



r.west@neu.edu



richardhwest



rwest

Reaction Mechanism Generator: Toward High-Throughput Transition State Calculations, and Interpreting Existing Kinetic Models

Richard H. West

4 August 2014

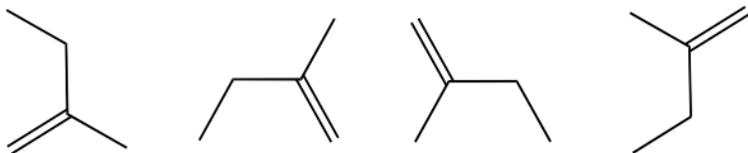


Northeastern.edu/comocheng

Presented at the 2nd International Workshop on Flame Chemistry in San Francisco

Automatic reaction mechanism generation needs methods to:

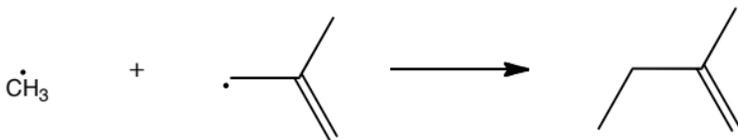
1. Represent molecules
(and identify duplicates)



The common theme to the two projects in the talk is Reaction Mechanism Generator software.
First, a brief introduction to it.

Automatic reaction mechanism generation needs methods to:

2. Create reactions (and then new species)



Automatic reaction mechanism generation needs methods to:

3. Choose which reactions to include
(and which to leave out)

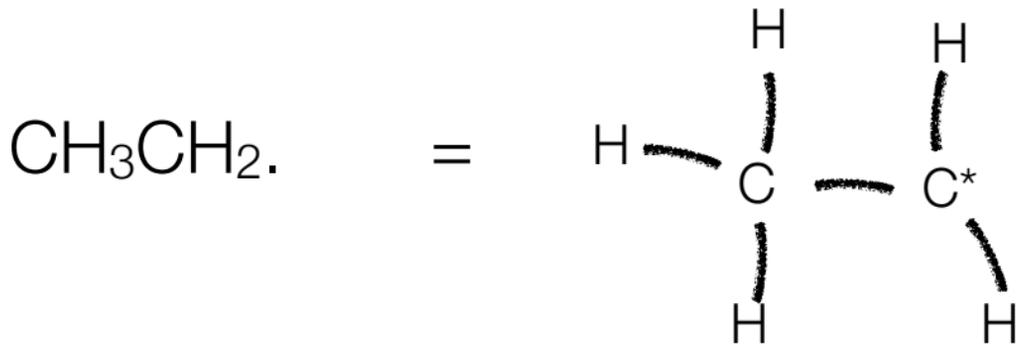
You'll notice a lot of slide designs are top-heavy.

This is because the room layout meant most people couldn't see the bottom of the screen!

Automatic reaction mechanism generation needs methods to:

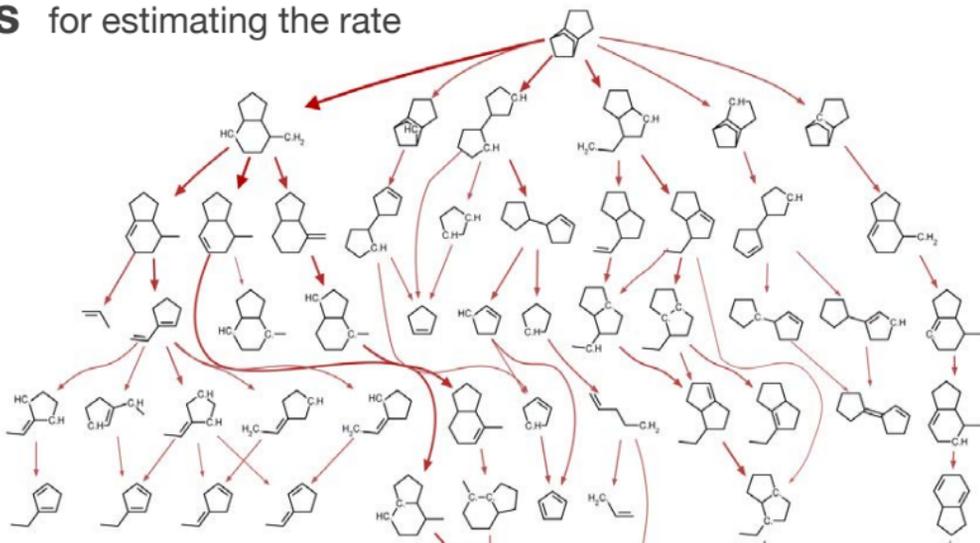
4. Estimate thermo and kinetic parameters
(quickly!)

Molecules are represented as 2D graphs



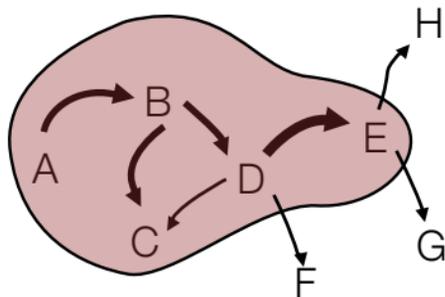
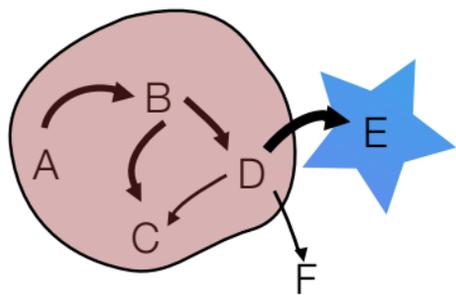
Reaction families propose all possible reactions with given species

- **Template** for recognizing reactive sites
- **Recipe** for changing the bonding at the site
- **Rules** for estimating the rate

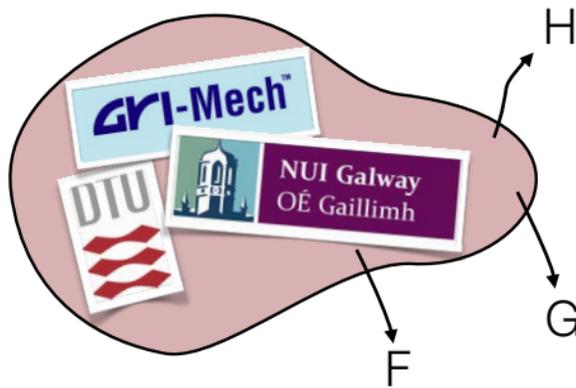


There are ~40 reaction families such as as Hydrogen abstraction, unimolecular homolysis, radical addition to a double bond...

All possible reactions are found, then core is expanded by following the fastest pathways.

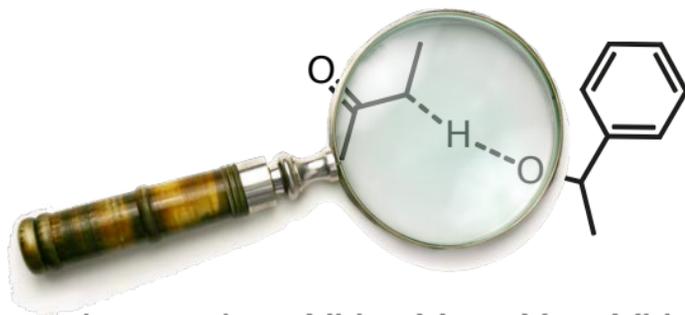


Model can be started from “seed mechanism” (if you have this in RMG format)



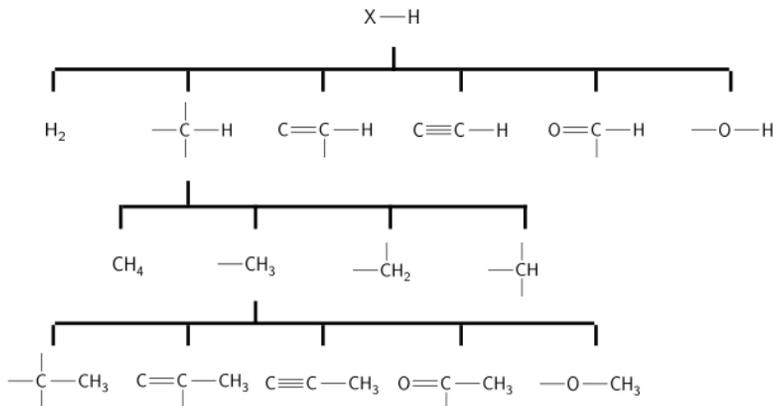
Getting it in “RMG format” is the hard part, and will be addressed in the second half of the talk.

Rate estimates are based on the local structure of the reacting sites.



- Hydrogen abstraction: $XH + Y. \rightarrow X. + YH$
- Rate depends on X and Y.

Rate estimation rules are organized in a tree



Part of the tree for X

The most generic expression goes at the top of the tree.

You use the most precise node that you can, then fall up until you find data.

New “RMG-Py” designed to be more extensible and developer-friendly.

- See poster on Wednesday



We presented two posters describing this work at the 35th International Symposium on Combustion occurring the week following this presentation.

New "RMG-Py" des more extensible and

- See poster on Wednesday





35th International Symposium on Combustion, August 3-8, 2014



RMG-Py: the Python version of Reaction Mechanism Generator

John W. Allen¹, Pierre L. Bhoorasingh¹, Connie W. Gao¹, William R. Richard II², Richard H. West¹
¹Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge MA, USA
²Department of Chemical Engineering, Northeastern University, Boston MA, USA

REACTION MECHANISM GENERATOR SOFTWARE



USE AUTOMATICALLY WRITTEN REACTION RULES



SKIP THE FIRST 3000 RELEASED IN 2004



WE ARE NOW WRITING THE FIRST INDEPENDENT VERSION



WE ARE NOW WRITING THE FIRST INDEPENDENT VERSION



WE ARE NOW WRITING THE FIRST INDEPENDENT VERSION



BASED ON CHEMISTRY RULES

MANY FEATURES HAVE BEEN ADDED

DESIGNED TO BE EASY TO EXTEND

Model Generation Algorithm

Rule-based algorithms expands a core reaction network relating reaction paths according to their flux, to create a kinetic model for any given starting species and reactor conditions.



Multiple conditions (TXX) can be run simultaneously to build a model valid over a wide range of situations.

Pruning of insignificant species from the "ridge" helps build large models with limited computer memory.

Reaction Kinetics Estimation

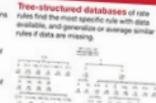
Reaction Families contain templates and recipes to generate sets of feasible reactions. Template:

$$R_1-X + Y \rightleftharpoons R_2-X + R_3-Y$$

Reaction:

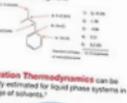
$$H_2 + H_2 \rightleftharpoons H + H - O - H$$

Tree-structured databases of rate using molecular graph theory calculations, and generation or average similar rates if data are missing.

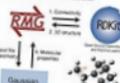


Species Property Estimation

Group Additive Thermochemistry estimates quickly find thermochemical properties by summing contributions from each molecular fragment.



On-the-fly Quantum Mechanics calculations improve thermochemistry estimates for fused rings species where group-additive methods are inaccurate.



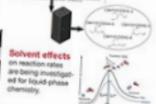
Ongoing Projects

Transition state geometry predictions for high throughput QM TST kinetic calculations. See poster WSP173.



Parallelization and algorithm modification are needed to make better use of computational resources for very large models.

Solvent effects on reaction rates are being investigated for liquid phase chemistry.



Recent Developments

Web UI and Tools allow an easy gateway to locate, estimate, verify, or improve kinetics and thermodynamic data. Visit <http://rmg.mit.edu/>



Nitrogen Chemistry required major modifications due to its diverging nomenclature, but recent changes have improved, and nitrogen data are now ready!



Acknowledgements

Authors would like to thank the Office of the Undergraduate Research Program (URP) for partial funding of this research. The progress of the development of RMG-Py has been supported by the National Science Foundation under Grant No. 4420771. Funding for the development of RMG-Py has also been provided by the National Science Foundation under Grant No. 08-0620110. Some work was funded in part by the National Energy Research Center (NERC) at the Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, Italy. We would like to thank the staff of the Department of Chemical Engineering at Northeastern University for providing us with the Department of Chemical Engineering facilities.

References

1. Bhoorasingh, P. L.; Allen, J. W.; Gao, C. W.; Richard, W. R.; West, R. H. "A Rule-Based Derivation of a Kinetic Model for Gas-Phase Reaction." *J. Phys. Chem. A* 2004, 108, 10211-10220.

2. Allen, J. W.; Gao, C. W.; Richard, W. R.; West, R. H. "A Rule-Based Derivation of a Kinetic Model for Gas-Phase Reaction." *J. Phys. Chem. A* 2004, 108, 10211-10220.

3. Allen, J. W.; Gao, C. W.; Richard, W. R.; West, R. H. "A Rule-Based Derivation of a Kinetic Model for Gas-Phase Reaction." *J. Phys. Chem. A* 2004, 108, 10211-10220.

4. Allen, J. W.; Gao, C. W.; Richard, W. R.; West, R. H. "A Rule-Based Derivation of a Kinetic Model for Gas-Phase Reaction." *J. Phys. Chem. A* 2004, 108, 10211-10220.

References

5. Gao, C. W.; Allen, J. W.; Gao, C. W.; Richard, W. R.; West, R. H. "A Rule-Based Derivation of a Kinetic Model for Gas-Phase Reaction." *J. Phys. Chem. A* 2004, 108, 10211-10220.

6. Allen, J. W.; Gao, C. W.; Richard, W. R.; West, R. H. "A Rule-Based Derivation of a Kinetic Model for Gas-Phase Reaction." *J. Phys. Chem. A* 2004, 108, 10211-10220.

7. Allen, J. W.; Gao, C. W.; Richard, W. R.; West, R. H. "A Rule-Based Derivation of a Kinetic Model for Gas-Phase Reaction." *J. Phys. Chem. A* 2004, 108, 10211-10220.

8. Allen, J. W.; Gao, C. W.; Richard, W. R.; West, R. H. "A Rule-Based Derivation of a Kinetic Model for Gas-Phase Reaction." *J. Phys. Chem. A* 2004, 108, 10211-10220.

We presented two posters describing this work at the 35th International Symposium on Combustion occurring the week following this presentation.

New "RMG- more extens

- See poster of



Northeastern University
College of Engineering

Improving Detailed Kinetic Models of Bio-oil Gasification via Rate Rule Calculations

Fariba Seyedzadeh Khanshan and Richard H. West
Department of Chemical Engineering, Northeastern University, Boston, MA, USA

RMG

Bio-oil has complex chemistry

- Bio-oil's main component is fragments of cellulose, hemicellulose, lignin and other oxygenated molecules.
- The presence of oxygen has important effects on bond dissociation energy (BDE) of a molecule, and BDEs determine if either C-C or C-O bond breaking is more feasible in primary decomposition of oxygenated molecules.

We want to build a bio-oil gasification model

RMG generates a reaction network from a set of input molecules and reaction rules. The network is then used to calculate the reaction rates and the overall reaction rates for a given set of conditions.

Specific reaction classes are needed

So far we have investigated three reaction classes for bio-oil gasification that are missing in RMG:

1. Range for changing the bonding
2. Rate coefficients for the rate parameter

The database should be updated

1. Examples for incorporating reaction sites

2. Range for changing the bonding

3. Rate coefficients for the rate parameter

Rate rules needed for new families

Kinetic parameter calculations using CBS-QB3 for a set of similar reactions show that rate rules can be used for these families.

New reaction classes have a significant effect on bio-oil gasification modeling

- RMG generated model is compared with two sets of experiments covering different temperature ranges and comparison shows that there are differences in CO₂ and CO yields between the experiments and RMG-built model.
- Products from C-O bond breaking reactions are predominantly H₂O and CO₂, the main gas products of bio-oil gasification, and have a significant impact on model prediction.

For future study, we plan to...

- continue kinetic studies of the decomposition of highly oxygenated compounds
- use *ab initio* calculations to determine unknown rate parameters of these reactions
- build more reliable detailed kinetic models for bio-oil

References

1. RMG: Reaction Mechanism Generator
http://www.oden.utah.edu/~rmg/

2. S. S. Gunzler, D. E. Reshler, G. E. Meitzner, J. C. Sauer, M. J. Pilling, *Chem. Phys. Lett.* 2007, 441, 1-10

3. M. J. Pilling, M. J. Pilling, *Chem. Phys. Lett.* 2007, 441, 1-10

4. J. S. Francisco, M. J. Pilling, *J. Phys. Chem. A* 2006, 110, 9106

35th International Symposium on Combustion
August 3-8, 2014

Mechanism Generator

West, Richard H. West
(Cambridge MA, USA
or Boston, MA, USA)

Free-structured databases of rate reactions

Free-structured databases of rate reactions that are not specific rule data available, and generate or average similar rules if data are missing.

Pressure-dependent unimolecular reaction networks explored and solved automatically via Master Equation calculations.

Libraries of known reaction rates are used as a preference for group-based estimates.

Mechanism importer tool takes models published in Chemical format and identifies species as the researchers can compare and learn from the models. See WSP109.

Solvent effects on reaction rates are being investigated for liquid phase kinetics.

Documentation and installation instructions are receiving renewed attention.

Presentations

1. A new generation software interface for unimolecular reaction networks

2. A new generation software interface for unimolecular reaction networks

References

1. RMG: Reaction Mechanism Generator
http://www.oden.utah.edu/~rmg/

2. S. S. Gunzler, D. E. Reshler, G. E. Meitzner, J. C. Sauer, M. J. Pilling, *Chem. Phys. Lett.* 2007, 441, 1-10

3. M. J. Pilling, M. J. Pilling, *Chem. Phys. Lett.* 2007, 441, 1-10

4. J. S. Francisco, M. J. Pilling, *J. Phys. Chem. A* 2006, 110, 9106

We presented two posters describing this work at the 35th International Symposium on Combustion occurring the week following this presentation.

Toward High-Throughput Transition State Calculations

Pierre L. Bhoorasingh



Acknowledgment is made to the Donors of
the American Chemical Society Petroleum
Research Fund for support of this research.

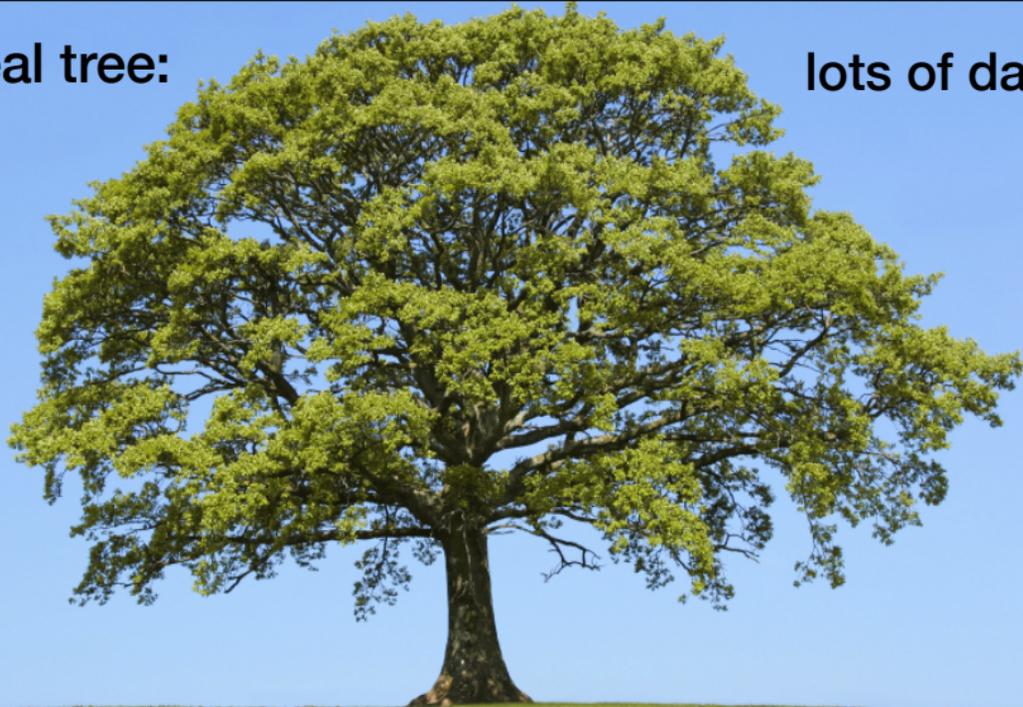


Northeastern.edu/comocheng

Now that you understand RMG, here's the first of the two projects.

Ideal tree:

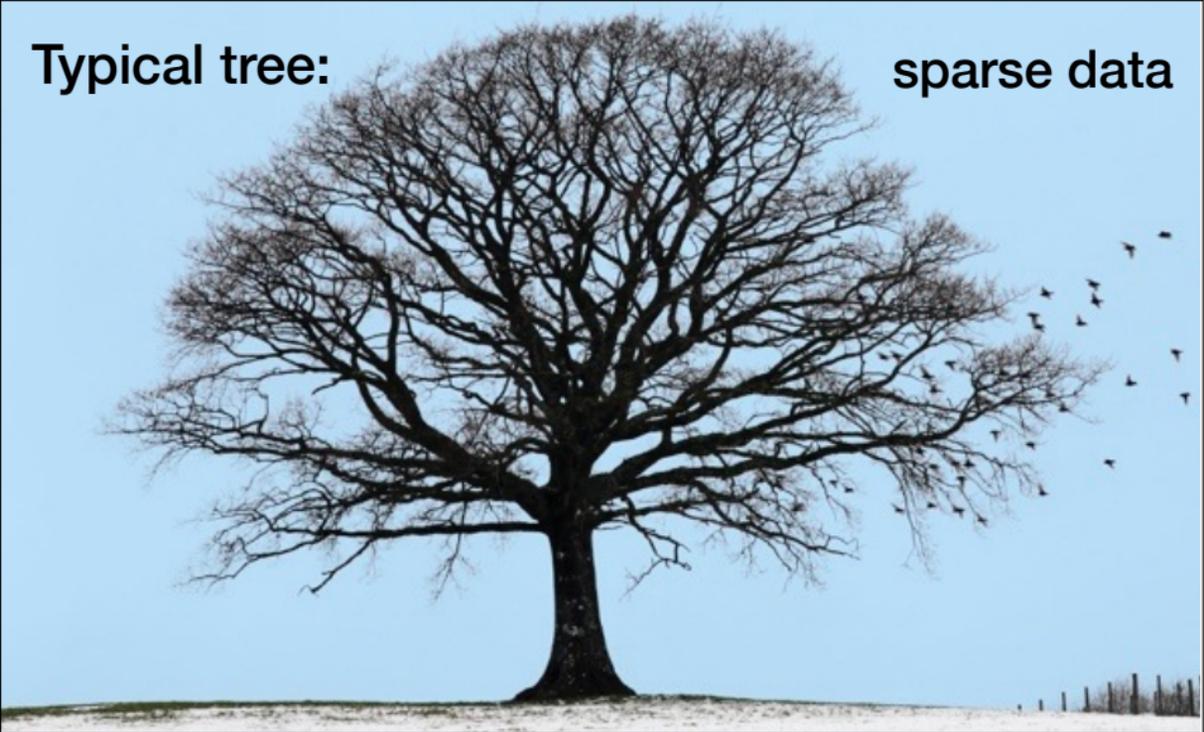
lots of data



Ideally we climb the tree, find the matching node for our reaction,
and pluck the leaf containing the rate expression.

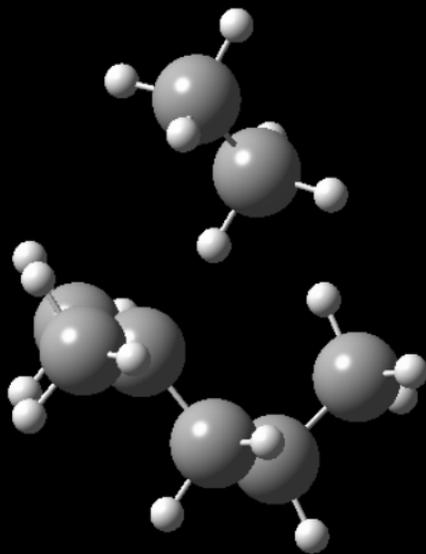
Typical tree:

sparse data

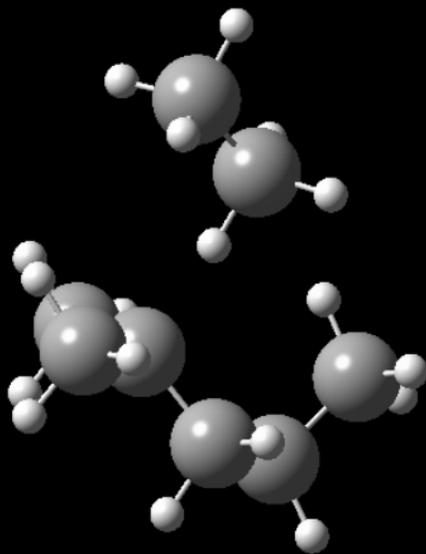


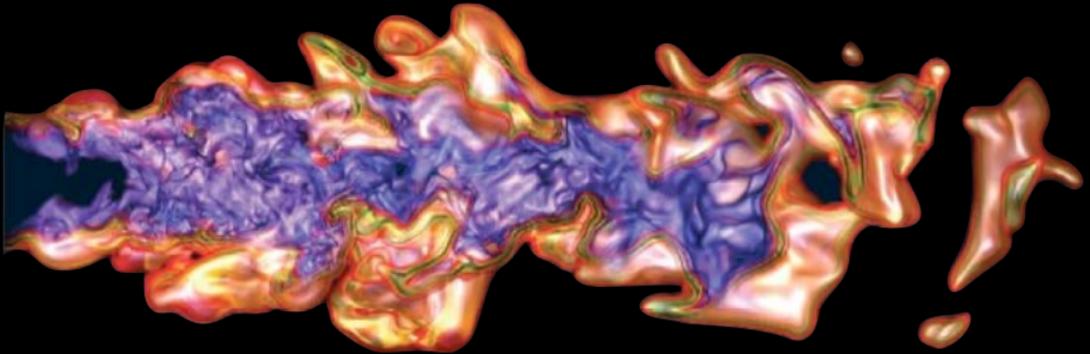
But what if there are no data? How can we make it?

Transition State calculations can provide missing data



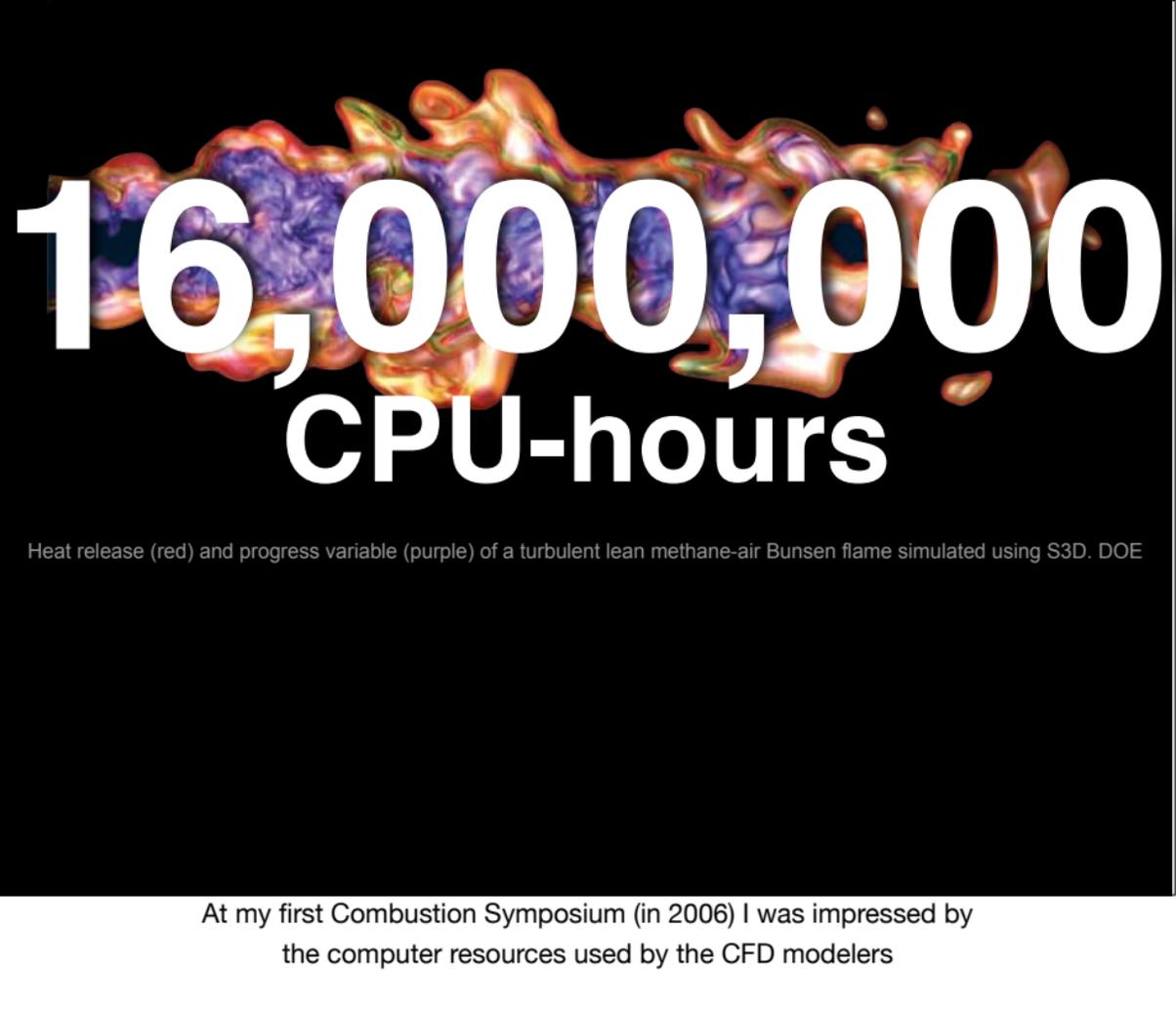
Transition State calculations can provide missing data





Heat release (red) and progress variable (purple) of a turbulent lean methane-air Bunsen flame simulated using S3D. DOE

At my first Combustion Symposium (in 2006) I was impressed by the computer resources used by the CFD modelers



16,000,000 CPU-hours

Heat release (red) and progress variable (purple) of a turbulent lean methane-air Bunsen flame simulated using S3D. DOE

At my first Combustion Symposium (in 2006) I was impressed by
the computer resources used by the CFD modelers



Assoc. Lib Director, ORNL Computing

This is one of the large federal computers they use.
Imagine what we could do if combustion *chemists* had 16M CPU hours!...

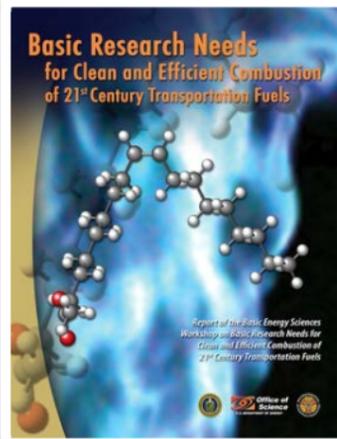


We'd need 16,000 graduate students to set up the calculations!



We'd need 16,000 graduate students to set up the calculations!

Automatic TS searches remain an important energy research goal

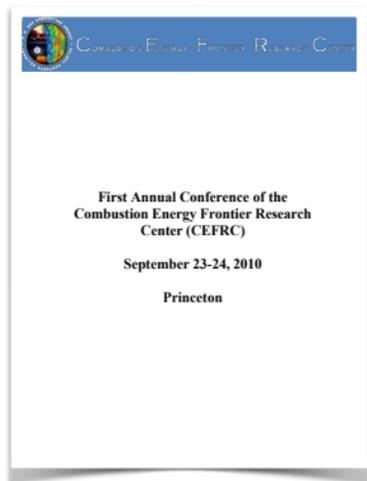


“An accurate description of the often intricate mechanisms of large-molecule reactions requires a characterization of all relevant transition states... Development of automatic means to search for chemically relevant configurations is the computational-kinetics equivalent of improved electronic structure methods.”

- Basic Research Needs for Clean and Efficient Combustion of 21st Century Transportation Fuels.
US Dept of Energy (2006)

I skipped this slide, because the previous talk had made the motivation clear.

Automatic TS searches remain an important energy research goal



“...transformation from by-hand calculations of single reactions to automated calculations of millions of reactions would be a game-changer for the field of chemistry, and would be a good ‘Grand Challenge’ target...”

- Combustion Energy Frontier Research Center (2010)

I skipped this slide too.

**Can you predict TS geometries
from molecular groups alone?**

(this would be great)

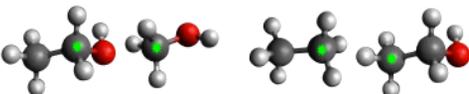
Can you predict TS geometries from molecular groups alone?

Molecule

Radical

Length of bond being broken,
at TS for Hydrogen abstraction

Can you predict TS geometries from molecular groups alone?



	1.149	1.155	1.168	1.179
	1.272	1.286	1.309	1.314
	1.278	1.295		1.355
	1.280	1.306	1.362	1.369

in Å with M06-2X/6-31+G(d,p)

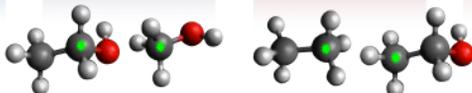
Can you identify the bond lengths in the following molecules?

1.168

1.272

1.345

1.372



1.149

1.155

1.168

1.179



1.272

1.286

1.309

1.314



1.278

1.295

1.355



1.280

1.306

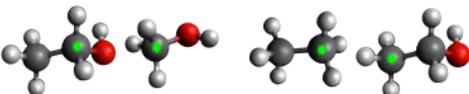
1.362

1.369

in Å with M06-2X/6-31+G(d,p)

Guess which number goes in the blank?

You can predict TS geometries from molecular groups alone!



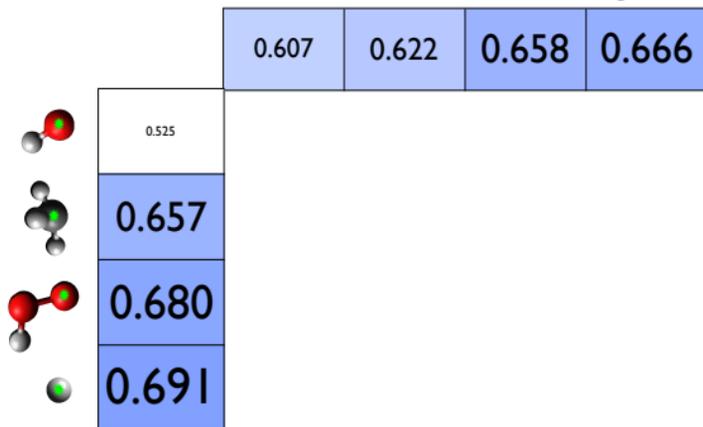
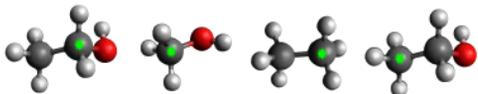
	1.149	1.155	1.168	1.179
	1.272	1.286	1.309	1.314
	1.278	1.295	1.345	1.355
	1.280	1.306	1.362	1.369

in Å with M06-2X/6-31+G(d,p)

So you can predict distances! But...

- (1) I gave you 15 numbers and you gave me one; and
- (2) you only gave me a distance, not a 3D geometry.

Add the two numbers to get any distance within 1 pm



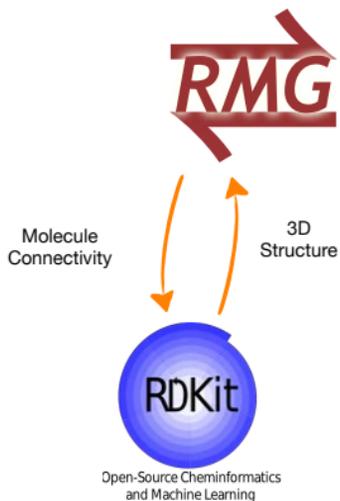
in Å with M06-2X/6-31+G(d,p)

First challenge: getting more data out than you put in.

Now, I give you 8 numbers and you can give me 16!

Arrange these groups in a tree as mentioned before, and job done.

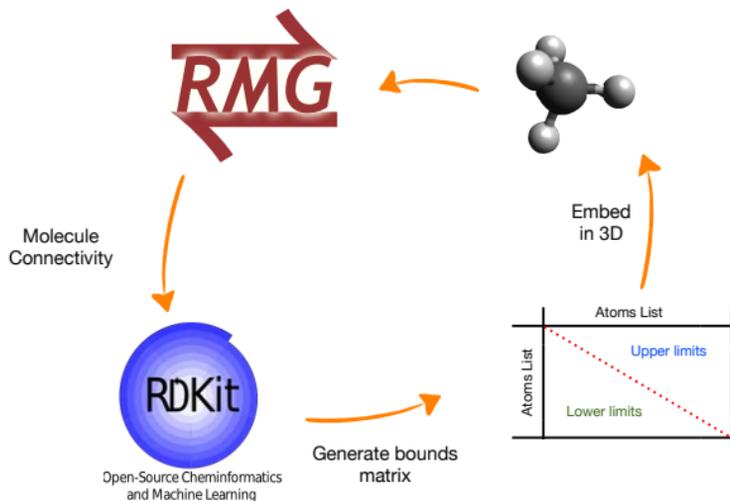
Use distance geometry to position atoms



Second challenge: getting from a distance to a 3D structure.

We use the distance-geometry algorithms in RDKit.

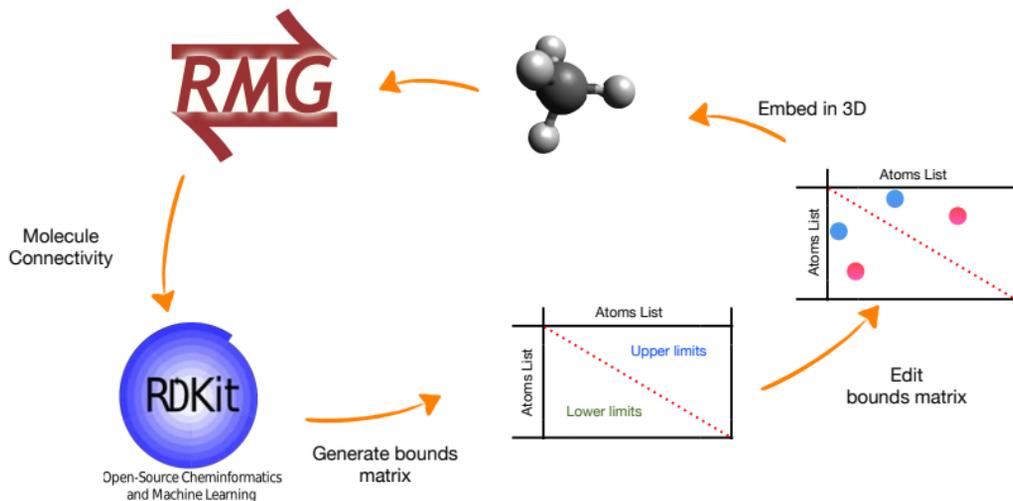
Use distance geometry to position atoms



Second challenge: getting from a distance to a 3D structure.

We use the distance-geometry algorithms in RDKit.

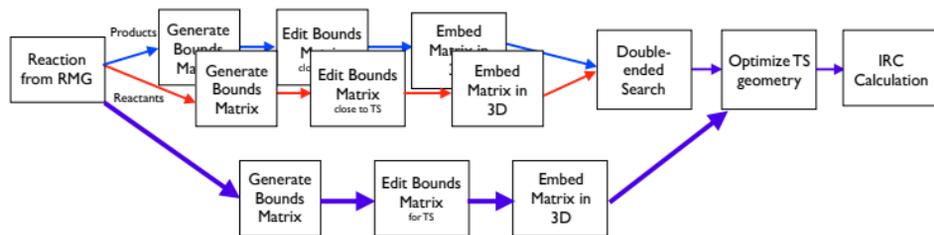
Use distance geometry to position atoms



Second challenge: getting from a distance to a 3D structure.

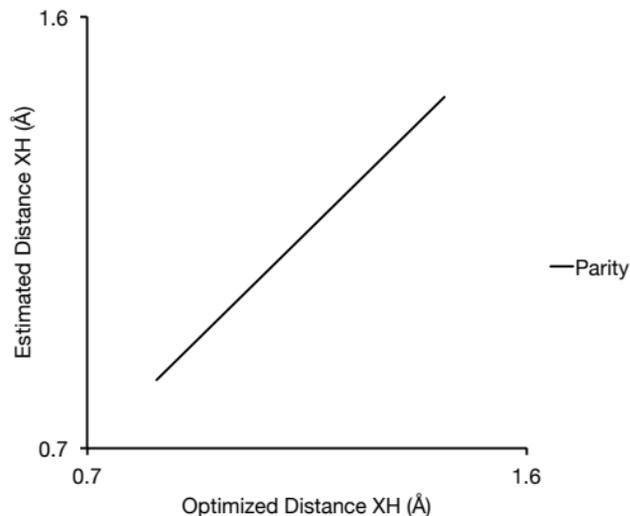
We use the distance-geometry algorithms in RDKit.

Estimate geometry directly via group additive distance estimates



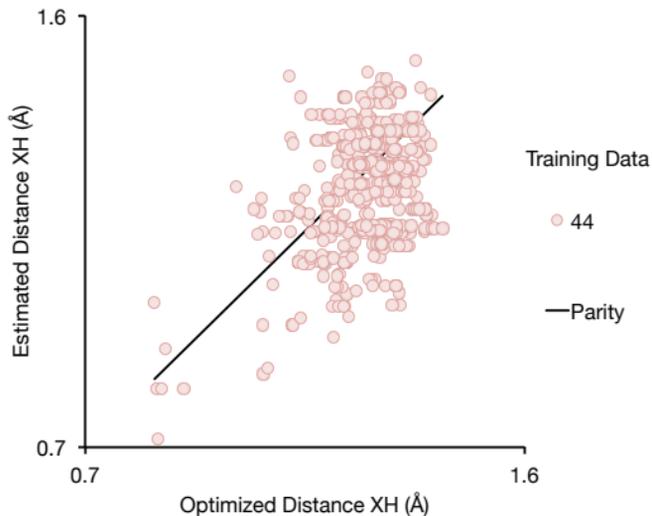
- Database arranged in tree structure as for kinetics
- Trained on successfully optimized transition states
- Direct guess much faster than double ended search
- Success depends on training data

Parity plots of Estimated against Optimized X-H distances at the transition state

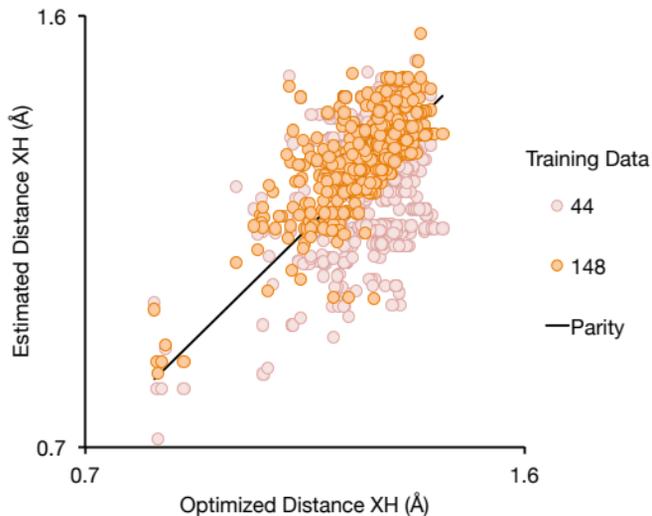


We'll see how well we predict the distances as we re-train the groups on more and more data.

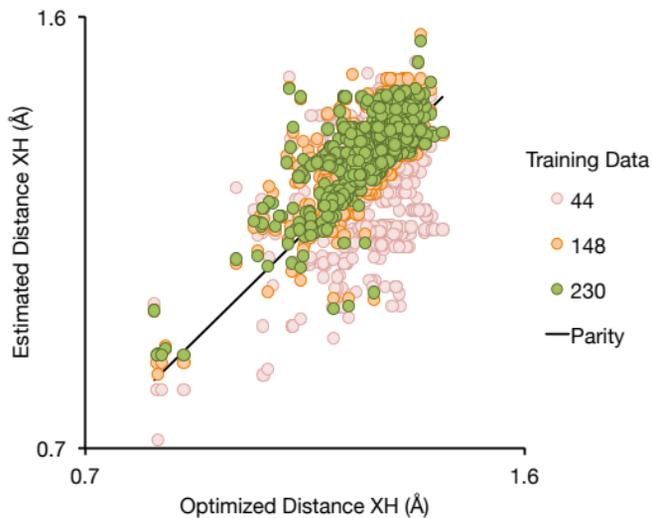
Trained on 44 reactions,
estimates not so great...



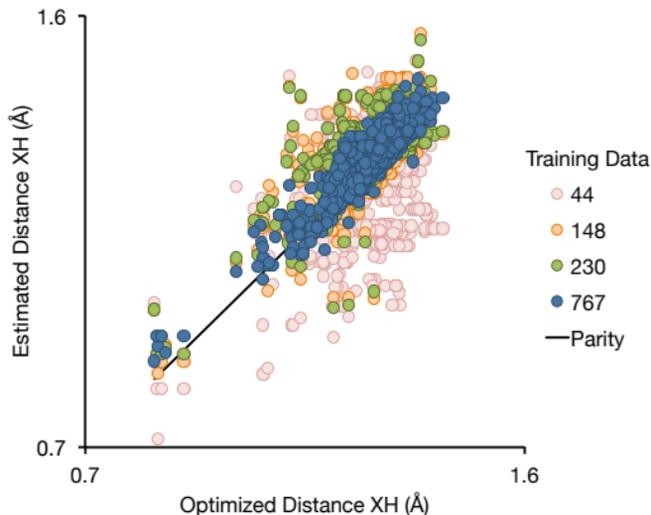
Trained on 148 reactions, estimates improve...



Trained on 230 reactions, estimates improve...



Trained on 767 reactions,
estimates are very good.



- See Pierre Bhoorasingh's poster on Wednesday



Northeastern University
 College of Engineering

20th International Symposium on Combustion
 August 3-8, 2014

Group contributions for determining transition state geometries

Pierre L. Bhoorasingh and Richard H. West
 Department of Chemical Engineering, Northeastern University, Boston, MA
www.northeastern.edu/chemeng

MOLECULAR GEOMETRIES REQUIRES DETAILED RESULTS FOR EACH MOLECULE

BUT WE CAN'T PAINT REACTION SCHEMES RED

WE REQUIRE A FASTER METHOD TO OBTAIN STABLE POINT GEOMETRIES

NO SPEED-UP AS YET FOR TRANSITION STATE THEORY CALCULATIONS

ACTUATING MOLECULAR GEOMETRIES INTO A MODEL

Inspiration

Group additive methods can predict many molecular properties.

Powered by Benson to estimate thermo-dynamic properties of a molecule [1], contributions from each molecular group are summed. This method has been extended to other applications, and we use it here to estimate transition state geometries.

Implementation

Distances between reacting atoms trend with neighbouring groups.

Group	1.200	1.214	1.216	1.222	1.248
1.200	1.395	1.395	1.395	1.395	1.395
1.214	1.395	1.395	1.395	1.395	1.395
1.216	1.395	1.395	1.395	1.395	1.395
1.222	1.395	1.395	1.395	1.395	1.395
1.248	1.429	1.432	1.432	1.432	1.432

Chemical groups of primary, secondary, tertiary, quaternary

All distances at the transition state are similar to distances for stable species, except those in the reaction center. The trends shown involve group additive estimation for the distances between reactive atoms. If abstraction transition states found at M06-2X/6-31+G(d,p).

Group values stored in a hierarchical tree.

```

graph TD
    Root[Root] --> H[H]
    Root --> C[C]
    Root --> N[N]
    Root --> O[O]
    Root --> F[F]
    Root --> Cl[Cl]
    Root --> Br[Br]
    Root --> I[I]
    Root --> S[S]
    Root --> Se[Se]
    Root --> Te[Te]
    Root --> Zn[Zn]
    Root --> Cd[Cd]
    Root --> Hg[Hg]
    Root --> Cu[Cu]
    Root --> Ag[Ag]
    Root --> Au[Au]
    Root --> Pt[Pt]
    Root --> Pd[Pd]
    Root --> Ni[Ni]
    Root --> Co[Co]
    Root --> Fe[Fe]
    Root --> Mn[Mn]
    Root --> Cr[Cr]
    Root --> V[V]
    Root --> Ti[Ti]
    Root --> Zr[Zr]
    Root --> Hf[Hf]
    Root --> Ta[Ta]
    Root --> Nb[Nb]
    Root --> Mo[Mo]
    Root --> W[W]
    Root --> Re[Re]
    Root --> Os[Os]
    Root --> Ir[Ir]
    Root --> Rh[Rh]
    Root --> Ru[Ru]
    Root --> Rhodium[Rhodium]
    Root --> Ruthenium[Ruthenium]
    Root --> Rhenium[Rhenium]
    Root --> Osmium[Osmium]
    Root --> Iridium[Iridium]
    Root --> Platinum[Platinum]
    Root --> Palladium[Palladium]
    Root --> Nickel[Nickel]
    Root --> Cobalt[Cobalt]
    Root --> Iron[Iron]
    Root --> Manganese[Manganese]
    Root --> Chromium[Chromium]
    Root --> Vanadium[Vanadium]
    Root --> Titanium[Titanium]
    Root --> Zirconium[Zirconium]
    Root --> Hafnium[Hafnium]
    Root --> Tantalum[Tantalum]
    Root --> Niobium[Niobium]
    Root --> Molybdenum[Molybdenum]
    Root --> Tungsten[Tungsten]
    Root --> Rhenium[Rhenium]
    Root --> Osmium[Osmium]
    Root --> Iridium[Iridium]
    Root --> Platinum[Platinum]
    Root --> Gold[Gold]
    Root --> Silver[Silver]
    Root --> Copper[Copper]
    Root --> Zinc[Zinc]
    Root --> Cadmium[Cadmium]
    Root --> Mercury[Mercury]
    Root --> Selenium[Selenium]
    Root --> Tellurium[Tellurium]
    Root --> Lead[Lead]
    Root --> B[B]
    Root --> Al[Al]
    Root --> Ga[Ga]
    Root --> In[In]
    Root --> Tl[Tl]
    Root --> Sc[Sc]
    Root --> Y[Y]
    Root --> La[La]
    Root --> Ce[Ce]
    Root --> Pr[Pr]
    Root --> Nd[Nd]
    Root --> Pm[Pm]
    Root --> Sm[Sm]
    Root --> Eu[Eu]
    Root --> Gd[Gd]
    Root --> Tb[Tb]
    Root --> Dy[Dy]
    Root --> Ho[Ho]
    Root --> Er[Er]
    Root --> Tm[Tm]
    Root --> Yb[Yb]
    Root --> Lu[Lu]
    Root --> Be[Be]
    Root --> Mg[Mg]
    Root --> Ca[Ca]
    Root --> Sr[Sr]
    Root --> Ba[Ba]
    Root --> Ra[Ra]
    Root --> Li[Li]
    Root --> Na[Na]
    Root --> K[K]
    Root --> Rb[Rb]
    Root --> Cs[Cs]
    Root --> Fr[Fr]
    Root --> H2[H2]
    Root --> H2O[H2O]
    Root --> H2O2[H2O2]
    Root --> H2O3[H2O3]
    Root --> H2O4[H2O4]
    Root --> H2O5[H2O5]
    Root --> H2O6[H2O6]
    Root --> H2O7[H2O7]
    Root --> H2O8[H2O8]
    Root --> H2O9[H2O9]
    Root --> H2O10[H2O10]
    Root --> H2O11[H2O11]
    Root --> H2O12[H2O12]
    Root --> H2O13[H2O13]
    Root --> H2O14[H2O14]
    Root --> H2O15[H2O15]
    Root --> H2O16[H2O16]
    Root --> H2O17[H2O17]
    Root --> H2O18[H2O18]
    Root --> H2O19[H2O19]
    Root --> H2O20[H2O20]
    Root --> H2O21[H2O21]
    Root --> H2O22[H2O22]
    Root --> H2O23[H2O23]
    Root --> H2O24[H2O24]
    Root --> H2O25[H2O25]
    Root --> H2O26[H2O26]
    Root --> H2O27[H2O27]
    Root --> H2O28[H2O28]
    Root --> H2O29[H2O29]
    Root --> H2O30[H2O30]
    Root --> H2O31[H2O31]
    Root --> H2O32[H2O32]
    Root --> H2O33[H2O33]
    Root --> H2O34[H2O34]
    Root --> H2O35[H2O35]
    Root --> H2O36[H2O36]
    Root --> H2O37[H2O37]
    Root --> H2O38[H2O38]
    Root --> H2O39[H2O39]
    Root --> H2O40[H2O40]
    Root --> H2O41[H2O41]
    Root --> H2O42[H2O42]
    Root --> H2O43[H2O43]
    Root --> H2O44[H2O44]
    Root --> H2O45[H2O45]
    Root --> H2O46[H2O46]
    Root --> H2O47[H2O47]
    Root --> H2O48[H2O48]
    Root --> H2O49[H2O49]
    Root --> H2O50[H2O50]
    Root --> H2O51[H2O51]
    Root --> H2O52[H2O52]
    Root --> H2O53[H2O53]
    Root --> H2O54[H2O54]
    Root --> H2O55[H2O55]
    Root --> H2O56[H2O56]
    Root --> H2O57[H2O57]
    Root --> H2O58[H2O58]
    Root --> H2O59[H2O59]
    Root --> H2O60[H2O60]
    Root --> H2O61[H2O61]
    Root --> H2O62[H2O62]
    Root --> H2O63[H2O63]
    Root --> H2O64[H2O64]
    Root --> H2O65[H2O65]
    Root --> H2O66[H2O66]
    Root --> H2O67[H2O67]
    Root --> H2O68[H2O68]
    Root --> H2O69[H2O69]
    Root --> H2O70[H2O70]
    Root --> H2O71[H2O71]
    Root --> H2O72[H2O72]
    Root --> H2O73[H2O73]
    Root --> H2O74[H2O74]
    Root --> H2O75[H2O75]
    Root --> H2O76[H2O76]
    Root --> H2O77[H2O77]
    Root --> H2O78[H2O78]
    Root --> H2O79[H2O79]
    Root --> H2O80[H2O80]
    Root --> H2O81[H2O81]
    Root --> H2O82[H2O82]
    Root --> H2O83[H2O83]
    Root --> H2O84[H2O84]
    Root --> H2O85[H2O85]
    Root --> H2O86[H2O86]
    Root --> H2O87[H2O87]
    Root --> H2O88[H2O88]
    Root --> H2O89[H2O89]
    Root --> H2O90[H2O90]
    Root --> H2O91[H2O91]
    Root --> H2O92[H2O92]
    Root --> H2O93[H2O93]
    Root --> H2O94[H2O94]
    Root --> H2O95[H2O95]
    Root --> H2O96[H2O96]
    Root --> H2O97[H2O97]
    Root --> H2O98[H2O98]
    Root --> H2O99[H2O99]
    Root --> H2O100[H2O100]
  
```

Distance = $1.204 + 0.224 + 0.044 + 0.114 = 1.586$

A database of known distances was created and used to populate hierarchical molecular group trees, one for each distance d for H-Abstraction. The reaction above matches the C(d) and O(d) groups, so each distance is calculated by summing the head node values and the group connectors. If a group is unrefined, its parent node value is used, and once optimized it can be applied to the training set to improve the groups. This continual improvement of group values is an important feature of our method.

MOLECULAR GEOMETRIES ARE REQUIRED

THESE GROUP CONTRIBUTIONS...

...AND USING GEOMETRIES IN COMPUTING POWER...

...IMPROVING MODELS...

...AND USING GEOMETRIES IN COMPUTING POWER...

...IMPROVING MODELS...

...AND USING GEOMETRIES IN COMPUTING POWER...

...IMPROVING MODELS...

CanTherm

Chemical Thermodynamics Package

AutoTST

RMG

ROK

Interpreting Existing Kinetic Models



Grant No. 1403171



Northeastern.edu/comocheng

Now for the second project, that is also built on RMG-Py.



Available online at www.sciencedirect.com

 ScienceDirect

Proceedings of the Combustion Institute 31 (2007) 125–140

Proceedings
of the
Combustion
Institute

www.elsevier.com/locate/proci

Transforming data into knowledge—Process Informatics for combustion chemistry

Michael Frenklach *

Department of Mechanical Engineering, University of California, and Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Abstract

The present frontier of combustion chemistry is the development of *predictive* reaction models, namely, chemical kinetics models capable of accurate numerical predictions with quantifiable uncertainties. While the usual factors like deficient knowledge of reaction pathways and insufficient accuracy of individual measurements and/or theoretical calculations impede progress, the key obstacle is the inconsistency of accumulating data and proliferating reaction mechanisms. Process Informatics introduces a new paradigm. It relies on three major components: proper organization of scientific data, availability of scientific tools for analysis and processing of these data, and engagement of the entire scientific community in the data collection and analysis. The proper infrastructure will enable a new form of scientific method by considering the entire content of information available, assessing and assuring mutual scientific consistency of the data, rigorously assessing data uncertainty, identifying problems with the available data, evaluating model predictability, suggesting new experimental and theoretical work with the highest possible impact, reaching community consensus, and merging the assembled data into new knowledge.

© 2006 The Combustion Institute. Published by Elsevier Inc. All rights reserved.

Keywords: Kinetics

Again, we go back to my first Combustion Symposium in 2006, where I was introduced to the idea of Process Informatics / Data Collaboration.

“the key obstacle is the inconsistency of accumulating data and proliferating reaction mechanisms.

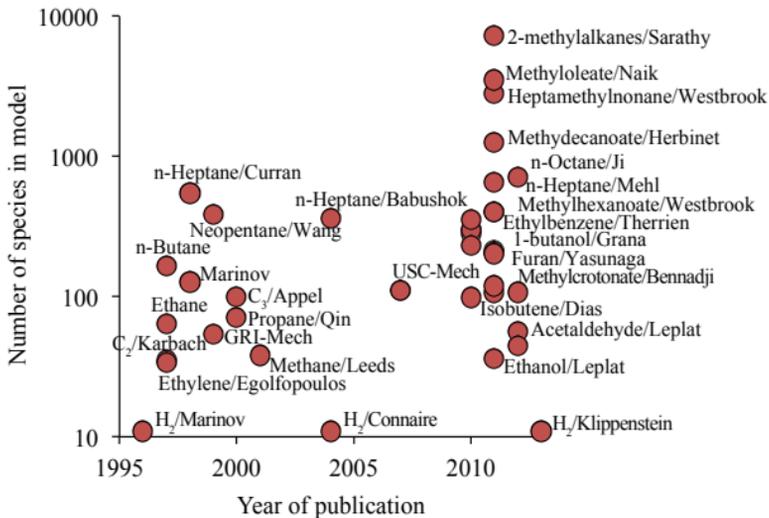
Process Informatics introduces a new paradigm. It relies on three major components:

- proper organization of scientific data,
- availability of scientific tools for analysis and processing of these data,
- and engagement of the entire scientific community in the data collection and analysis.”

M. Frenklach, *Proc. Combust. Inst.* 31 (2006)

assumed the entire scientific community would go home and properly organize their scientific data.

Model Complexity is Increasing



So is everything published since 2006 in PrImE format?

No. It's mostly in CHEMKIN format.

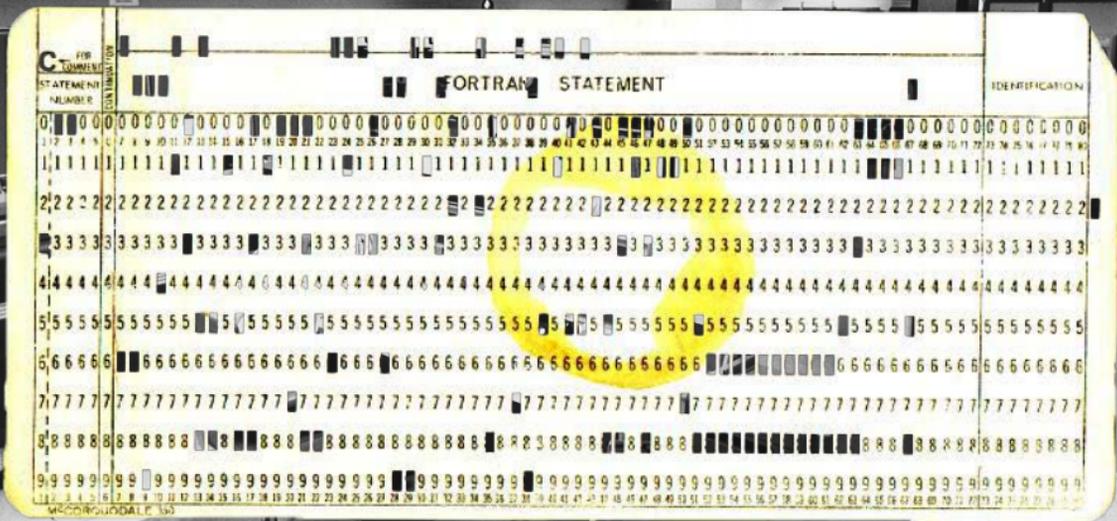
(This chart is very incomplete random sampling)



And again, here is a photo of a large federal computer. This time at NASA



This is the computer model Gordon and McBride used when they devised the NASA polynomial form for thermochemistry (now used in CHEMKIN). It was designed to fit on 80-column punch-cards.



is the computer model Gordon and McBride used when they devised the NASA polynomial form for thermochemistry (now used in CHEMKIN). It was designed to fit on 80-column punch-cards.

APPEND D
THERMO DATA (FORMAT AND LISTING)

The order and format of the input data cards in this appendix are given in the following table:

Card order	Contents	Format	Card column
1	THERMO	3A4	1 to 6
2	Temperature ranges for 2 sets of coefficients: (lower, T, common T, and highest T)	2F10.3	1 to 30
3	Species name	3A4	1 to 12
	Date	2A2	13 to 24
	Atomic symbols and formula	4IA2.F3.0	25 to 44
	Phase of species (S, L, or G for solid, liquid, or gas, respectively)	A1	45
	Temperature range	2F10.3	46 to 65
4	Integer 1	5CE15.6	1 to 75
	Coefficients a_i (i = 1 to 5) in equations (90) to (92) (for upper temperature interval)	15	80
	Integer 2	5CE15.6	1 to 75
5	Coefficients in equations (90) to (92) (a_0 , a_1 for upper temperature interval, and a_1 , a_2 , and a_3 for lower)	15	80
	Integer 3	5CE15.6	1 to 75
6	Coefficients in equations (90) to (92) (a_0 , a_1 , a_2 , a_3 for lower temperature interval)	15	80
	Integer 4	5CE15.6	1 to 75
(a)	Repeat cards numbered 4 to 6 in cc 80 for each species	10	80
(Final card)	END (Indicates end of thermodynamic data)	3A4	1 to 3

^aGasous species and condensed species with only one condensed phase can be in any order. However, the sets for two or more condensed phases of the same species must be adjacent. If there are more than two condensed phases of a species, their sets must be either in increasing or decreasing order according to their temperature intervals.

ORIGINAL PAGE IS
OF POOR QUALITY

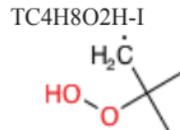
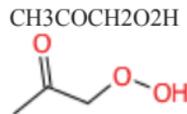
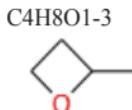
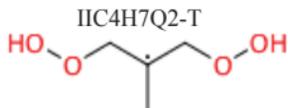
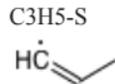
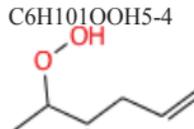
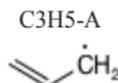
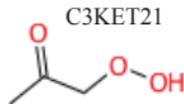
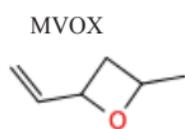
is the computer model Gordon and McBride used when they devised the NASA polynomial form for thermochemistry (now used in CHEMKIN). It was designed to fit on 80-column punch-cards.

NASA (Chemkin) format is very dense – not much room for species identifiers

Species Name Chemical Formula Parameters for H(T), S(T)

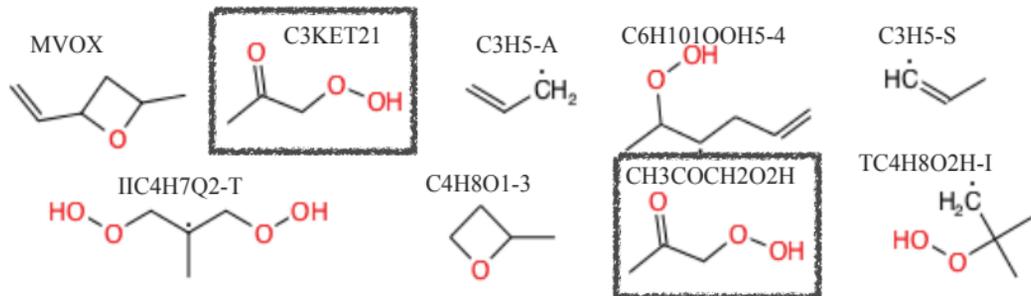
ch3chocho	11/14/95	thermc	3h	5o	2	0g	300.000	5000.000	1395.000	21
	1.27827539e+01	1.21529347e-02	-4.21228103e-06				6.59844007e-10	-3.85228684e-14		2
	-2.15946846e+04	-3.94512036e+01	5.06096897e-01				4.26121702e-02	-3.37628379e-05		3
	1.39422914e-08	-2.35317319e-12	-1.74952136e+04				2.58742916e+01			4
ch3chcho	12/ 1/99	thermc	3h	5o	1	0g	300.000	5000.000	1385.000	21
	1.02647528E+01	1.17295061E-02	-4.06267561E-06				6.36086751E-10	-3.71220190E-14		2
	-7.25910488E+03	-2.91646638E+01	1.54196914E+00				3.03297391E-02	-1.88922690E-05		3
	5.92263358E-09	-7.53762316E-13	-4.01761098E+03				1.83707011E+01			4
sc3h5cho	11/15/95	therc	4h	6o	1	0g	300.000	5000.000	1390.000	21
	1.31695904e+01	1.42484434e-02	-4.90843998e-06				7.65789041e-10	-4.45834896e-14		2
	-2.04032613e+04	-4.43673205e+01	4.35795171e-01				4.48719314e-02	-3.36582931e-05		3
	1.33066870e-08	-2.17839128e-12	-1.60394651e+04				2.37597452e+01			4
sc3h5co	11/15/95	thermc	4h	5o	1	0g	300.000	5000.000	1392.000	21
	1.25514754e+01	1.22521948e-02	-4.22382101e-06				6.59184896e-10	-3.83818826e-14		2
	-4.25349795e+03	-4.02864145e+01	1.74191343e+00				3.97229536e-02	-3.20061901e-05		3
	1.38227925e-08	-2.46272017e-12	-6.64428100e+02				1.70762023e+01			4
ic3h6cho	2/22/96	thermc	4h	7o	1	0g	300.000	5000.000	1390.000	31
	1.33102250e+01	1.62097959e-02	-5.57575891e-06				8.69003718e-10	-5.05554202e-14		2
	-7.62177931e+03	-4.25050854e+01	5.21481767e-01				4.43114357e-02	-2.86617314e-05		3
	9.30319894e-09	-1.20761563e-12	-2.99677086e+03				2.68182130e+01			4
c3h6cho-1	9/27/95	thermc	4h	7o	1	0g	300.000	5000.000	1379.000	31
	1.20222054e+01	1.62418273e-02	-5.54288124e-06				8.50722695e-10	-4.08450726e-14		2

Space constraints led to “creative” naming schemes



Notice hydroperoxypropan-2-one has two different names, often in the same mechanism file!

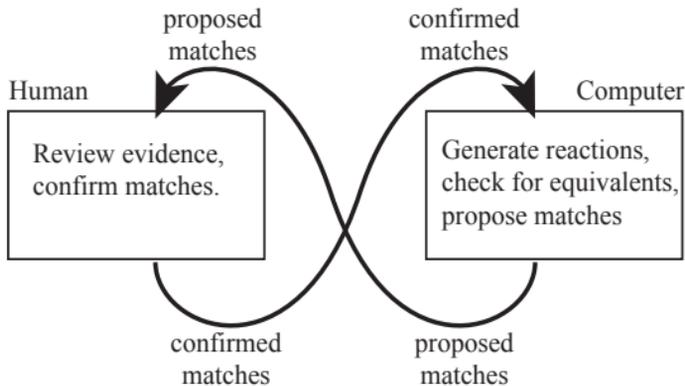
Space constraints led to “creative” naming schemes



- C3KET21 is generated from alkyl peroxy radical isomerization pathway
- CH3COCH2O2H is generated from low-temperature oxidation of acetone

Notice hydroperoxypropan-2-one has two different names, often in the same mechanism file!

Human-Computer team can identify species more quickly

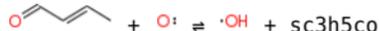
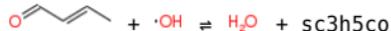
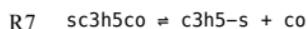
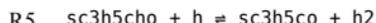
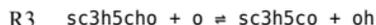
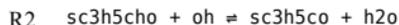
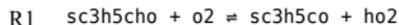


- Our new tool uses RMG to generate reactions to compare with the target model
- A human reviews the evidence and confirms matches.

Identify small molecules first

- Identify species with only one possible structure
 - CO_2
 - H_2O
 - C_3H_8
- Then species with "borrowed" thermochemistry
- Then boot-strap based on how these react

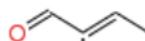
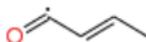
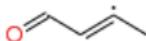
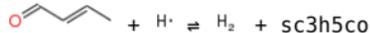
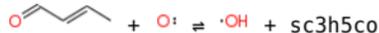
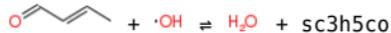
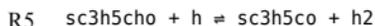
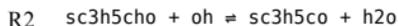
Identifying 'sc3h5co' from its reactions



The first 6 reactions are all H-abstractions giving 4 possibilities for sc3h5co.

The 7th reaction is beta-scission, also giving 4 possibilities.

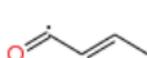
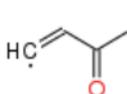
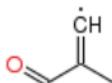
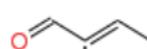
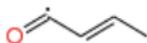
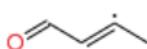
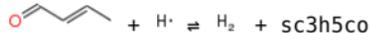
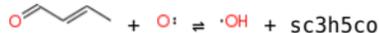
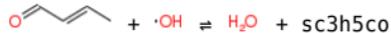
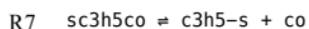
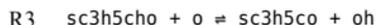
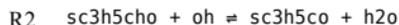
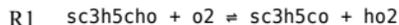
Identifying 'sc3h5co' from its reactions



The first 6 reactions are all H-abstractions giving 4 possibilities for sc3h5co.

The 7th reaction is beta-scission, also giving 4 possibilities.

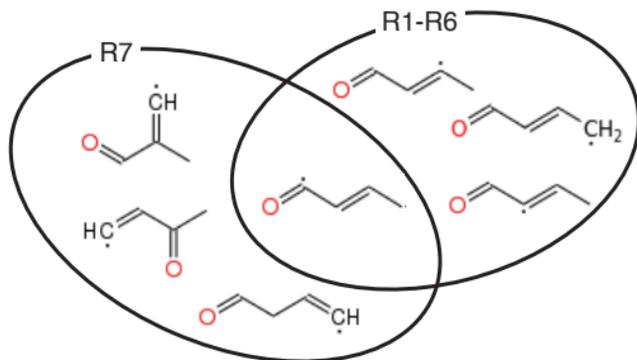
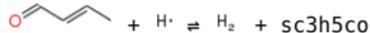
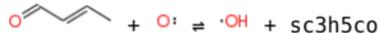
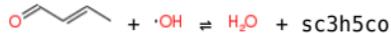
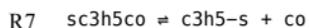
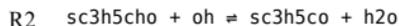
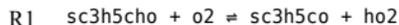
Identifying 'sc3h5co' from its reactions



The first 6 reactions are all H-abstractions giving 4 possibilities for sc3h5co.

The 7th reaction is beta-scission, also giving 4 possibilities.

Identifying 'sc3h5o' from its reactions



By considering all seven reactions, we can deduce the most likely isomer.

'sc3h5o' is but-2-enoyl



Interact with a Web UI

- Computers make mistakes, and humans can see patterns
- All matches must be approved by an operator using a web user interface

Name	Molecule	$\Delta H_f^\circ(298\text{K})$	Matching Thermo	
CH2O	$\text{H}_2\text{C}=\text{O}$	0.9 kJ/mol	GRI-Mech3.0	confirm
HCO	$\text{H}\dot{\text{C}}=\text{O}$	1.0 kJ/mol	-	confirm
HOCHO	$\text{O}=\text{C}-\text{OH}$	0.8 kJ/mol	-	confirm
C2H2	$\text{HC}\equiv\text{CH}$	-3.2 kJ/mol	-	confirm
CH2CO	$\text{O}=\text{C}=\text{CH}_2$	-3.0 kJ/mol	-	confirm



Two undergraduates this summer imported models from recent volumes of the journals *Combustion & Flame* and *Proceedings of the Combustion Institute*. More details on our poster.





Northeastern University



35th International Symposium on Combustion, August 3-8, 2014

Identification and Comparison of Detailed Chemical Mechanisms

Esma Goktekin, Anthony J. Silvi, Victor R. Lambert, Fariba Seyedzadeh Khanshan and Richard H. West
Department of Chemical Engineering, Northeastern University, Boston, MA

Each new user brings a new model.

Introduction to our mechanism importer and compare tool

Collection of Published Models

The importer uses two main sets of data to build the kinetic models. The first set of data is the thermochemical data associated with the chemical species within the model. The thermochemical data set must be formatted as NASA polynomial coefficients, or listed as NASA polynomial coefficients, or listed as NASA polynomial coefficients, or listed as NASA polynomial coefficients.

Species	Order	Temperature Range (K)	Reference
H ₂	7	200-1000	1
O ₂	7	200-1000	1
H ₂ O	7	200-1000	1
CO	7	200-1000	1
CO ₂	7	200-1000	1
H ₂ O ₂	7	200-1000	1
HO ₂	7	200-1000	1
OH	7	200-1000	1
H	7	200-1000	1
O	7	200-1000	1
NO	7	200-1000	1
NO ₂	7	200-1000	1
H ₂ CO	7	200-1000	1
H ₂ CO ₂	7	200-1000	1
H ₂ CO ₃	7	200-1000	1
H ₂ CO ₄	7	200-1000	1
H ₂ CO ₅	7	200-1000	1
H ₂ CO ₆	7	200-1000	1
H ₂ CO ₇	7	200-1000	1
H ₂ CO ₈	7	200-1000	1
H ₂ CO ₉	7	200-1000	1
H ₂ CO ₁₀	7	200-1000	1
H ₂ CO ₁₁	7	200-1000	1
H ₂ CO ₁₂	7	200-1000	1
H ₂ CO ₁₃	7	200-1000	1
H ₂ CO ₁₄	7	200-1000	1
H ₂ CO ₁₅	7	200-1000	1
H ₂ CO ₁₆	7	200-1000	1
H ₂ CO ₁₇	7	200-1000	1
H ₂ CO ₁₈	7	200-1000	1
H ₂ CO ₁₉	7	200-1000	1
H ₂ CO ₂₀	7	200-1000	1
H ₂ CO ₂₁	7	200-1000	1
H ₂ CO ₂₂	7	200-1000	1
H ₂ CO ₂₃	7	200-1000	1
H ₂ CO ₂₄	7	200-1000	1
H ₂ CO ₂₅	7	200-1000	1
H ₂ CO ₂₆	7	200-1000	1
H ₂ CO ₂₇	7	200-1000	1
H ₂ CO ₂₈	7	200-1000	1
H ₂ CO ₂₉	7	200-1000	1
H ₂ CO ₃₀	7	200-1000	1
H ₂ CO ₃₁	7	200-1000	1
H ₂ CO ₃₂	7	200-1000	1
H ₂ CO ₃₃	7	200-1000	1
H ₂ CO ₃₄	7	200-1000	1
H ₂ CO ₃₅	7	200-1000	1
H ₂ CO ₃₆	7	200-1000	1
H ₂ CO ₃₇	7	200-1000	1
H ₂ CO ₃₈	7	200-1000	1
H ₂ CO ₃₉	7	200-1000	1
H ₂ CO ₄₀	7	200-1000	1
H ₂ CO ₄₁	7	200-1000	1
H ₂ CO ₄₂	7	200-1000	1
H ₂ CO ₄₃	7	200-1000	1
H ₂ CO ₄₄	7	200-1000	1
H ₂ CO ₄₅	7	200-1000	1
H ₂ CO ₄₆	7	200-1000	1
H ₂ CO ₄₇	7	200-1000	1
H ₂ CO ₄₈	7	200-1000	1
H ₂ CO ₄₉	7	200-1000	1
H ₂ CO ₅₀	7	200-1000	1
H ₂ CO ₅₁	7	200-1000	1
H ₂ CO ₅₂	7	200-1000	1
H ₂ CO ₅₃	7	200-1000	1
H ₂ CO ₅₄	7	200-1000	1
H ₂ CO ₅₅	7	200-1000	1
H ₂ CO ₅₆	7	200-1000	1
H ₂ CO ₅₇	7	200-1000	1
H ₂ CO ₅₈	7	200-1000	1
H ₂ CO ₅₉	7	200-1000	1
H ₂ CO ₆₀	7	200-1000	1
H ₂ CO ₆₁	7	200-1000	1
H ₂ CO ₆₂	7	200-1000	1
H ₂ CO ₆₃	7	200-1000	1
H ₂ CO ₆₄	7	200-1000	1
H ₂ CO ₆₅	7	200-1000	1
H ₂ CO ₆₆	7	200-1000	1
H ₂ CO ₆₇	7	200-1000	1
H ₂ CO ₆₈	7	200-1000	1
H ₂ CO ₆₉	7	200-1000	1
H ₂ CO ₇₀	7	200-1000	1
H ₂ CO ₇₁	7	200-1000	1
H ₂ CO ₇₂	7	200-1000	1
H ₂ CO ₇₃	7	200-1000	1
H ₂ CO ₇₄	7	200-1000	1
H ₂ CO ₇₅	7	200-1000	1
H ₂ CO ₇₆	7	200-1000	1
H ₂ CO ₇₇	7	200-1000	1
H ₂ CO ₇₈	7	200-1000	1
H ₂ CO ₇₉	7	200-1000	1
H ₂ CO ₈₀	7	200-1000	1
H ₂ CO ₈₁	7	200-1000	1
H ₂ CO ₈₂	7	200-1000	1
H ₂ CO ₈₃	7	200-1000	1
H ₂ CO ₈₄	7	200-1000	1
H ₂ CO ₈₅	7	200-1000	1
H ₂ CO ₈₆	7	200-1000	1
H ₂ CO ₈₇	7	200-1000	1
H ₂ CO ₈₈	7	200-1000	1
H ₂ CO ₈₉	7	200-1000	1
H ₂ CO ₉₀	7	200-1000	1
H ₂ CO ₉₁	7	200-1000	1
H ₂ CO ₉₂	7	200-1000	1
H ₂ CO ₉₃	7	200-1000	1
H ₂ CO ₉₄	7	200-1000	1
H ₂ CO ₉₅	7	200-1000	1
H ₂ CO ₉₆	7	200-1000	1
H ₂ CO ₉₇	7	200-1000	1
H ₂ CO ₉₈	7	200-1000	1
H ₂ CO ₉₉	7	200-1000	1
H ₂ CO ₁₀₀	7	200-1000	1

The importer also reads the reaction mechanism file data. The mechanism data must be supplied in the CHEMKIN format.

Debug Parameter Files for Import

Many published papers do not supply full thermochemical data for all species. The importer uses the associated thermochemical data to fill in the missing data in an uninterpretable format (e.g., PCP float). Some papers include the correct kinetic and thermochemical parameters, but contain mistakes in their large data files.

Some common errors include:

- Duplicate Thermochemical Entries
- Missing Thermochemical Entries
- Missing NASA polynomial fits
- Incorrect CHEMKIN formatting and spacing
- Excessive purification guesses, commas, spaces

Run Importer and Select Species

The importer reads in the thermochemical data and reaction mechanisms to build the kinetic model. The tool searches previously imported models to generate proposed species.

After the importer generates proposed matches, the user needs to confirm the match. The proposed match is displayed in an easy-to-use interface with the reaction mechanism and an entropy estimation. After choosing matches, the importer imports the model to generate new matches.

Compare Imported Models

Another tool is used to compare the models. The Compare tool reads in the combined species, thermochemical data, and reaction rates at 500K, 1000K, and 1500K of all imported models. The Compare generates thermochemical data, reaction rates, and standard deviation for each species and reaction.

Analysis of the 34th International Symposium on Combustion

Progress Importing Last Year's Models

Of the papers submitted to Proceedings of the Combustion Institute Volume 34, only a fraction are able to be imported. 12 papers are marked for comparison, their progress is shown below. Currently, 88% of all the species have been identified and added in their respective papers. The progress is shown in the bar chart below. As more user queries are added to the importer, they will become easier to identify.

Species	Identified	Not Identified
H ₂	100%	0%
O ₂	100%	0%
H ₂ O	100%	0%
CO	100%	0%
CO ₂	100%	0%
H ₂ O ₂	100%	0%
HO ₂	100%	0%
OH	100%	0%
H	100%	0%
O	100%	0%
NO	100%	0%
NO ₂	100%	0%
H ₂ CO	100%	0%
H ₂ CO ₂	100%	0%
H ₂ CO ₃	100%	0%
H ₂ CO ₄	100%	0%
H ₂ CO ₅	100%	0%
H ₂ CO ₆	100%	0%
H ₂ CO ₇	100%	0%
H ₂ CO ₈	100%	0%
H ₂ CO ₉	100%	0%
H ₂ CO ₁₀	100%	0%
H ₂ CO ₁₁	100%	0%
H ₂ CO ₁₂	100%	0%
H ₂ CO ₁₃	100%	0%
H ₂ CO ₁₄	100%	0%
H ₂ CO ₁₅	100%	0%
H ₂ CO ₁₆	100%	0%
H ₂ CO ₁₇	100%	0%
H ₂ CO ₁₈	100%	0%
H ₂ CO ₁₉	100%	0%
H ₂ CO ₂₀	100%	0%
H ₂ CO ₂₁	100%	0%
H ₂ CO ₂₂	100%	0%
H ₂ CO ₂₃	100%	0%
H ₂ CO ₂₄	100%	0%
H ₂ CO ₂₅	100%	0%
H ₂ CO ₂₆	100%	0%
H ₂ CO ₂₇	100%	0%
H ₂ CO ₂₈	100%	0%
H ₂ CO ₂₉	100%	0%
H ₂ CO ₃₀	100%	0%
H ₂ CO ₃₁	100%	0%
H ₂ CO ₃₂	100%	0%
H ₂ CO ₃₃	100%	0%
H ₂ CO ₃₄	100%	0%
H ₂ CO ₃₅	100%	0%
H ₂ CO ₃₆	100%	0%
H ₂ CO ₃₇	100%	0%
H ₂ CO ₃₈	100%	0%
H ₂ CO ₃₉	100%	0%
H ₂ CO ₄₀	100%	0%
H ₂ CO ₄₁	100%	0%
H ₂ CO ₄₂	100%	0%
H ₂ CO ₄₃	100%	0%
H ₂ CO ₄₄	100%	0%
H ₂ CO ₄₅	100%	0%
H ₂ CO ₄₆	100%	0%
H ₂ CO ₄₇	100%	0%
H ₂ CO ₄₈	100%	0%
H ₂ CO ₄₉	100%	0%
H ₂ CO ₅₀	100%	0%
H ₂ CO ₅₁	100%	0%
H ₂ CO ₅₂	100%	0%
H ₂ CO ₅₃	100%	0%
H ₂ CO ₅₄	100%	0%
H ₂ CO ₅₅	100%	0%
H ₂ CO ₅₆	100%	0%
H ₂ CO ₅₇	100%	0%
H ₂ CO ₅₈	100%	0%
H ₂ CO ₅₉	100%	0%
H ₂ CO ₆₀	100%	0%
H ₂ CO ₆₁	100%	0%
H ₂ CO ₆₂	100%	0%
H ₂ CO ₆₃	100%	0%
H ₂ CO ₆₄	100%	0%
H ₂ CO ₆₅	100%	0%
H ₂ CO ₆₆	100%	0%
H ₂ CO ₆₇	100%	0%
H ₂ CO ₆₈	100%	0%
H ₂ CO ₆₉	100%	0%
H ₂ CO ₇₀	100%	0%
H ₂ CO ₇₁	100%	0%
H ₂ CO ₇₂	100%	0%
H ₂ CO ₇₃	100%	0%
H ₂ CO ₇₄	100%	0%
H ₂ CO ₇₅	100%	0%
H ₂ CO ₇₆	100%	0%
H ₂ CO ₇₇	100%	0%
H ₂ CO ₇₈	100%	0%
H ₂ CO ₇₉	100%	0%
H ₂ CO ₈₀	100%	0%
H ₂ CO ₈₁	100%	0%
H ₂ CO ₈₂	100%	0%
H ₂ CO ₈₃	100%	0%
H ₂ CO ₈₄	100%	0%
H ₂ CO ₈₅	100%	0%
H ₂ CO ₈₆	100%	0%
H ₂ CO ₈₇	100%	0%
H ₂ CO ₈₈	100%	0%
H ₂ CO ₈₉	100%	0%
H ₂ CO ₉₀	100%	0%
H ₂ CO ₉₁	100%	0%
H ₂ CO ₉₂	100%	0%
H ₂ CO ₉₃	100%	0%
H ₂ CO ₉₄	100%	0%
H ₂ CO ₉₅	100%	0%
H ₂ CO ₉₆	100%	0%
H ₂ CO ₉₇	100%	0%
H ₂ CO ₉₈	100%	0%
H ₂ CO ₉₉	100%	0%
H _{2</}		

Identified 8,299 species from 58 models

- Malformed chemkin files
- Incorrect glossary entries
- 100 kJ/mol disagreements in enthalpies of formation
- 30 orders of magnitude disagreements in rates
- Only about 10% pressure-dependent

We frequently hear we can get 1 kJ/mol errors in energy and 30% in rate constants, yet our supplementary material differ by 100 kJ/mol and 10³⁰ respectively!
These discrepancies are usually undetected.

Some recent developments in RMG

- Developer-friendly "RMG-Py" being released
 - New features (Nitrogen!)
 - Web tools
- Group-Additive Transition State Estimates
 - Fast estimates of TS geometries
 - Gets better at guessing the more it guesses
- Mechanism Importer tool
 - Facilitates identification of species in chemkin files
 - Started to collect and curate data

Richard H. West

4 August 2014



r.west@neu.edu



richardhwest



rwest



Northeastern.edu/comocheng

And of course, <http://www.slideshare.net/richardhwest>