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

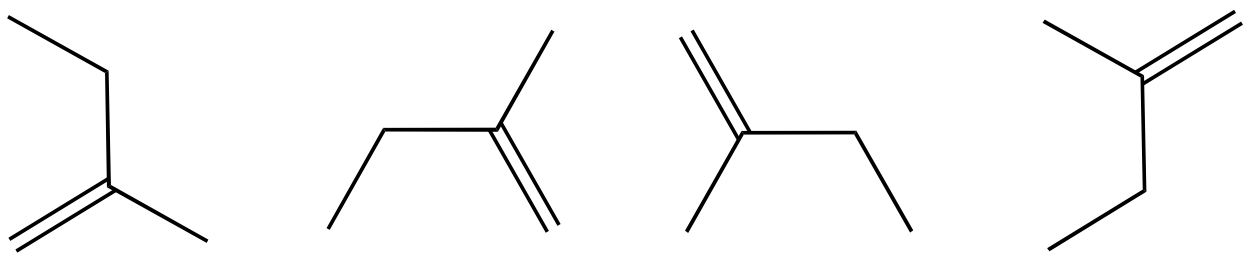
RMG

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)

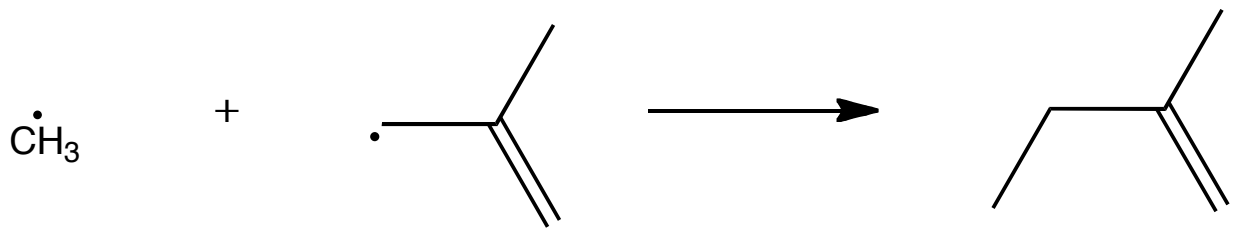


The common theme to the two projects in the talk is Reaction Mechanism Generator software.

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.

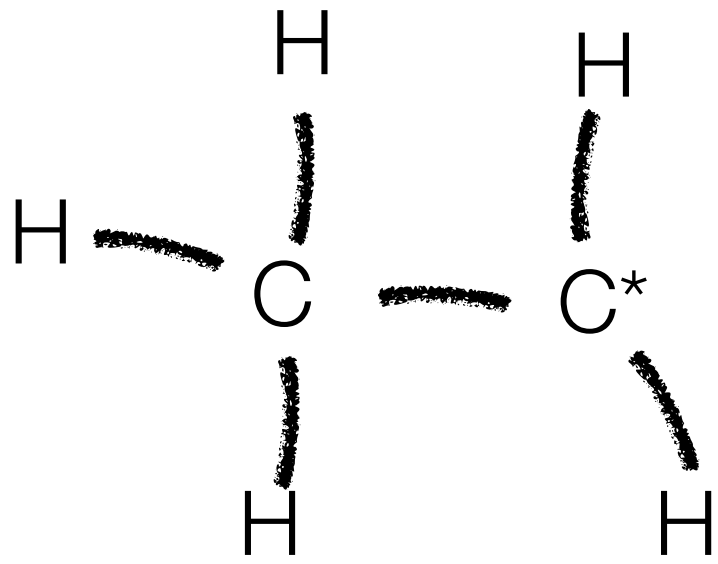
This is because the room layout meant most people couldn't see the bottom of the screen!

Automatic reaction mechanism generation needs methods to:

4. Estimate thermo and kinetic parameters (quickly!)

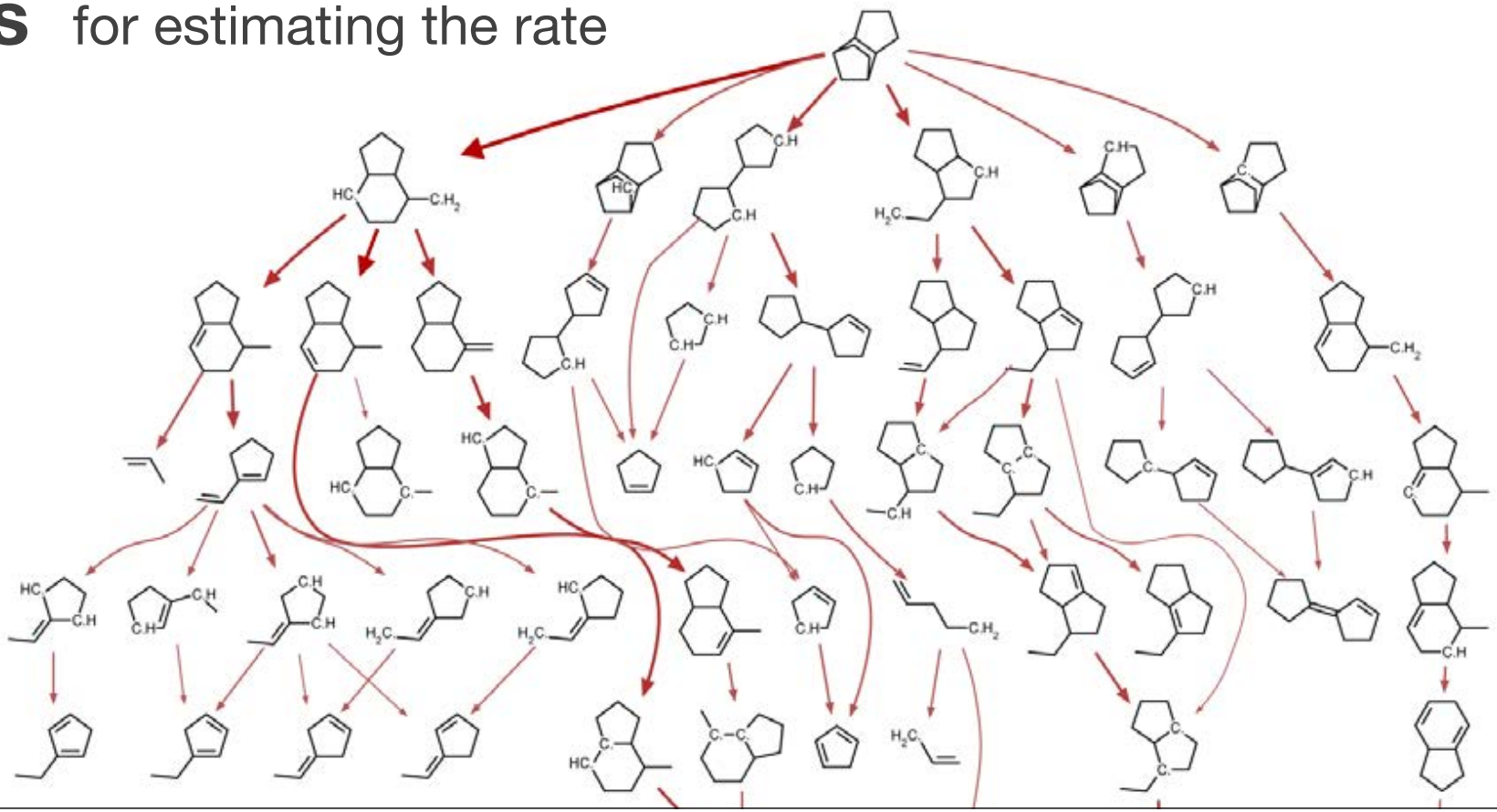
Molecules are represented as 2D graphs

$CH_3CH_2.$ =



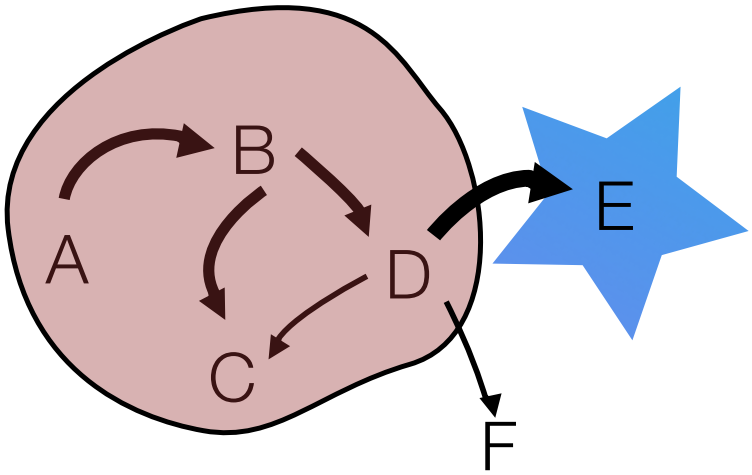
Reaction families propose all possible reactions with given species

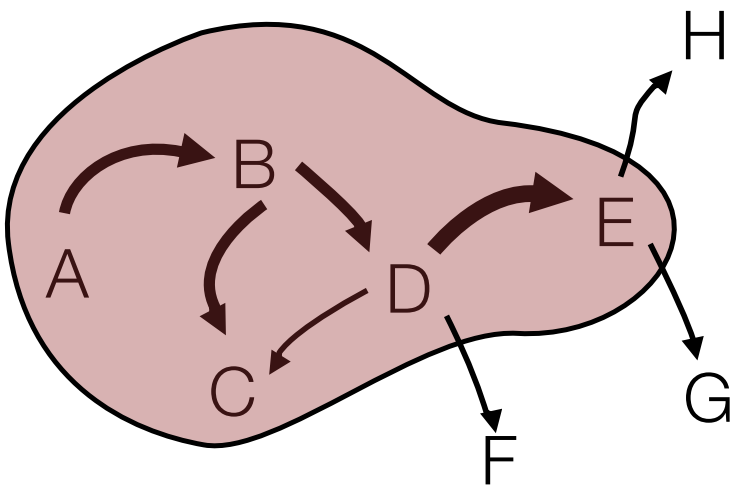
- **Template** for recognizing reactive sites
- **Recipe** for changing the bonding at the site
- **Rules** for estimating the rate



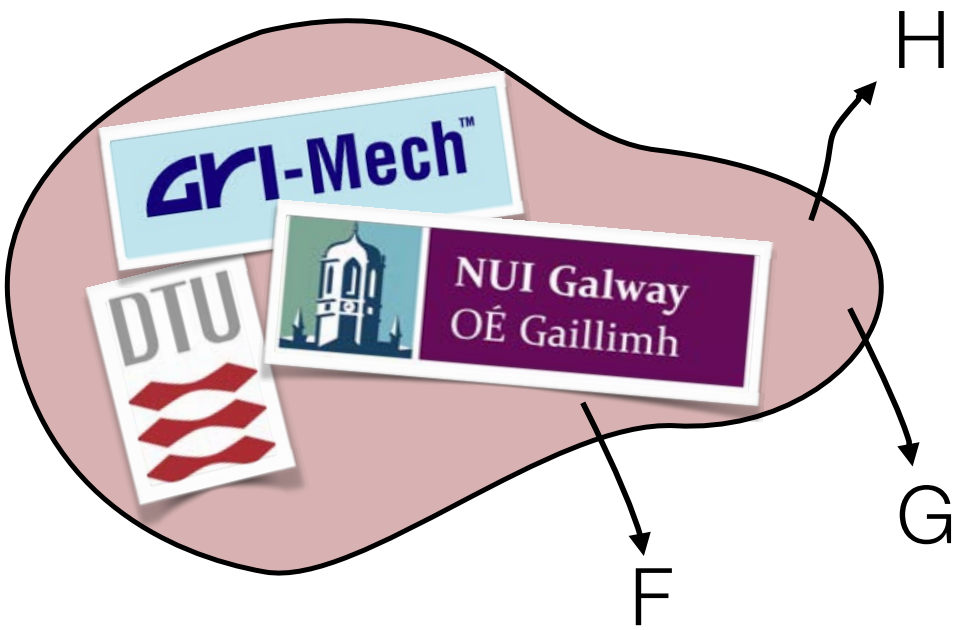
There are ~40 reaction families such as as Hydrogen abstraction, unimolecular homolysis, radical addition to a double bond…

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



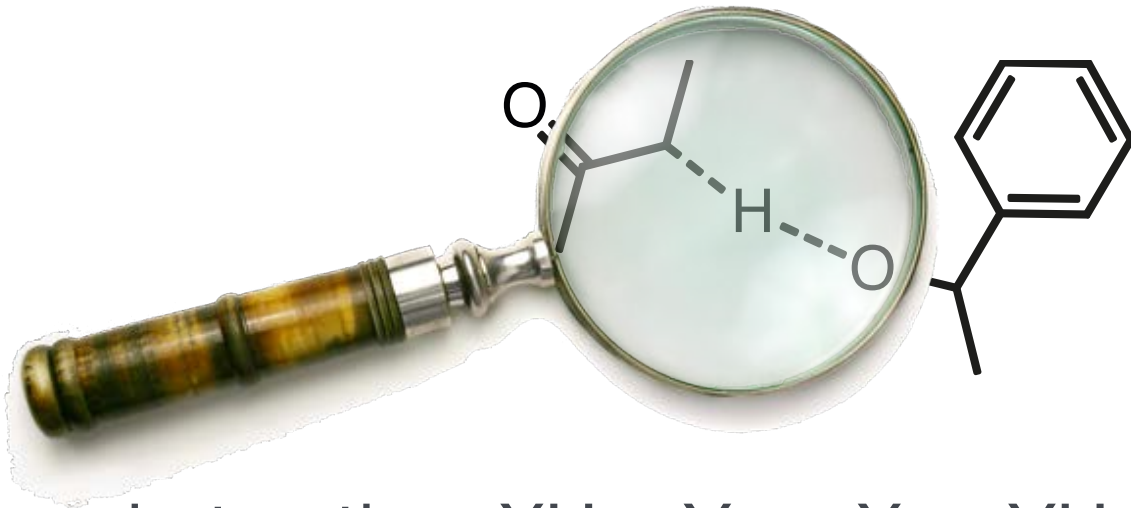


Model can be started from “seed mechanism” (if you have this in RMG format)



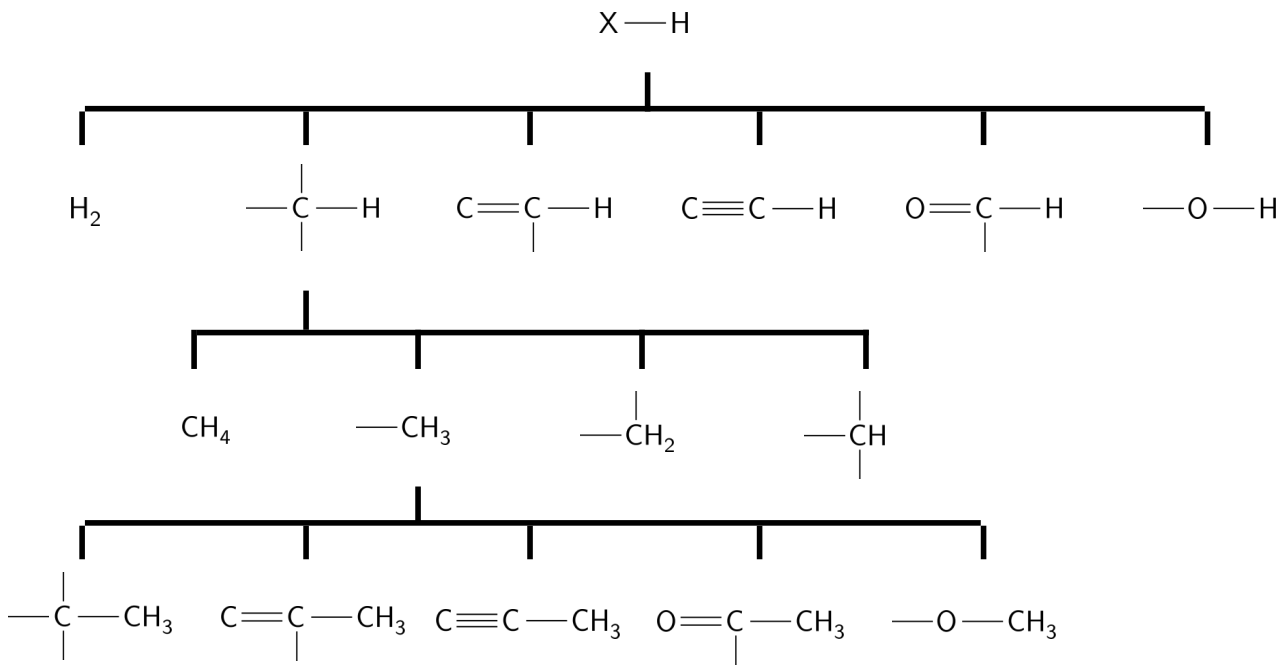
Getting it in “RMG format” is the hard part, and will be addressed in the second half of the talk.

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



- Hydrogen abstraction: XH + Y. → X. + YH
- Rate depends on X and Y.

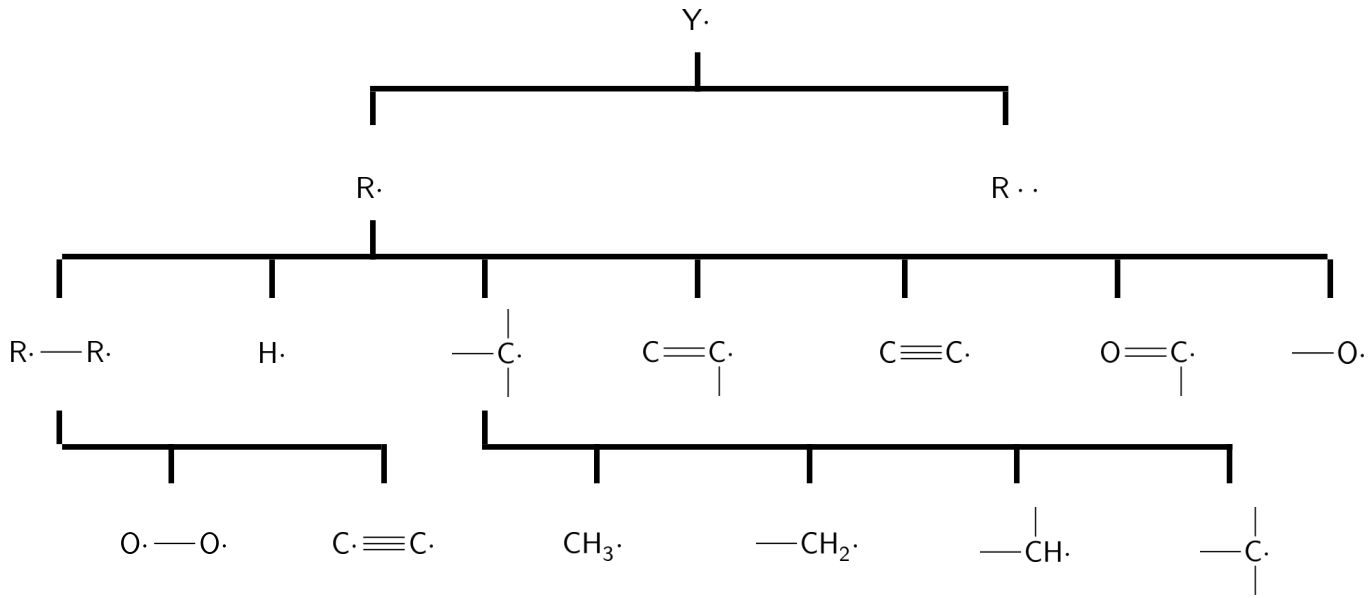
Rate estimation rules are organized in a tree



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.



Part of the tree for Y

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

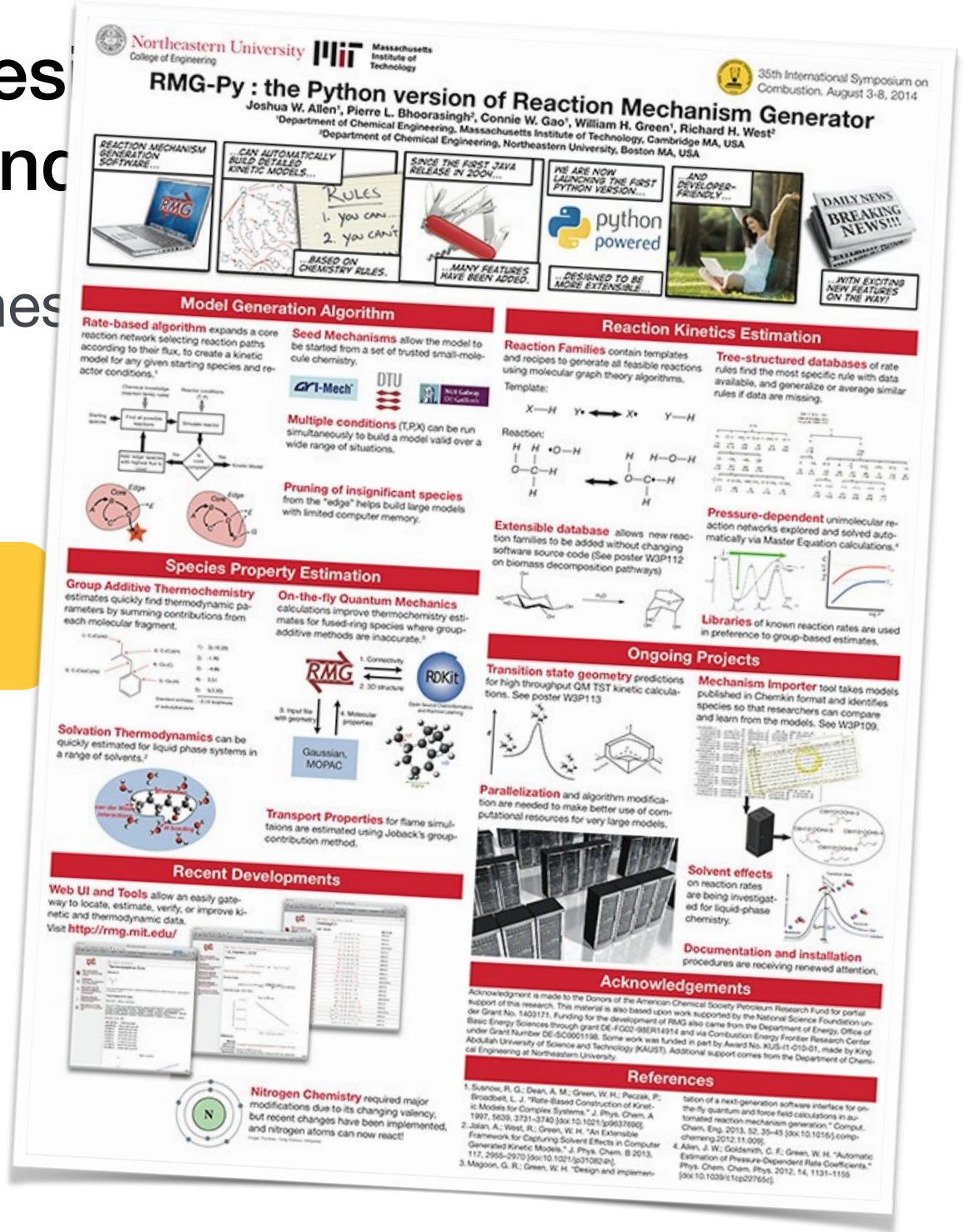
- See poster on Wednesday

python powered

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

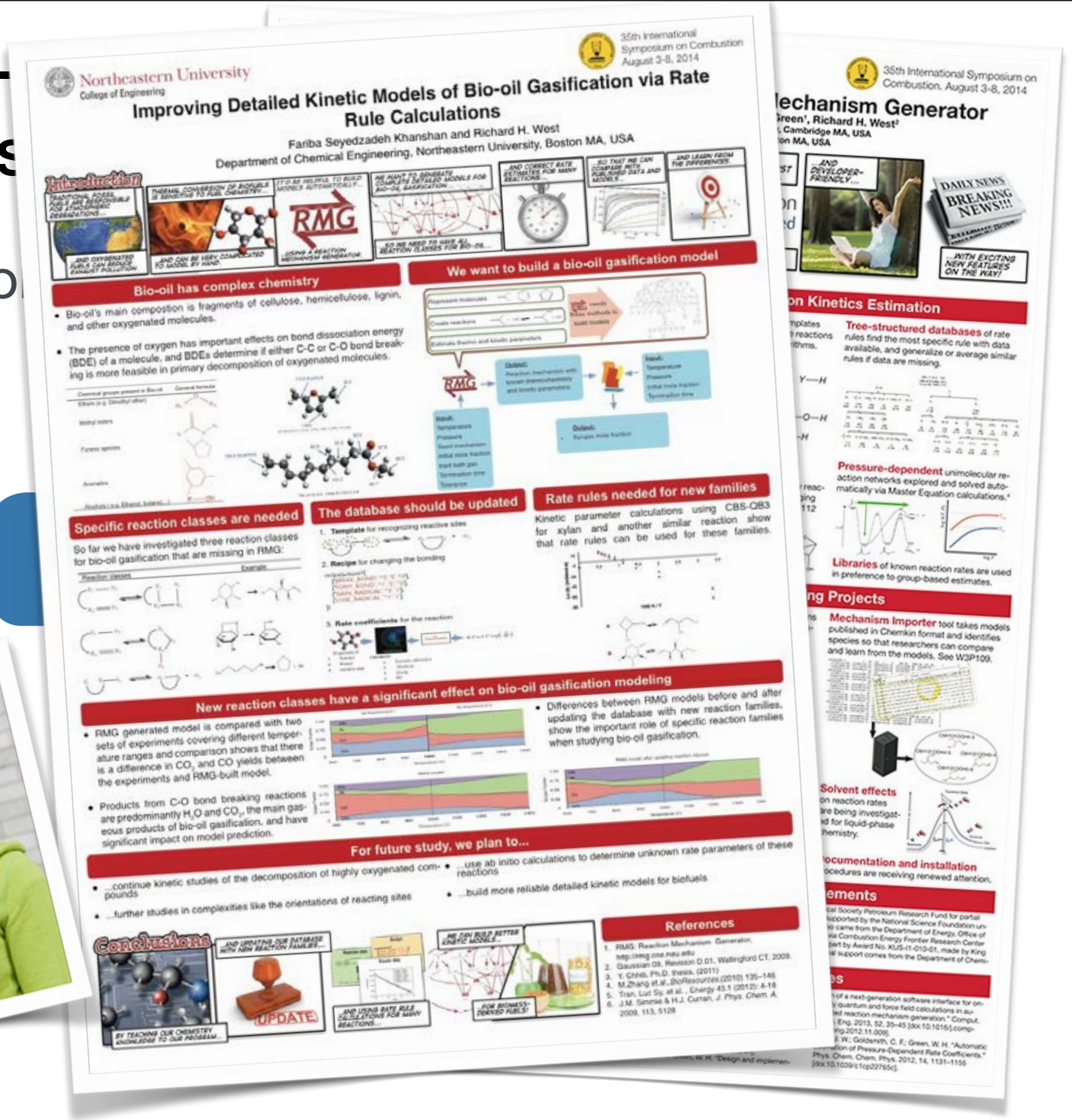
- See poster on Wednes



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 o



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

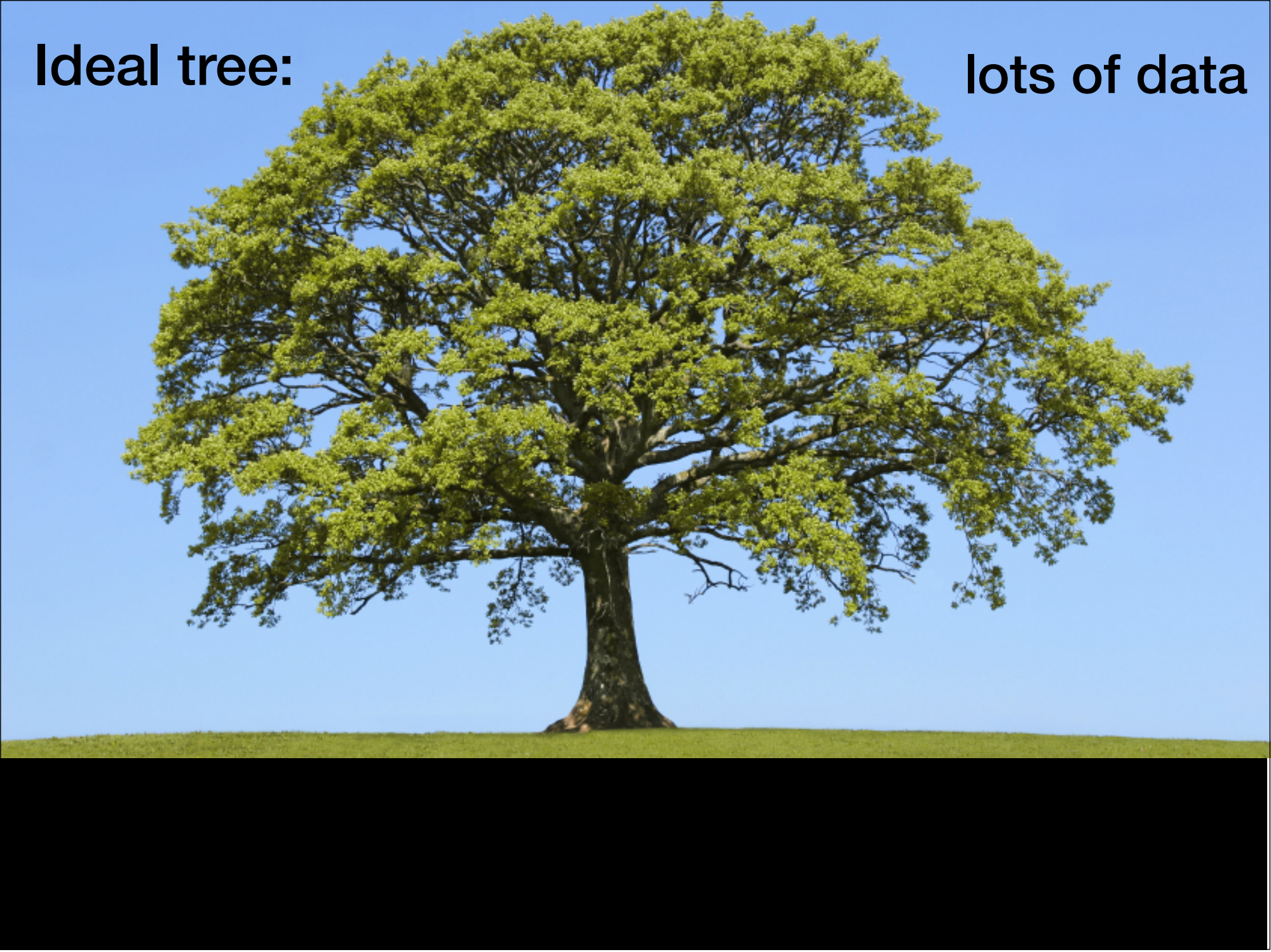
Toward High-Throughput Transition State Calculations

Pierre L. Bhoorasingh

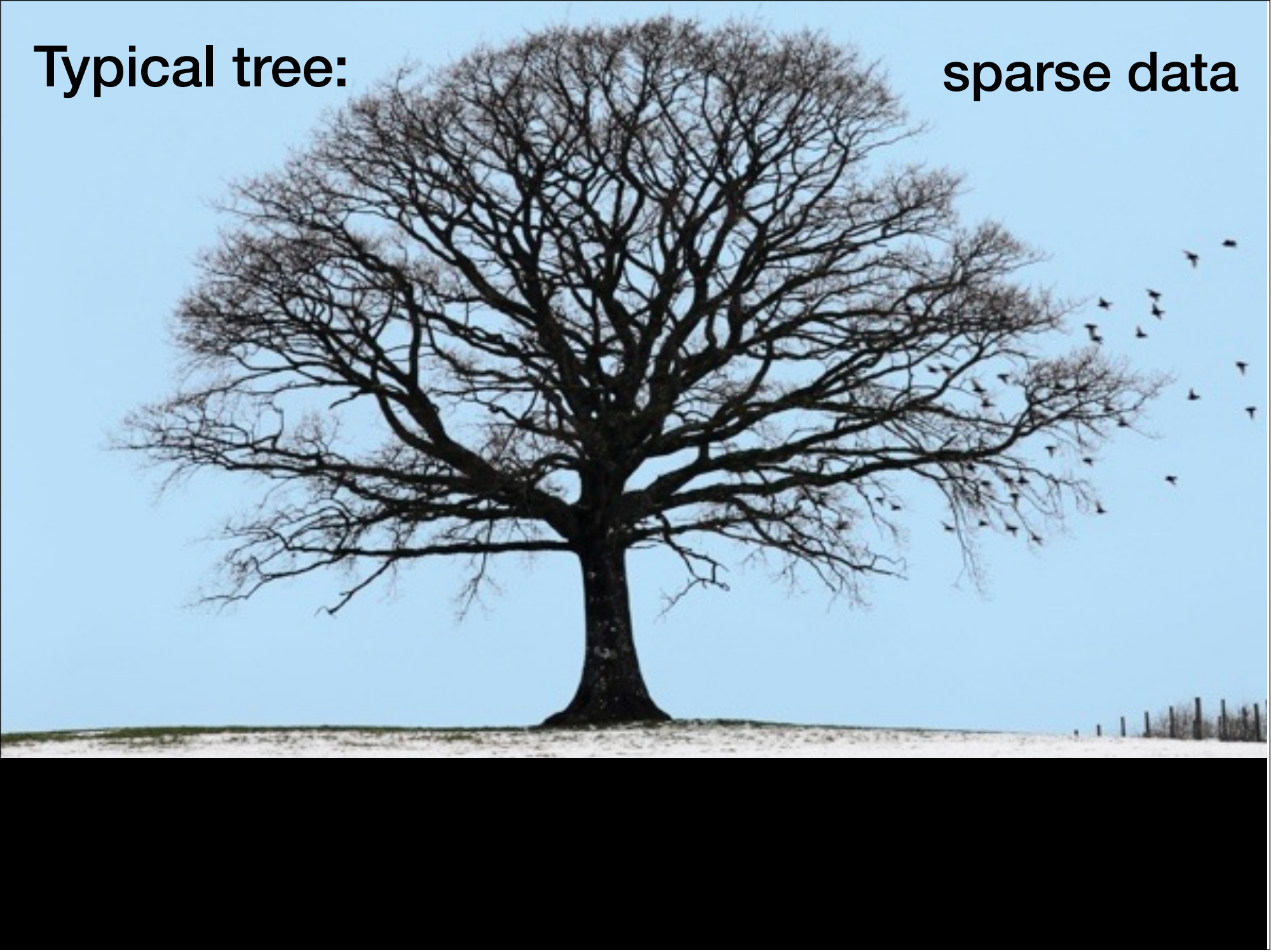
Acknowledgment is made to the Donors of the American Chemical Society Petroleum Research Fund for support of this research.

Northeastern.edu/comocheng

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

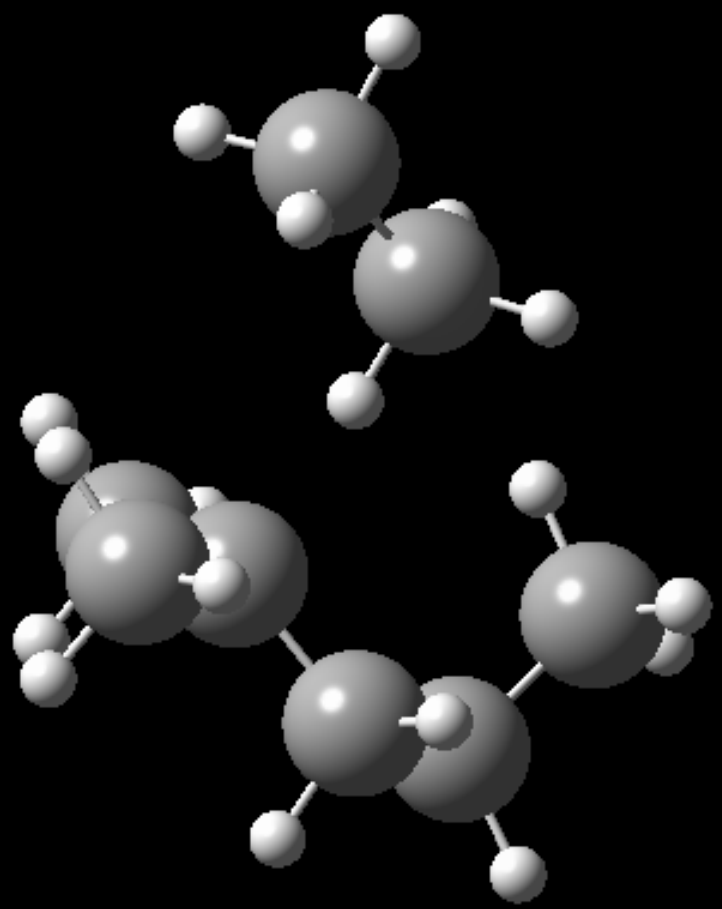


Ideally we climb the tree, find the matching node for our reaction,
and pluck the leaf containing the rate expression.

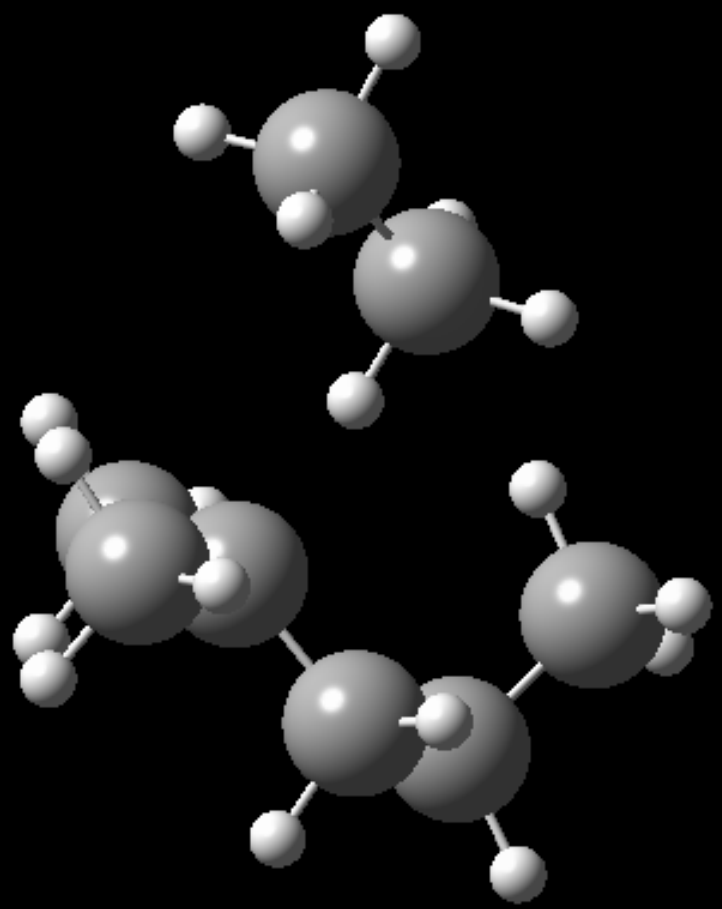


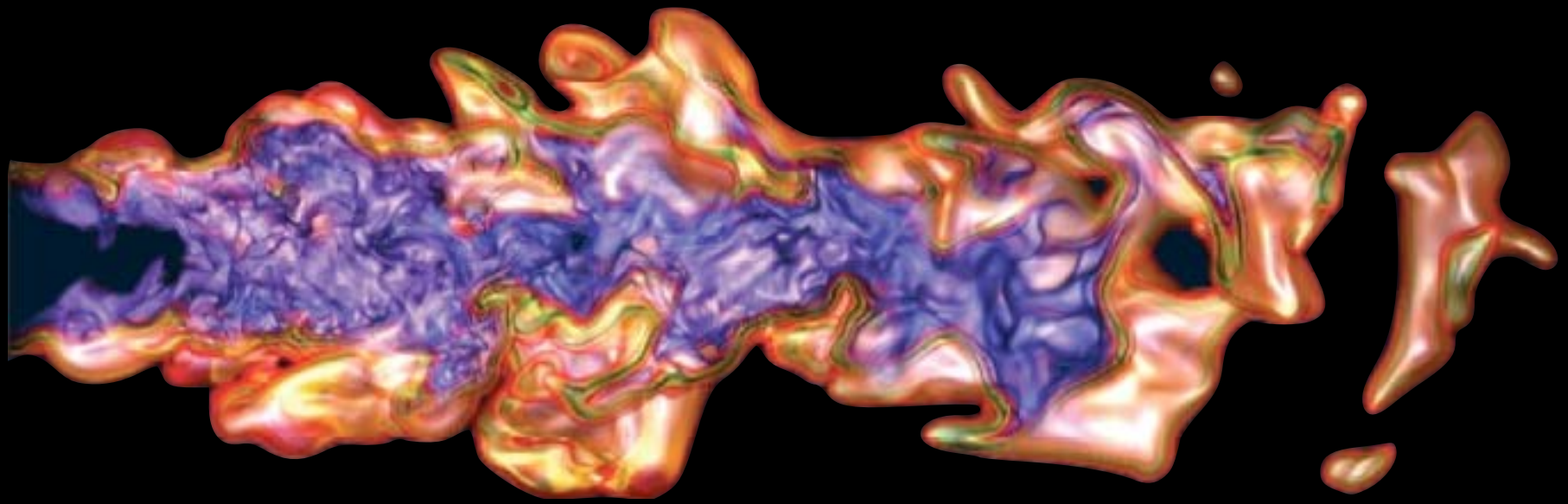
But what if there are no data? How can we make it?

Transition State calculations can provide missing data



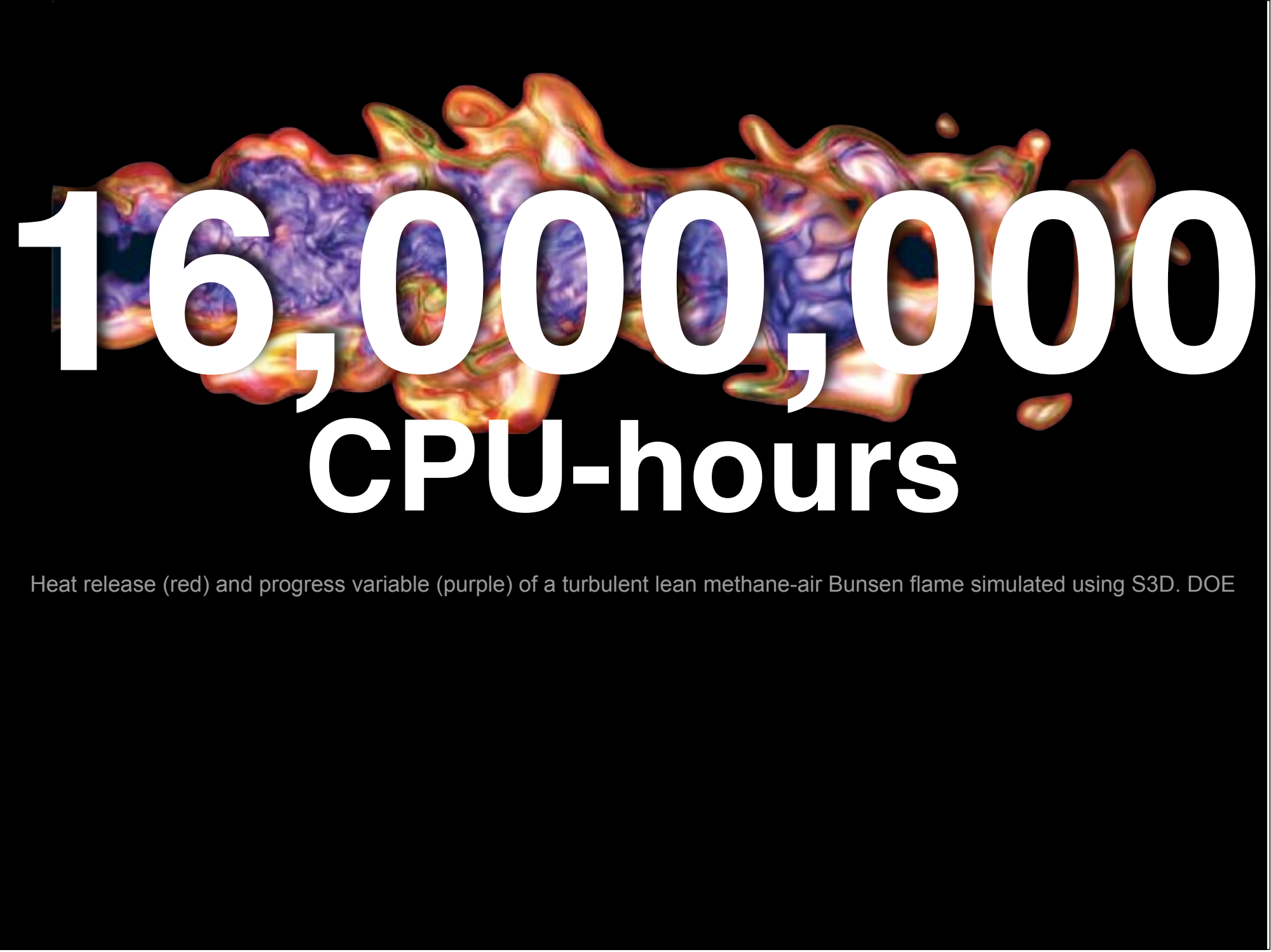
Transition State calculations can provide missing data





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



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

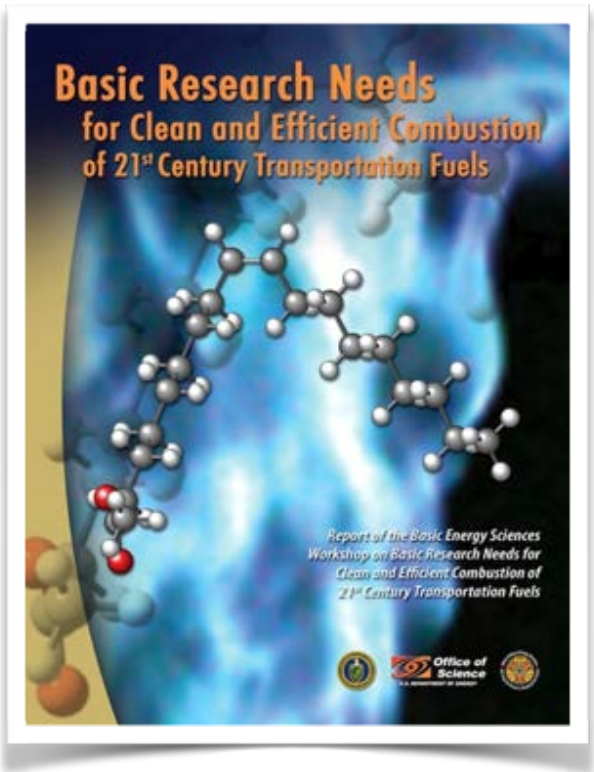
This is one of the large federal computers they use.

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

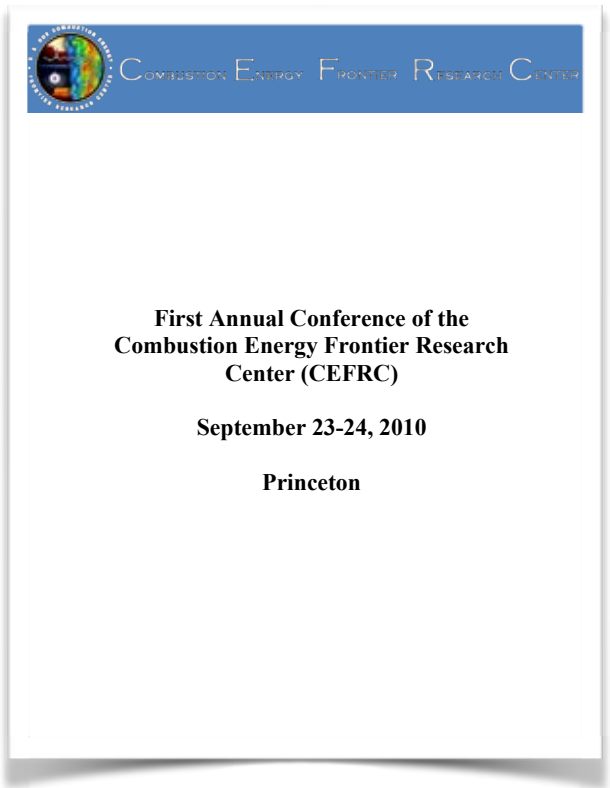


"An accurate description of the often intricate mechanisms of large-molecule reactions requires a characterization of all relevant transition states... Development of automatic means to search for chemically relevant configurations is the computational-kinetics equivalent of improved electronic structure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

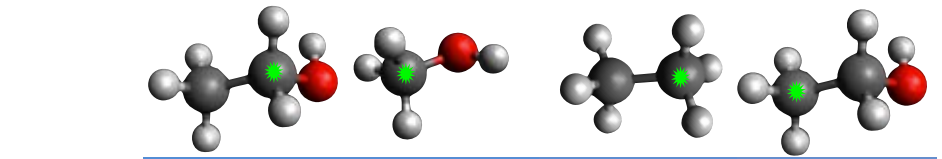
Can you predict TS geometries from molecular groups alone?

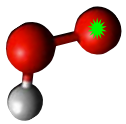
Molecule

Radical

Length of bond being broken,
at TS for Hydrogen abstraction

Can you predict TS geometries from molecular groups alone?

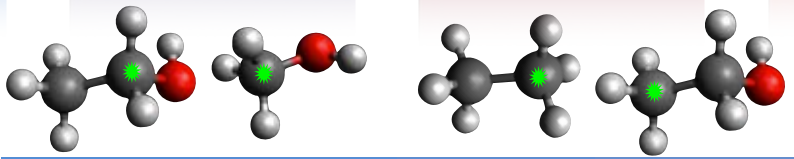


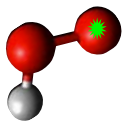
	1.149	1.155	1.168	1.179
	1.272	1.286	1.309	1.314
	1.278	1.295		1.355
	1.280	1.306	1.362	1.369

in Å with M06-2X/6-31+G(d,p)

Can y di TS es
from r la os

1.168	1.272	1.345	1.372

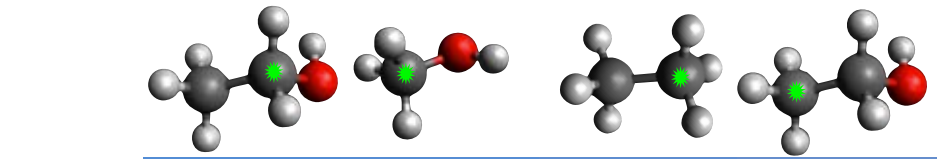


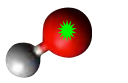
	1.149	1.155	1.168	1.179
	1.272	1.286	1.309	1.314
	1.278	1.295		1.355
	1.280	1.306	1.362	1.369

in Å with M06-2X/6-31+G(d,p)

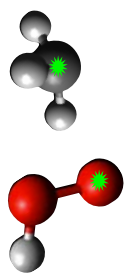
Guess which number goes in the blank?

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



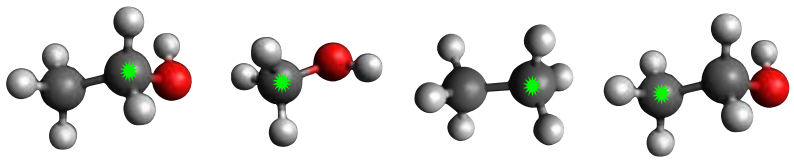
in Å with M06-2X/6-31+G(d,p)

So you can predict distances! But…

(1) I gave you 15 numbers and you gave be one; and

(2) you only gave me a distance, not a 3D geometry.

Add the two numbers to get any distance within 1 pm

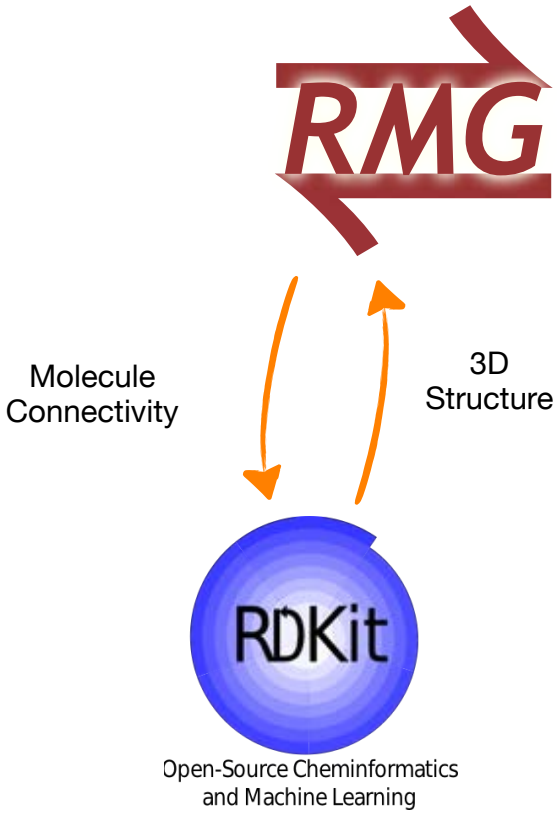


	0.607	0.622	0.658	0.666
0.525				
0.657				
0.680				
0.691				

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.

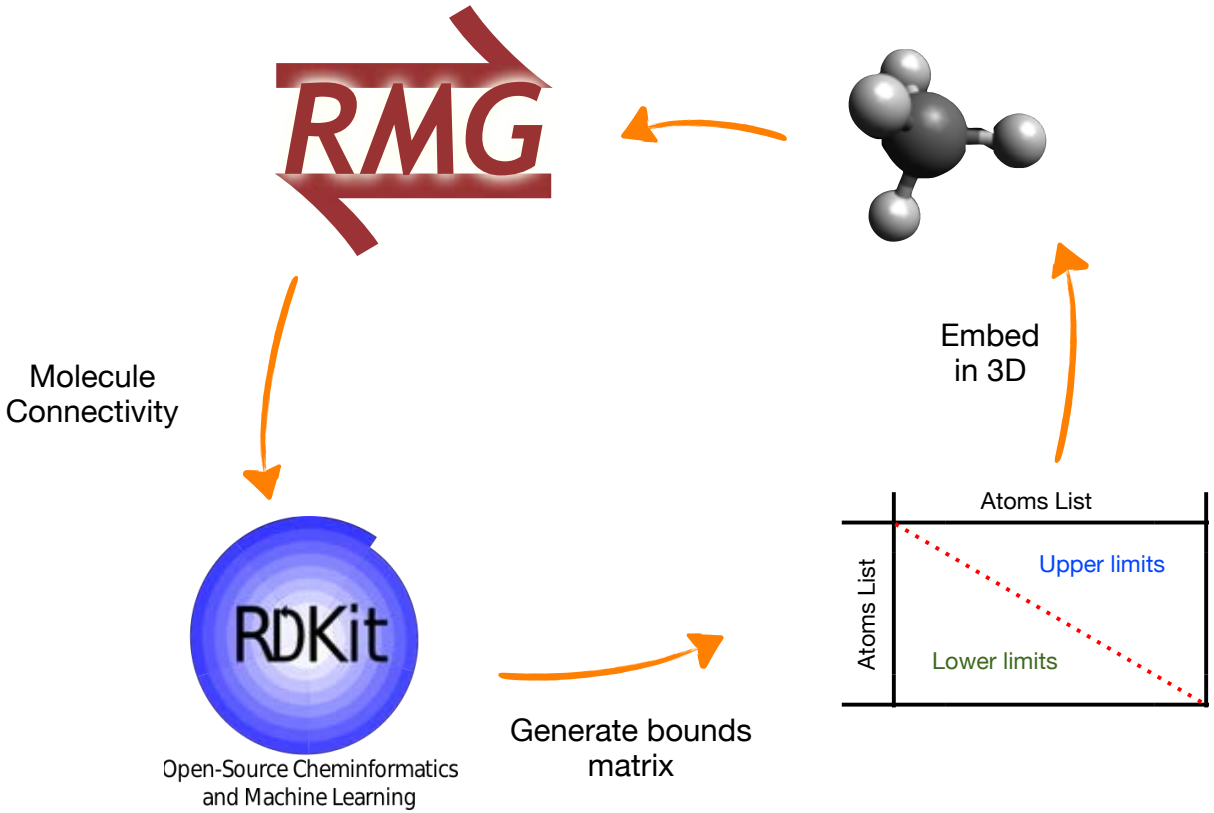
Use distance geometry to position atoms



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

We use the distance-geometry algorithms in RDKit.

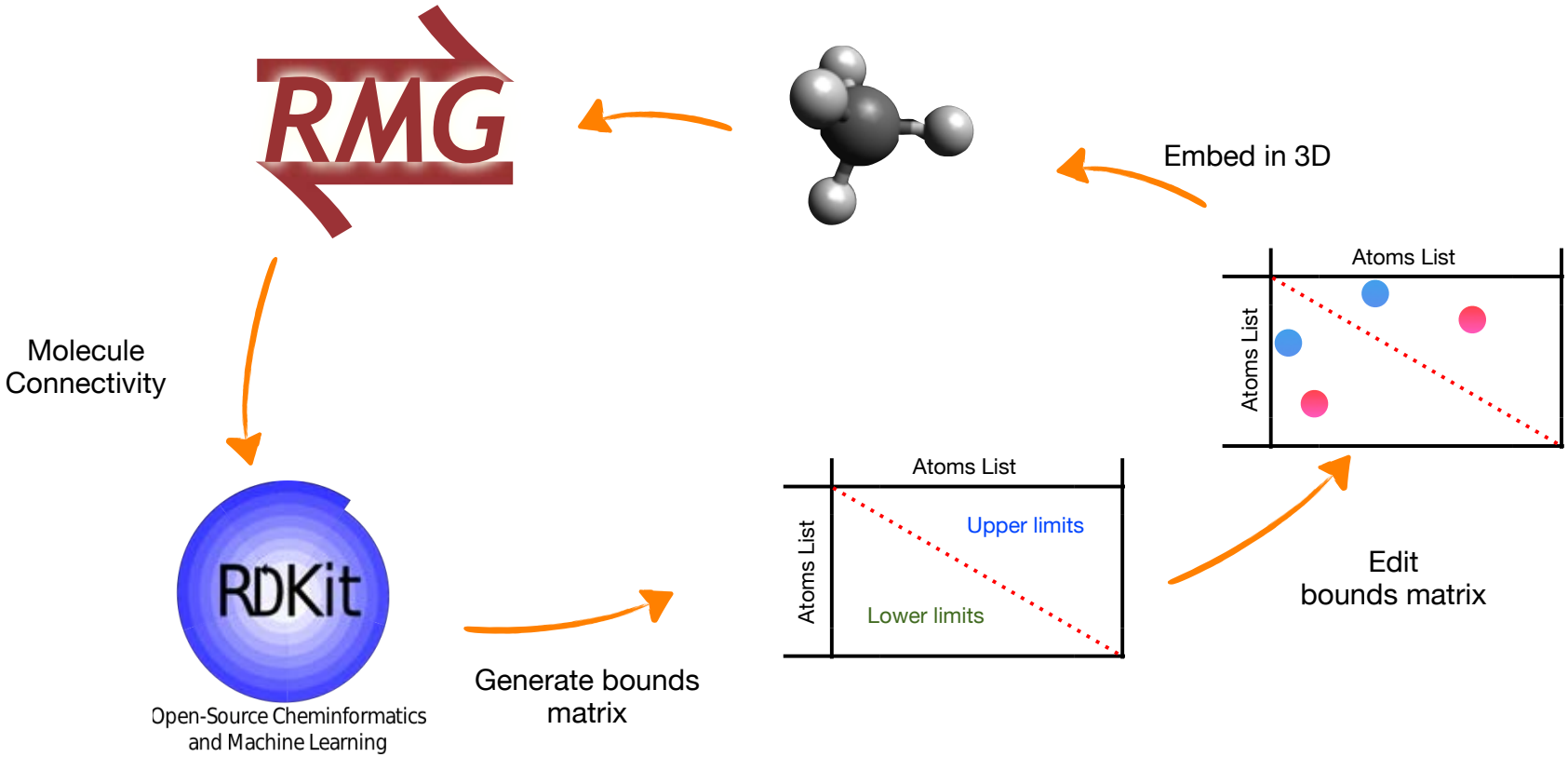
Use distance geometry to position atoms



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

We use the distance-geometry algorithms in RDKit.

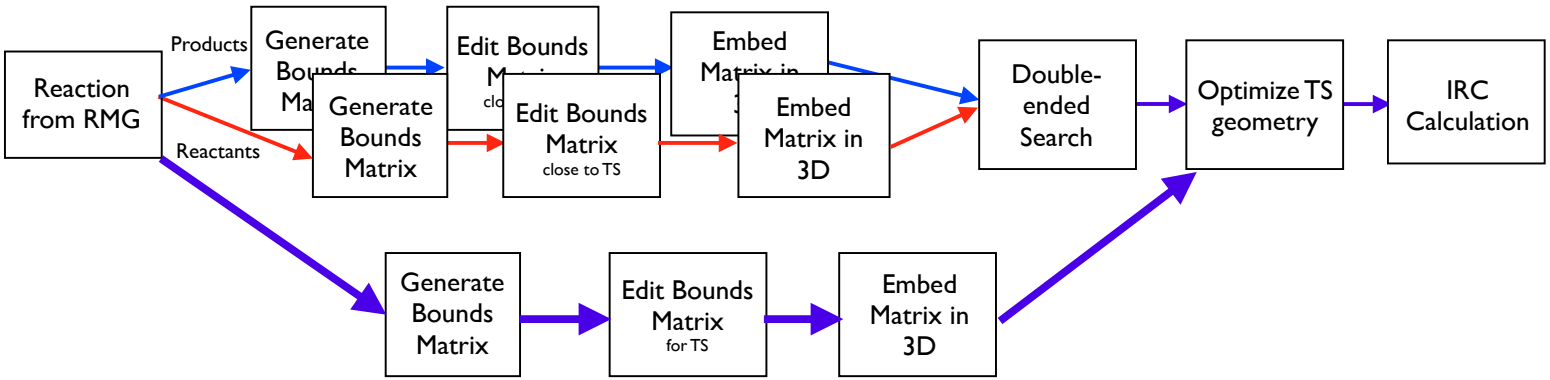
Use distance geometry to position atoms



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

Test the group-additive method on H-abstraction reactions in Diisopropyl Ketone

Allen, J. W. et al. Combust. Flame 161, 711–724 (2014).

Combustion and Flame 161 (2014) 711–724

Contents lists available at ScienceDirect

Combustion and Flame

journal homepage: www.elsevier.com/locate/combustflame

A coordinated investigation of the combustion chemistry of diisopropyl ketone, a prototype for biofuels produced by endophytic fungi

Joshua W. Allen [a], Adam M. Scheer [b], Connie W. Gao [a], Shamel S. Merchant [a], Subith S. Vasu [c], Oliver Welz [b], John D. Savee [b], David L. Osborn [b], Changyoul Lee [d], Stijn Vranckx [d], Zhandong Wang [e], Fei Qi [e], Ravi X. Fernandes [d,f], William H. Green [a,*], Masood Z. Hadi [g,1], Craig A. Taatjes [b,*]

[a] Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[b] Combustion Research Facility, Sandia National Laboratories Livermore, CA 94551, USA
[c] Mechanical and Aerospace Engineering, University of Central Florida, Orlando, FL 32708, USA
[d] Physico-Chemical Fundamentals of Combustion, RWTH Aachen University, 52056 Aachen, Germany
[e] National Synchrotron Radiation Laboratory, University of Science and Technology of China, Hefei, Anhui 230029, PR China
[f] Physikalisch-Technische Bundesanstalt, Bundesallee 100, D38116 Braunschweig, Germany
[g] Biomass Conversion Technologies, Sandia National Laboratories, Livermore, CA 94551, USA

ARTICLE INFO

Article history:
Received 25 July 2013
Received in revised form 18 September 2013
Accepted 18 October 2013
Available online 20 November 2013

Keywords:
Diisopropyl ketone
Automatic mechanism generation
Ignition delay
Pyrolysis
Combustion
Detailed kinetics modeling

ABSTRACT

Several classes of endophytic fungi have been recently identified that convert cellulosic biomass to a range of ketones and other oxygenated molecules, which are potentially viable as biofuels, but whose oxidation chemistry is not yet well understood. In this work, we present a predictive kinetics model describing the pyrolysis and oxidation of diisopropyl ketone (DIPK) that was generated automatically using the Reaction Mechanism Generator (RMG) software package. The model predictions are evaluated against three experiments that cover a range of temperatures, pressures, and oxygen concentrations: (1) Synchrotron photoionization mass spectrometry (PIMS) measurements of pyrolysis in the range 800–1340 K at 30 Torr and 760 Torr; (2) Synchrotron PIMS measurements of laser photolytic Cl-initiated oxidation from 550 K to 700 K at 8 Torr; and (3) Rapid-compression machine measurements of ignition delay between 591 K and 720 K near 10 bar. Improvements made to the model parameters, particularly in the areas of hydrogen abstraction from the initial DIPK molecule and low-temperature peroxy chemistry, are discussed. Our ability to automatically generate this model and systematically improve its parameters without fitting to the experimental results demonstrates the usefulness of the predictive chemical kinetics paradigm.

1. Introduction

Liquid petroleum-based fuels account for more than 30% of world energy consumption, the majority occurring in the transportation sector, where liquid fuels dominate because of their high energy density and ease of storage and distribution [1]. However, insecurity of oil supply and negative impacts of climate change are motivating increased development of alternative liquid fuels [2]. Lignocellulosic biofuels are a promising renewable alternative to conventional petroleum fuels that could lower greenhouse gas emissions. Due to the recalcitrance of lignocellulosic materials, new means for breaking down biomass into fermentable sugars are a central area of biofuels research [3]. However, the products that are most efficiently produced from biomass may not have well-characterized combustion properties. Furthermore, advanced clean, efficient combustion strategies, particularly those that rely on compression ignition, are often very sensitive to fuel chemistry [4,5]. To address these issues comprehensively, coordinated efforts towards biofuel-engine co-development are required.

Figure 1 depicts the Sandia collaborative biofuel development strategy. Synthetic biologists working in close collaboration with combustion researchers develop fundamental mechanisms for the production and combustion of potential biofuels. Ignition and engine trials then provide feasibility tests for fuels and mixtures and yield recommendations for the bioengineering scale-up of specific metabolic pathways. This coupling of fundamental and applied combustion chemistry and synthetic biology is a strategy to identify and investigate the most promising fuel compounds through mutual feedback. Moreover, the development of combustion models will provide the predictive capability needed for eventual efficient utilization of the new biofuel stream. A similar

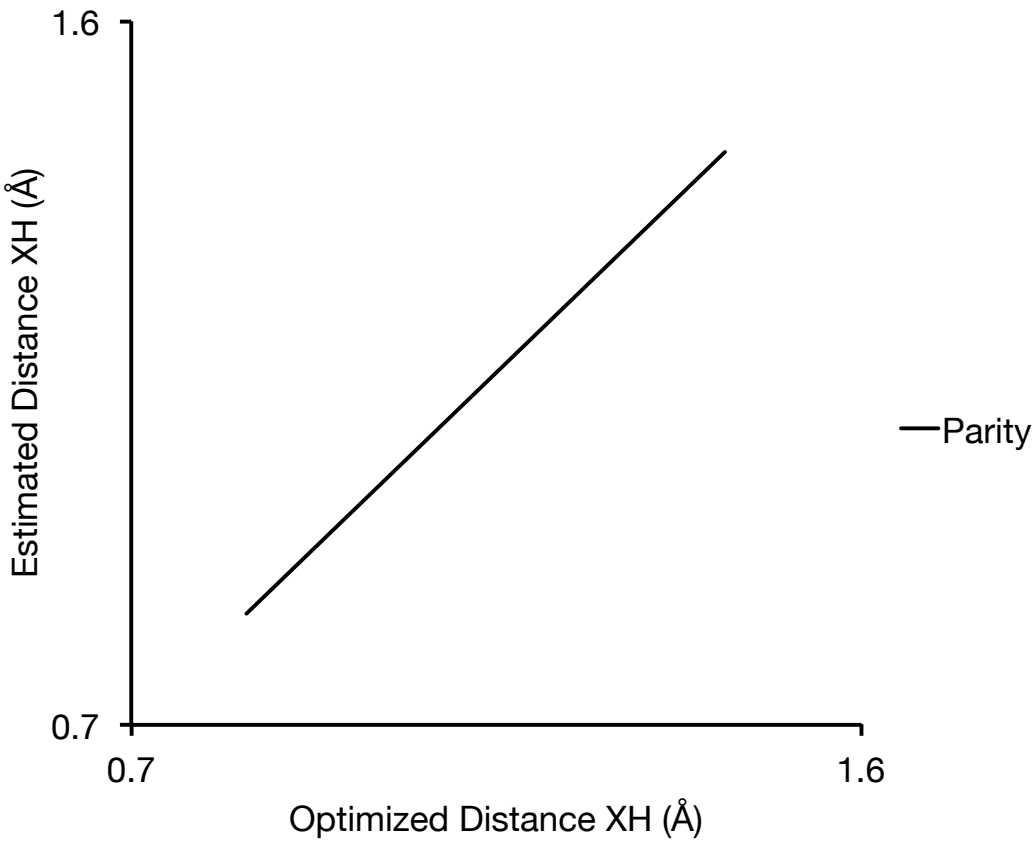
* Corresponding authors.
E-mail addresses: whgreen@mit.edu (W.H. Green), cataatj@sandia.gov (C.A. Taatjes).
[1] Present address: NASA Ames Research Center, Moffett Field, CA 94035, USA.

0010-2180/$ - see front matter © 2013 The Combustion Institute. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.combustflame.2013.10.019

These data were hot off the press.

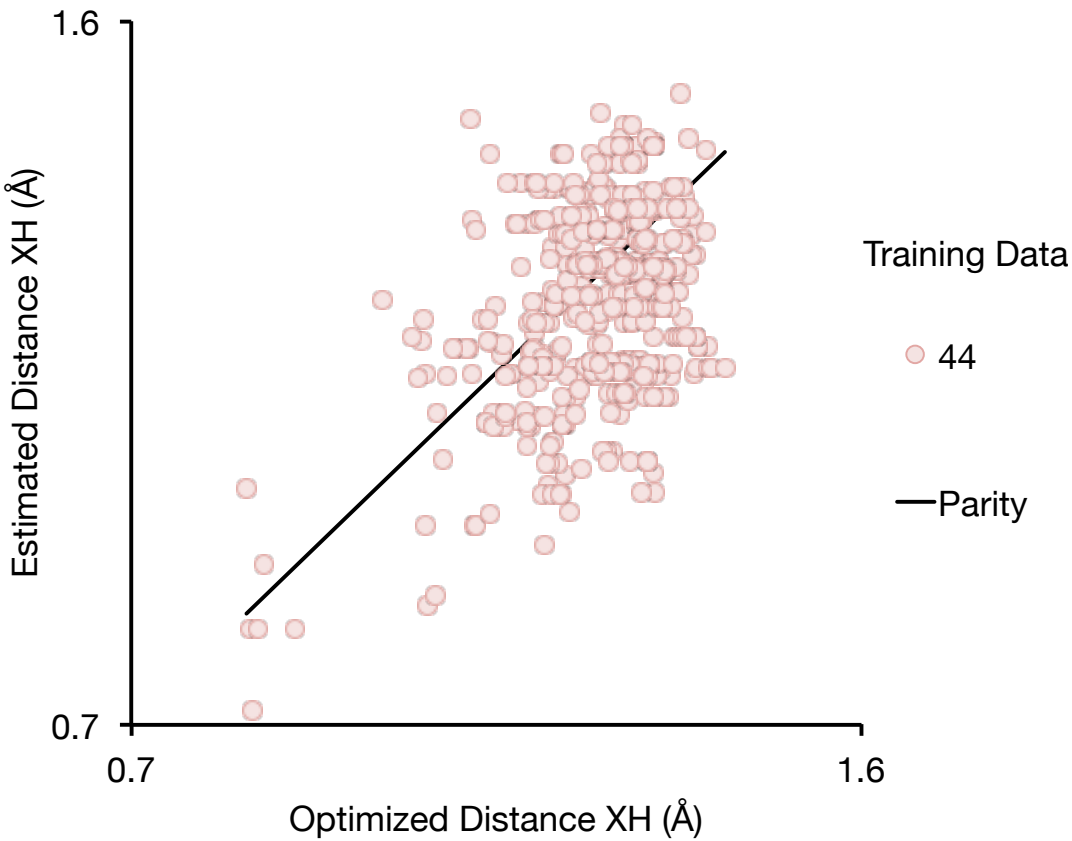
We tried finding every H-abstraction reaction in this large published kinetic model.

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

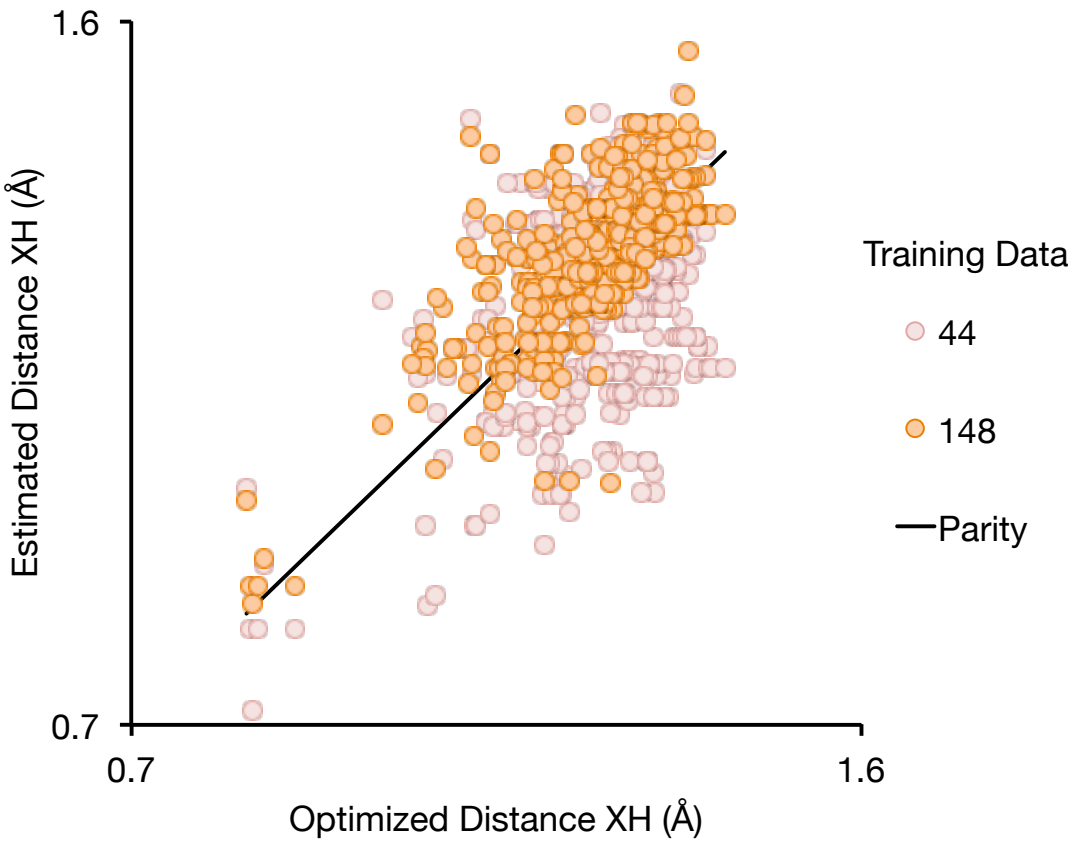


We'll see how well we predict the distances as we re-train the groups on more and more data.

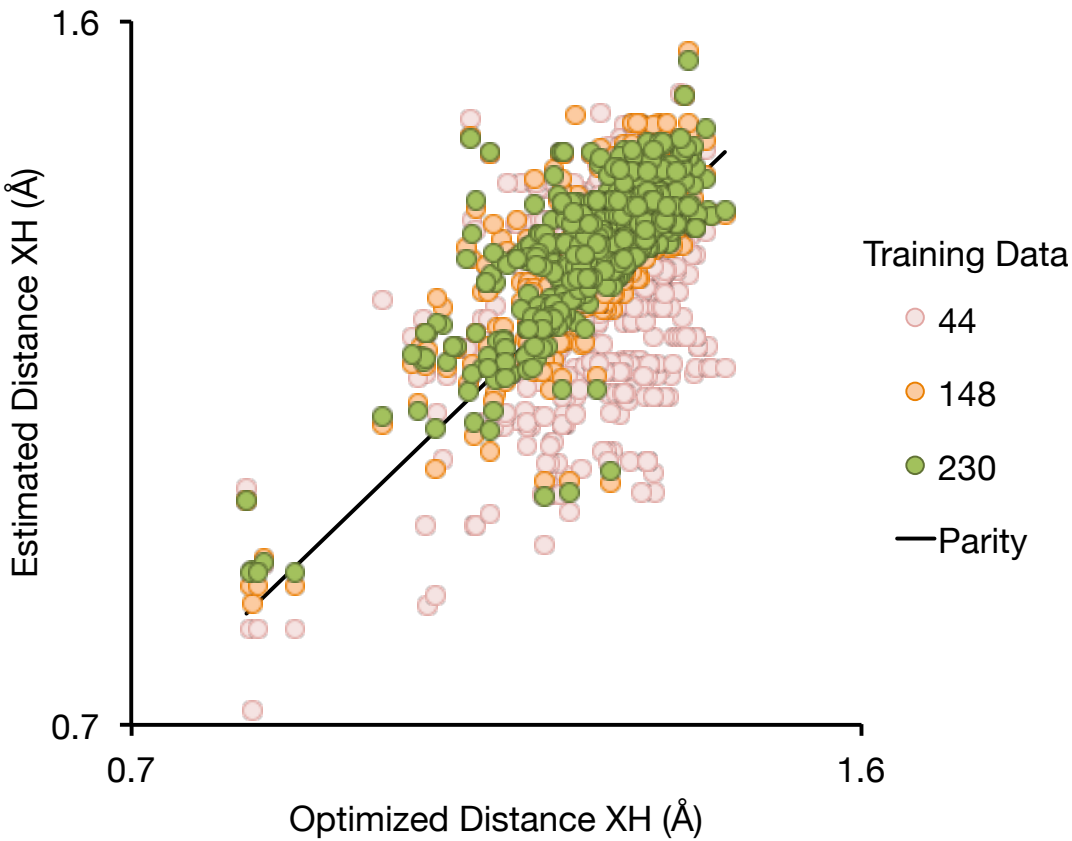
Trained on 44 reactions, estimates not so great…



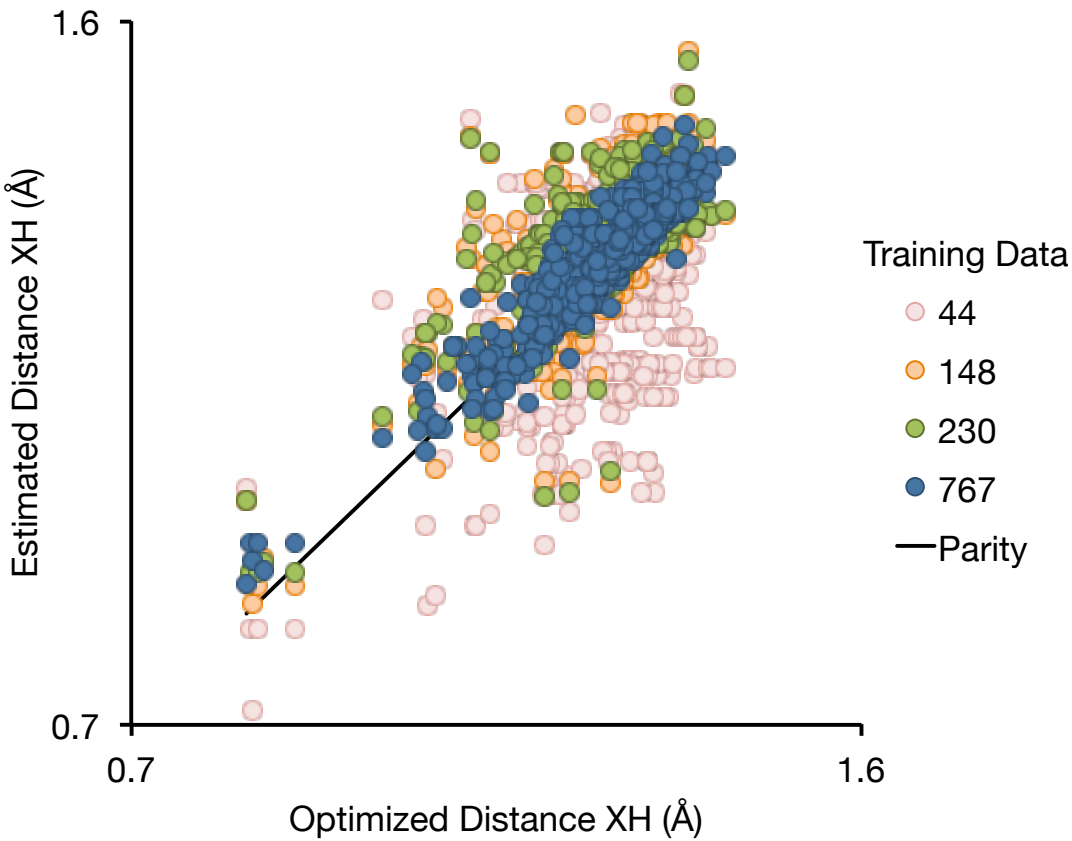
Trained on 148 reactions, estimates improve…



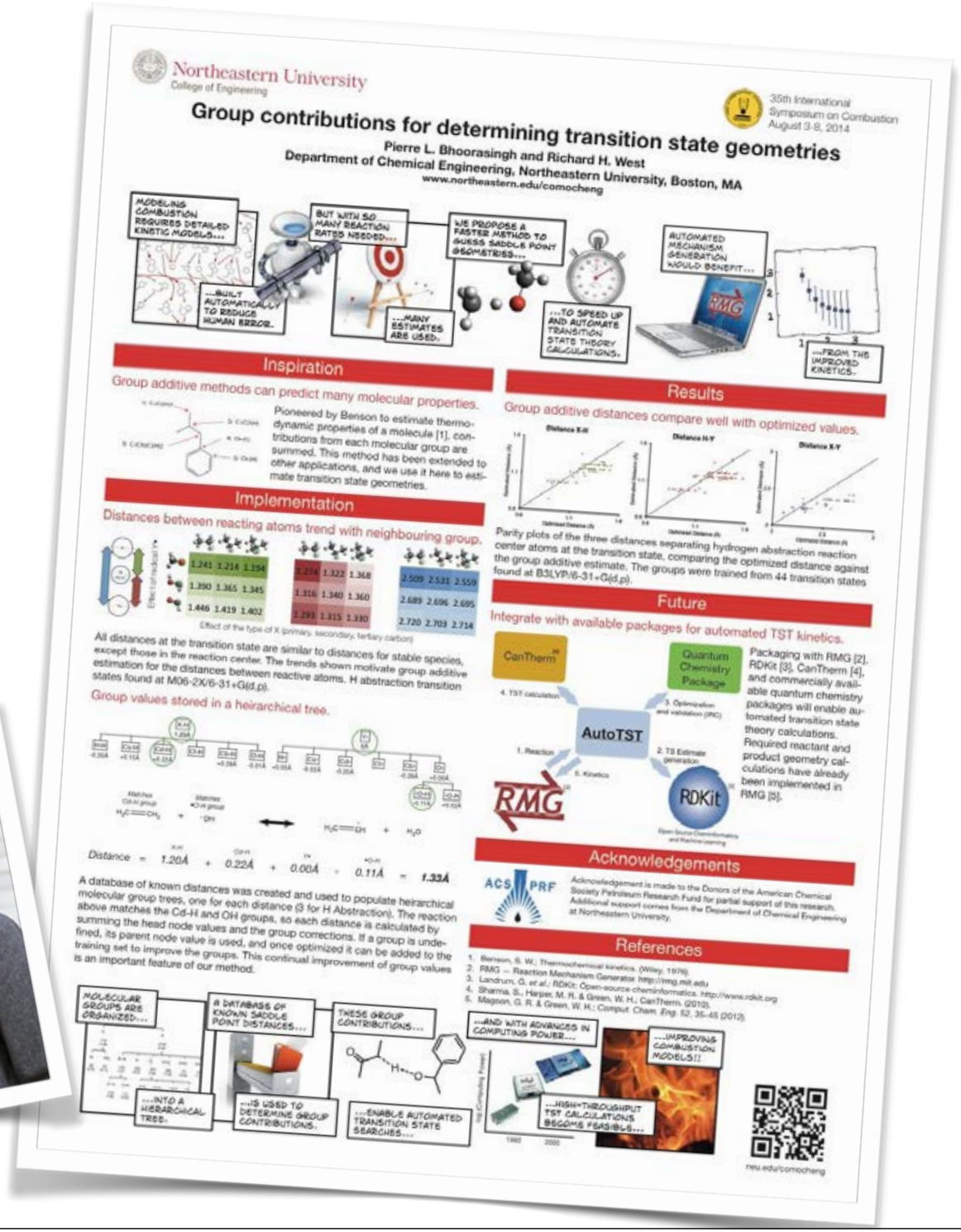
Trained on 230 reactions, estimates improve…



Trained on 767 reactions, estimates are very good.



- See Pierre Bhoorasingh's poster on Wednesday



For more details, talk to Pierre (who was up all night generating those results)

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.

Keywords: Kinetics; Mo

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.

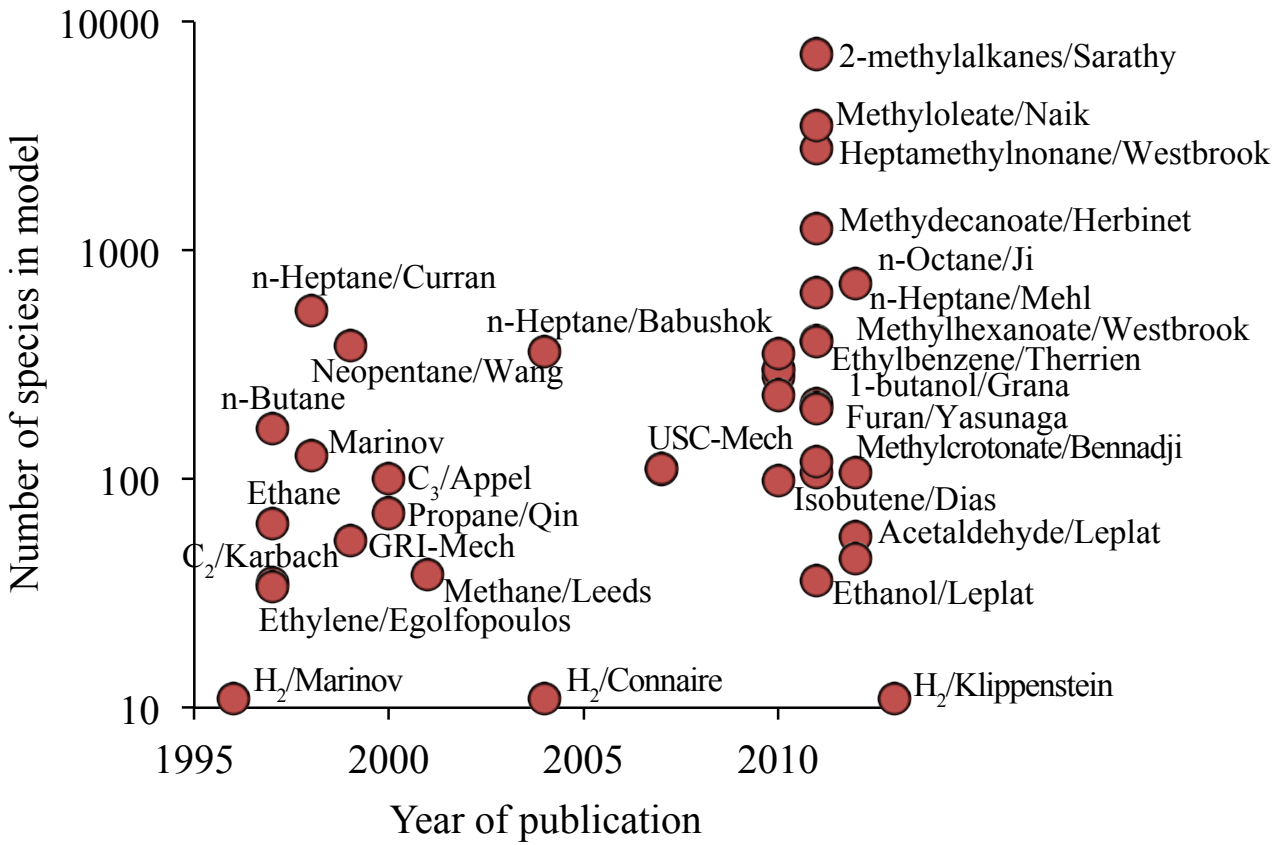
Process Informatics introduces a new paradigm. It relies on three major components:

- 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



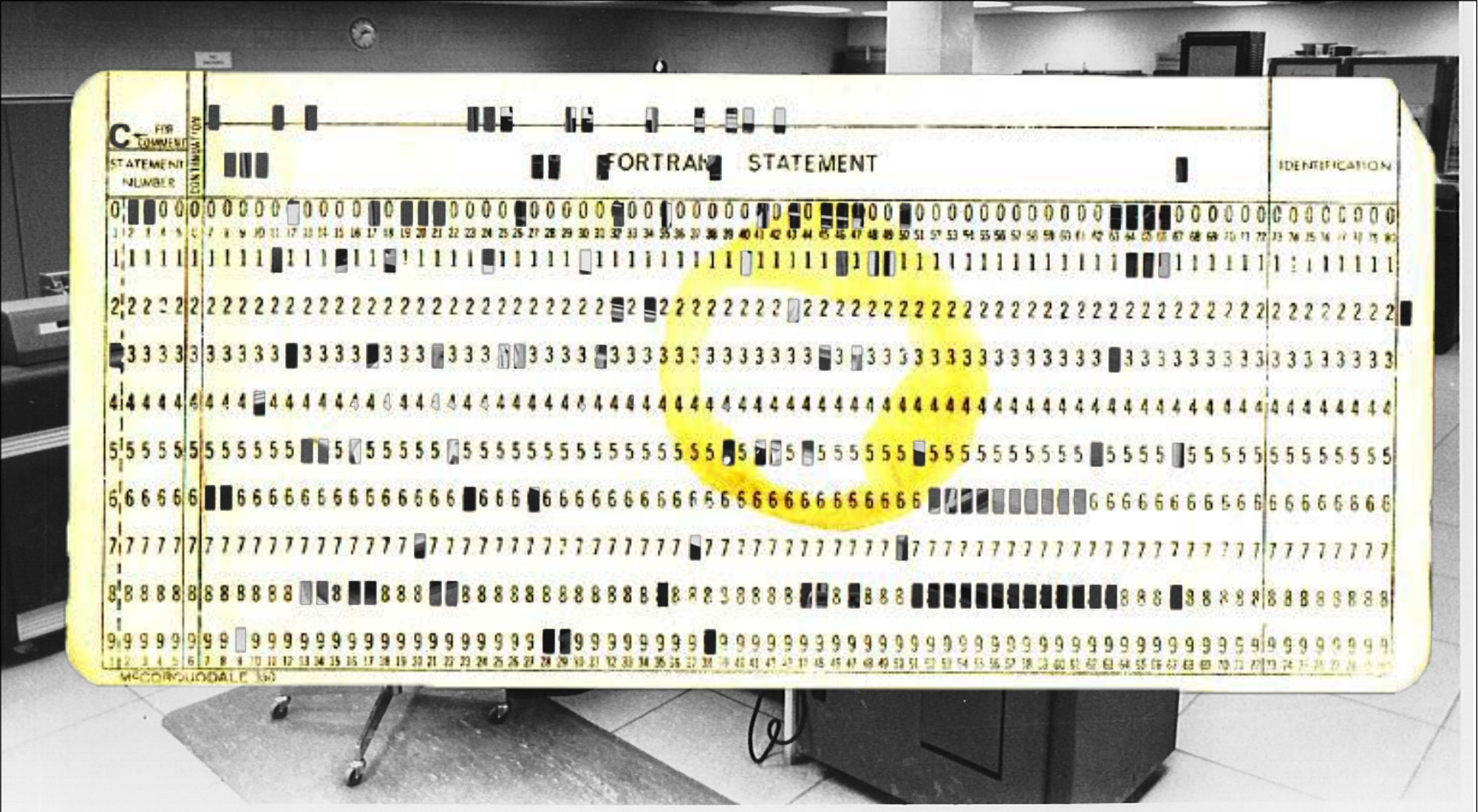
So is everything published since 2006 in PrIMe format?

No. It's mostly in CHEMKIN format.

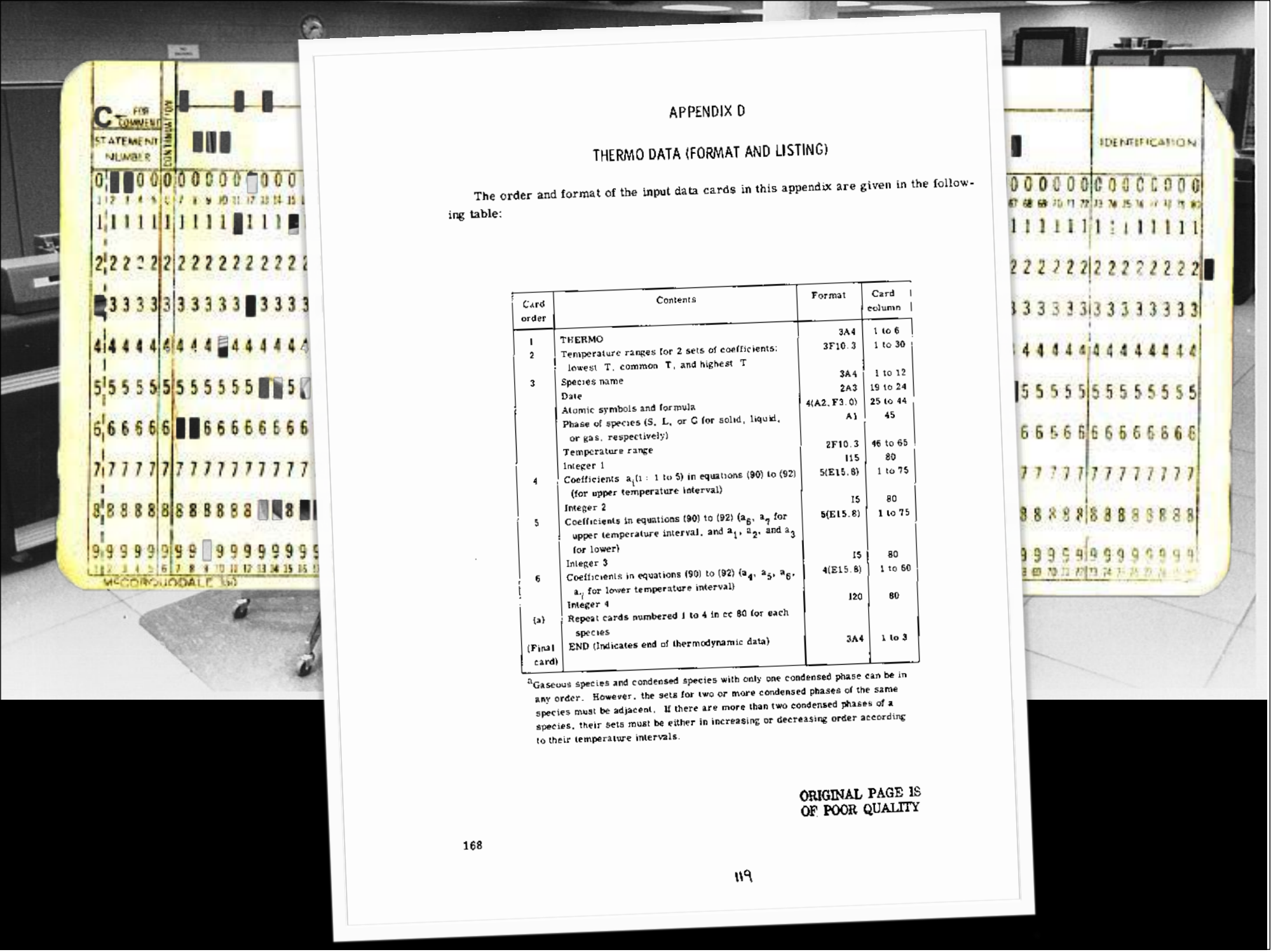
(This chart is very incomplete random sampling)

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

his 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.



his 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.



APPENDIX 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: lowest T, common T, and highest T	3F10.3	1 to 30
3	Species name	3A4	1 to 12
	Date	2A3	19 to 24
	Atomic symbols and formula	4(A2, 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
	Integer 1	I15	80
4	Coefficients a_i(i = 1 to 5) in equations (90) to (92) (for upper temperature interval)	5(E15.8)	1 to 75
	Integer 2	I5	80
5	Coefficients in equations (90) to (92) (a_6, a_7 for upper temperature interval, and a_1, a_2, and a_3 for lower)	5(E15.8)	1 to 75
	Integer 3	I5	80
6	Coefficients in equations (90) to (92) (a_4, a_5, a_6, a_7 for lower temperature interval)	4(E15.8)	1 to 60
	Integer 4	I20	80
(a)	Repeat cards numbered 1 to 4 in cc 80 for each species		
(Final card)	END (Indicates end of thermodynamic data)	3A4	1 to 3

[a]Gaseous 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.

his 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

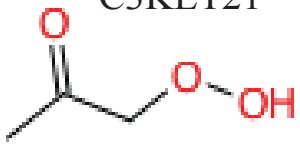
Species Name Chemical Formula Parameters for H(T), S(T)

```
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
```

Space constraints led to “creative” naming schemes

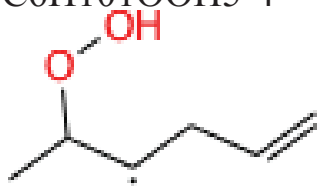
MVOX

C3KET21



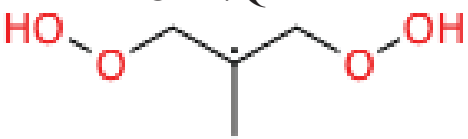
C3H5-A

C6H10JOOH5-4

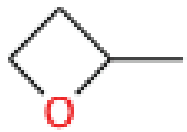


C3H5-S

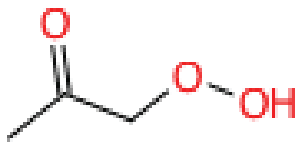
IIC4H7Q2-T



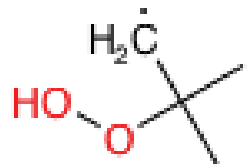
C4H8O1-3



CH3COCH2O2H



TC4H8O2H-I

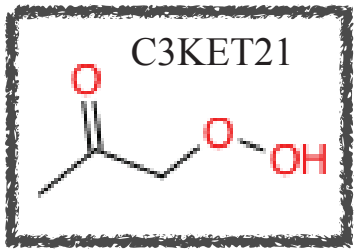


Notice hydroperoxypropan-2-one has two different names, often in the same mechanism file!

Space constraints led to “creative” naming schemes

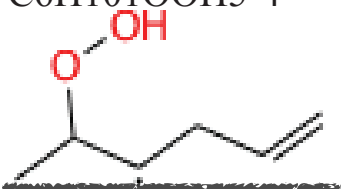
MVOX

C3KET21



C3H5-A

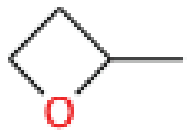
C6H101OOH5-4



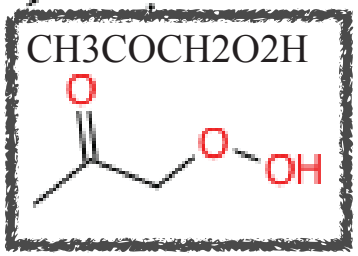
C3H5-S

IIC4H7Q2-T

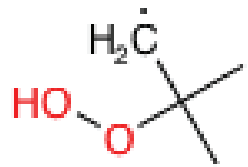
C4H8O1-3



CH3COCH2O2H



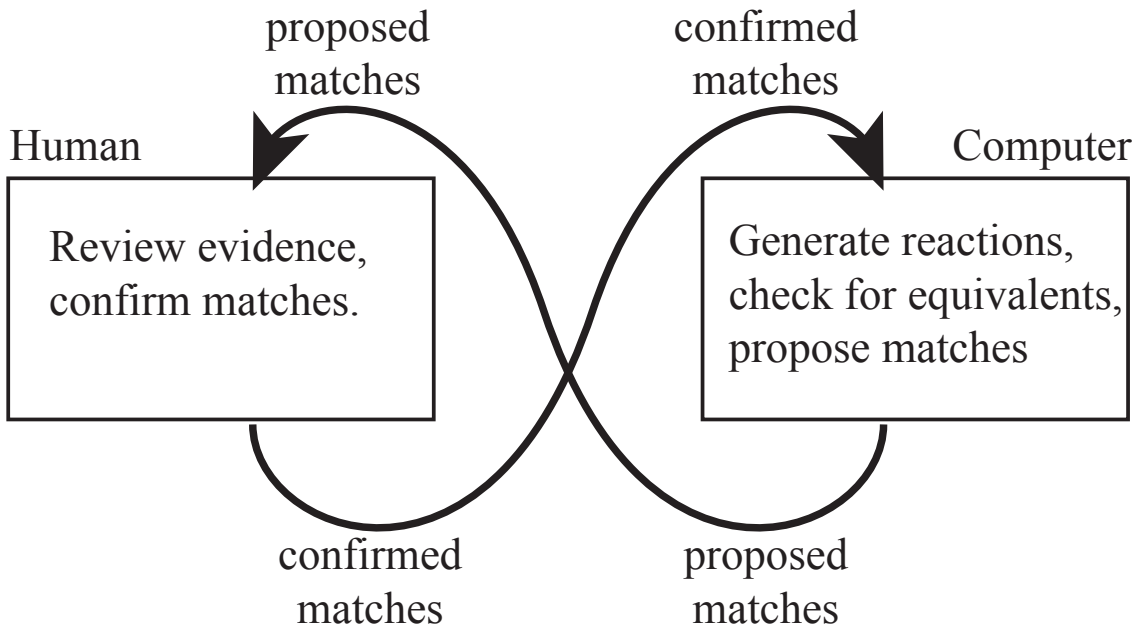
TC4H8O2H-I



- 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



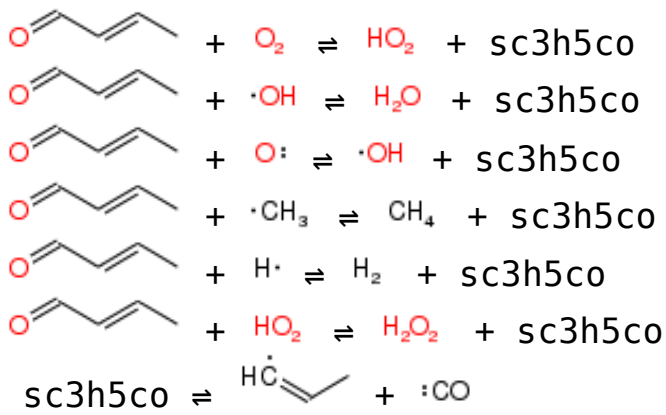
- Our new tool uses RMG to generate reactions to compare with the target model
- A human reviews the evidence and confirms matches.

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

R1	sc3h5cho + o2 ⇌ sc3h5co + ho2
R2	sc3h5cho + oh ⇌ sc3h5co + h2o
R3	sc3h5cho + o ⇌ sc3h5co + oh
R4	sc3h5cho + ch3 ⇌ sc3h5co + ch4
R5	sc3h5cho + h ⇌ sc3h5co + h2
R6	sc3h5cho + ho2 ⇌ sc3h5co + h2o2
R7	sc3h5co ⇌ c3h5-s + co

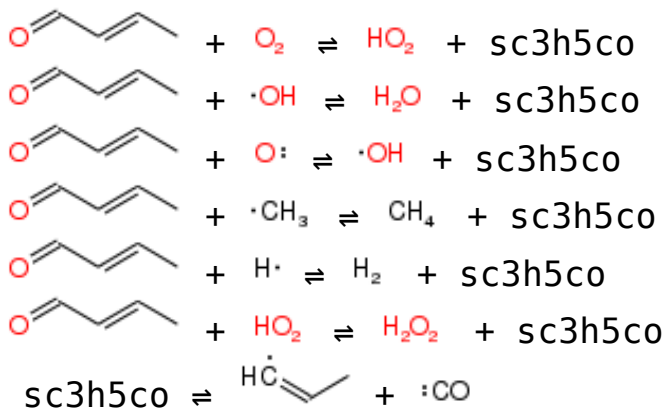


The first 6 reactions are all H-abstractions giving 4 possibilities for sc3h5co.

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

Identifying ‘sc3h5co’ from its reactions

R1	sc3h5cho + o2 ⇌ sc3h5co + ho2
R2	sc3h5cho + oh ⇌ sc3h5co + h2o
R3	sc3h5cho + o ⇌ sc3h5co + oh
R4	sc3h5cho + ch3 ⇌ sc3h5co + ch4
R5	sc3h5cho + h ⇌ sc3h5co + h2
R6	sc3h5cho + ho2 ⇌ sc3h5co + h2o2
R7	sc3h5co ⇌ c3h5-s + co

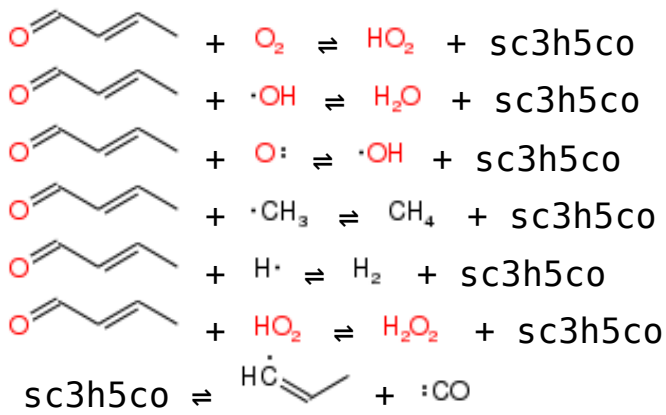


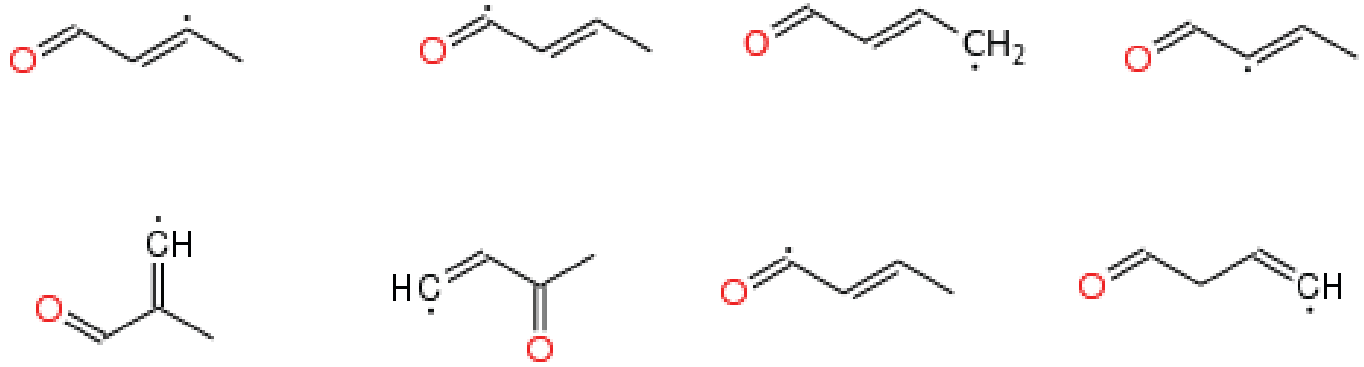
The first 6 reactions are all H-abstractions giving 4 possibilities for sc3h5co.

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

Identifying ‘sc3h5co’ from its reactions

R1	sc3h5cho + o2 ⇌ sc3h5co + ho2
R2	sc3h5cho + oh ⇌ sc3h5co + h2o
R3	sc3h5cho + o ⇌ sc3h5co + oh
R4	sc3h5cho + ch3 ⇌ sc3h5co + ch4
R5	sc3h5cho + h ⇌ sc3h5co + h2
R6	sc3h5cho + ho2 ⇌ sc3h5co + h2o2
R7	sc3h5co ⇌ c3h5-s + co





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

R1 sc3h5cho + o2 ⇌ sc3h5co + ho2

R2 sc3h5cho + oh ⇌ sc3h5co + h2o

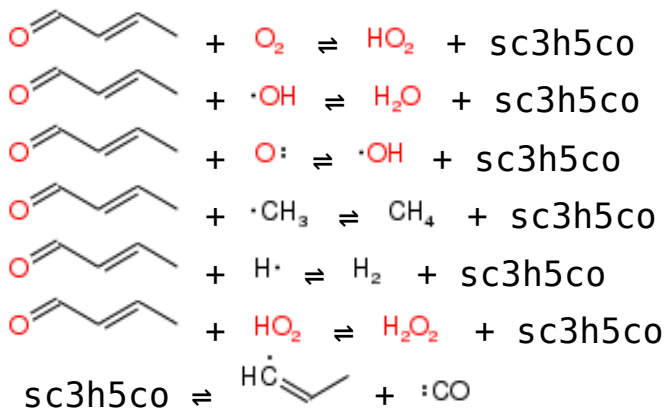
R3 sc3h5cho + o ⇌ sc3h5co + oh

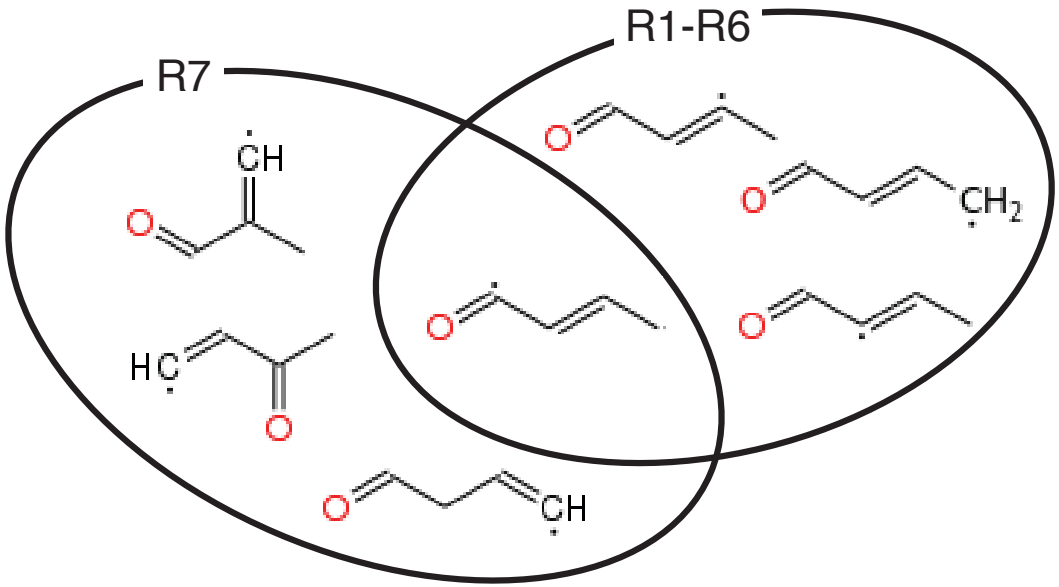
R4 sc3h5cho + ch3 ⇌ sc3h5co + ch4

R5 sc3h5cho + h ⇌ sc3h5co + h2

R6 sc3h5cho + ho2 ⇌ sc3h5co + h2o2

R7 sc3h5co ⇌ c3h5-s + co



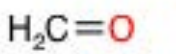
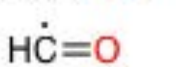
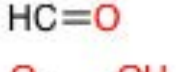
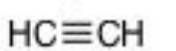


By considering all seven reactions, we can deduce the most likely isomer.

‘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	$\Delta H°_f(298K)$	Matching Thermo	
CH2O		0.9 kJ/mol	GRI-Mech3.0	confirm
HCO		1.0 kJ/mol	-	confirm
HOCHO		0.8 kJ/mol	-	confirm
C2H2		-3.2 kJ/mol	-	confirm
CH2CO		-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.

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.

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 10^{30} 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