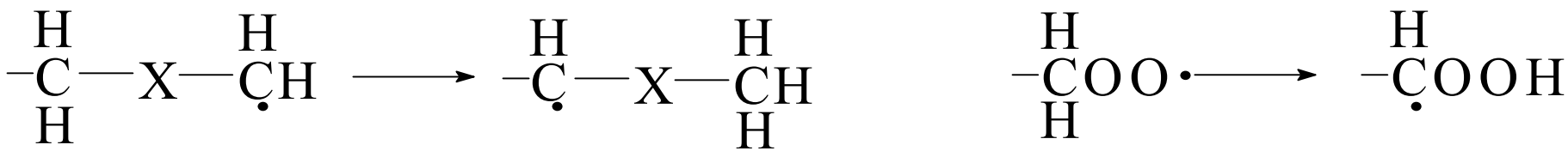
COMGEN selected reaction types available in literature

Bond fission and beta-scission

Radical recombination:

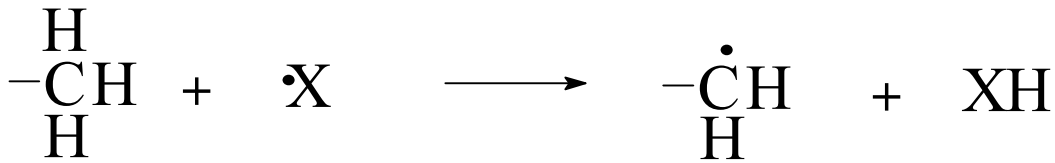


Shifts

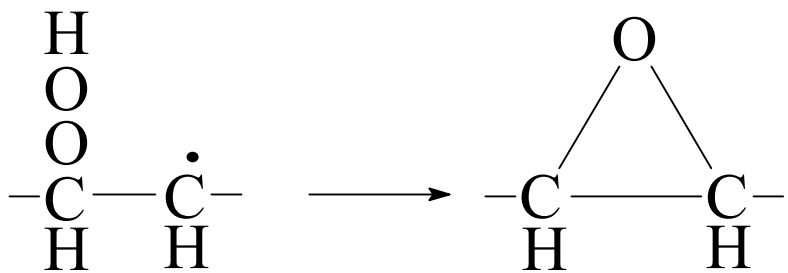


COMGEN selected reaction types available in literature

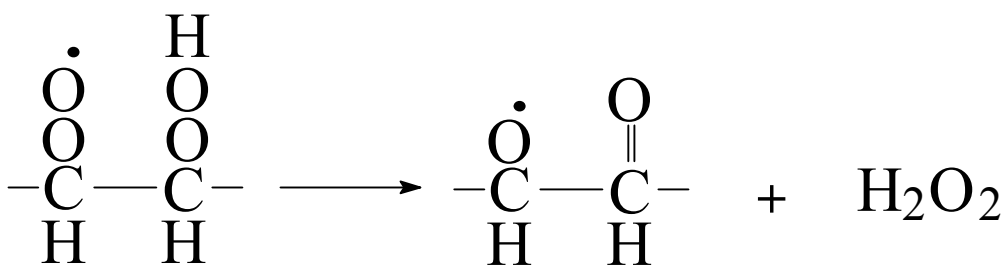
Meta - thesis



Cyclic ether formation

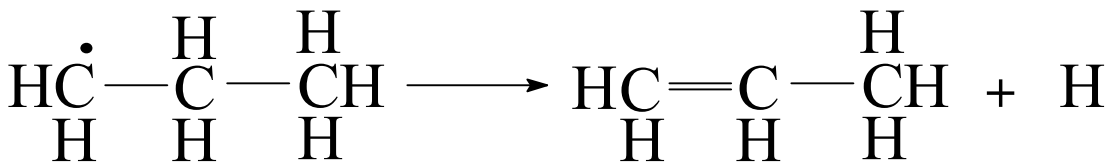


Branching agent formation

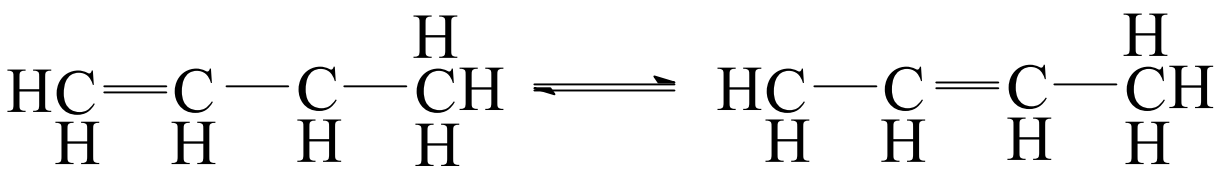


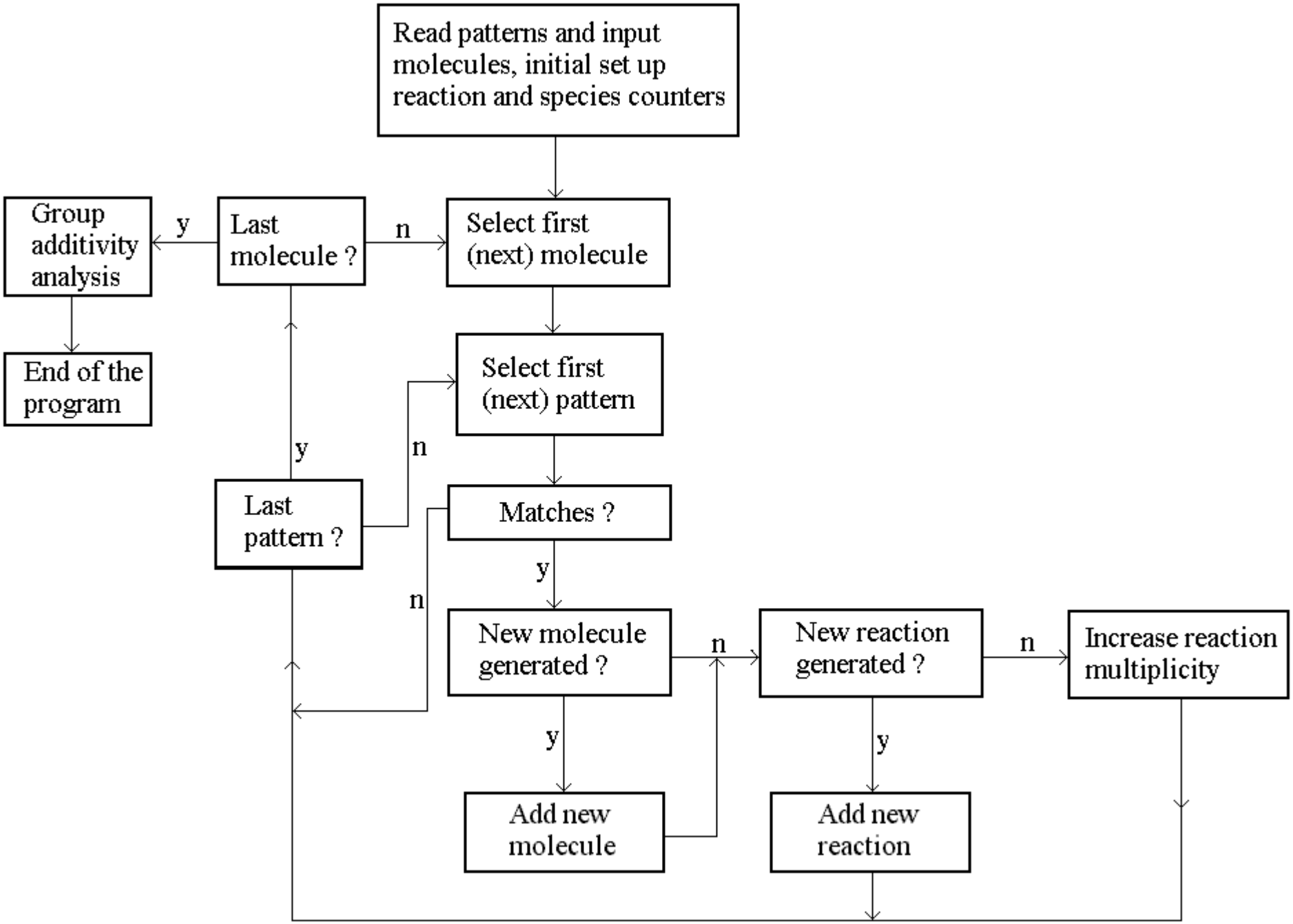
COMGEN selected reaction types not available in literature

Diradicals are automatically converted into alkenes:



The program is able to recognize resonance structures:




Read patterns and input molecules, initial set up reaction and species counters
Select first (next) molecule
Last molecule ?
y
n
Group additivity analysis
End of the program
Select first (next) pattern
Last pattern ?
y
n
Matches ?
y
n
New molecule generated ?
n
New reaction generated ?
n
Increase reaction multiplicity
y
y
Add new molecule
Add new reaction

COMGEN Thermochemistry

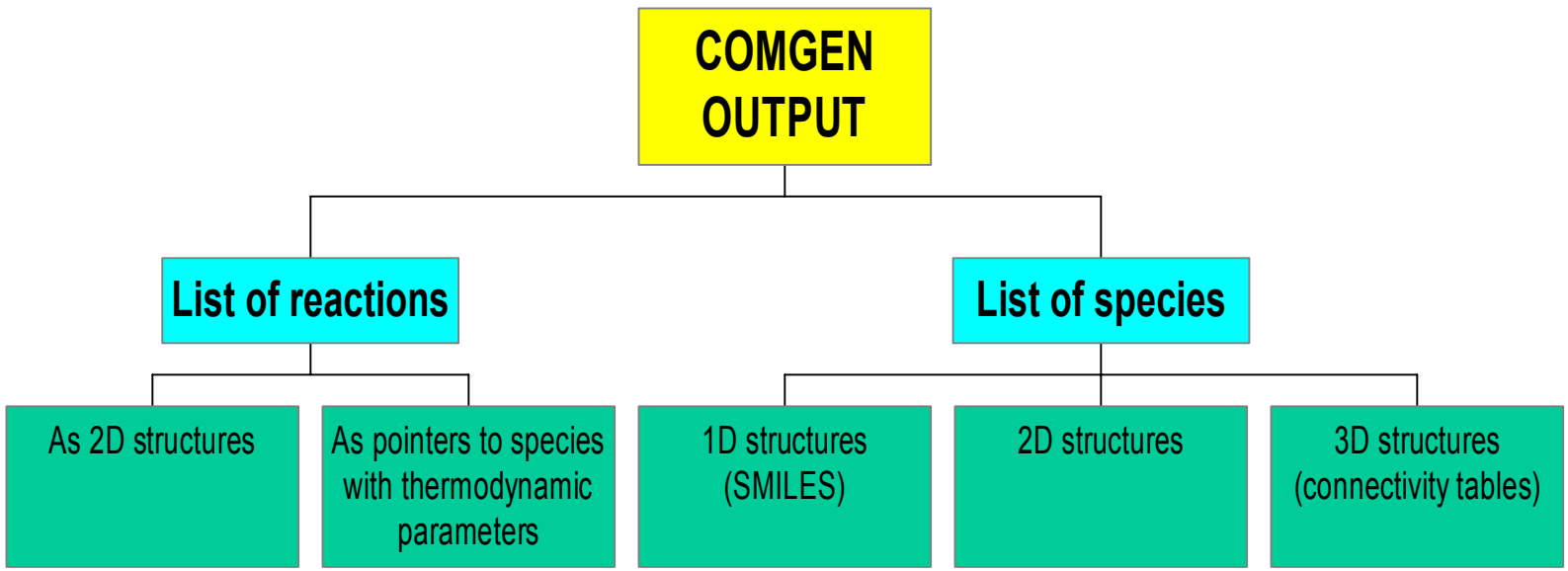
A Benson group additivity approach is used. Each compound is treated as a sum of groups, each of them having been defined as an atom (ligancy >1) and its ligands.
Values of thermochemical parameters are generated "on the fly". An appropriate ring and symmetry corrections are applied, the program is able to determine symmetry number of each molecule.

The mean accuracy claimed by Benson for molecules is ± 0.5 kcal/mol for H_f^0 and ± 1 e.u./mol for S^0 and C_p^0, for free radicals it it claimed to be ± 1 kcal/mol for H_f^0 and ± 2 e.u./mol for S^0 and C_p^0 .

COMGEN example results

n – propane oxidation up to C3 species – 244 species and 4952 reactions with thermochemical parameters

n – heptane oxidation up to C7 species – 14584 species and 44956 reactions with thermochemical parameters



Acknowledgments

The authors are grateful for financial support from the National Science Foundation by the NSF-NATO Postdoctoral Fellowship in Science and Engineering (grant DGE-000798).