9th US National Combustion Meeting
Organized by the Central States Section of the Combustion Institute
May 17–20, 2015
Cincinnati, Ohio

Identification, Correction, and Comparison of Detailed Kinetic Models

Victor R. Lambert[1] and Richard H. West[1,]*

[1] Department of Chemical Engineering, Northeastern University, Boston MA 02115, USA
**Corresponding Author Email: r.west@neu.edu*

For decades, detailed kinetic models have been published in CHEMKIN format, in which each species given a name with a limited number of characters. As models have grown in size it has become harder to uniquely describe each species, and the proliferation of nicknames and naming conventions has made the error-free comparison or merging of models almost impossible.

We have developed tools to facilitate the identification of chemical species in a kinetic model "Chemkin file", and then to allow comparison of the models. The tools, built on top of the Python version of Reaction Mechanism Generator software (RMG-Py), also enable error identification, self-consistency checking, and will facilitate model merging, gap-filling, and model validation and updating.

We are using our tool to import all the models published in supplementary material to the Proceedings of the 34th Combustion Symposium into a unified and cross-referencable database, and to analyze them to reveal agreements, disagreements, consistencies, and areas for the combustion kinetics community to focus on in the future.

1 Introduction

In recent decades, detailed kinetic models have become integral to combustion research. Progress has been significant and the latest models can explain many complicated combustion phenomena, and allow increasingly accurate engine simulations. These models can be very large (eg. the LLNL model for 2-methylalkanes has over 7,000 species and 30,000 reactions [1]) and there are now dozens of large published models. Unfortunately, these ever-proliferating detailed kinetic models are incompatible and inconsistent, are seldom compared directly, and often contain undetected mistakes. We hear these days that it's possible to achieve errors a low as 1 kJ/mol in energy and 30% in rate constants, yet in many published and trusted models these parameters can differ by 100 kJ/mol and 30 orders of magnitude respectively.

Collaborative optimization, solution mapping, and data curation projects such as PrIMe offer great promise [2], but adoption is low and the preferred publication format remains a "CHEMKIN file". This CHEMKIN format [3, 4], devised in the 1970's when input was limited by the width of 80-column punch-cards, forces model-builders to abbreviate species' names, thereby losing their chemical identity, and to discard other metadata. The main challenge in comparing these

models is in recognizing, for example, that the name "C3KET12" in one model represents 1-hydroperoxypropan-2-one, which another research group may have named "CH3COCH2O2H" in a different model.

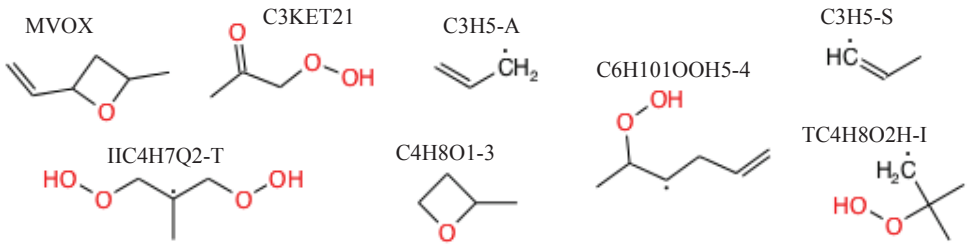

Figure 1: *Some of the species and names from the LLNL model for n-heptane[5]*

There have been some efforts to aid comparison of CHEMKIN files over the years. In 2004 Simmie and co-workers developed some VisualBasic applications: CHEMClean to tidy up the formatting of a CHEMKIN file, and CHEMDiffs to compare two such files for common reactions. They applied this to five methane combustion models, revealing how many reaction rates had been re-used between them. [6] A similar tool from the same group allows users to compare thermodynamics files [7].

More recently Ratkiewicz and Truong [8] proposed a canonical form for reaction mechanisms, useful for merging and comparison of mechanisms, and developed software called CANMECH (CANonizedMECHanism) to translate mechanisms into such a form. They used their tool to compare six methane combustion models, and demonstrated the manipulations that become possible from merging kinetic models and interchanging sub-mechanisms between models. The CANMECH software requires an input file identifying the molecular structure of each species in the CHEMKIN file. They lament that "since there is no universal algorithm of decoding species nomenclature used by different researchers, this step cannot be fully automated" and "may be difficult for some users"[8]. This is the challenge addressed by the current work.

2 Methods

2.1 Importer Tool algorithm

To illustrate the algorithm, consider trying to identify the species named 'sc3h5cho' in a large methly butanoate combustion model [9]. The data available (Figure 2) are a block of thermochemistry data including the chemical formula, and a list of seven reactions that it participates in (from the 1,219 reactions in the model).

The first six reactions are all hydrogen abstractions from but-2-enal (assume for now that this species has already been identified), which has four types of hydrogen atom, implying sc3h5co could be one of four possible radicals. The seventh reaction is the decomposition into propenyl and carbon monoxide, also limiting sc3h5co to four possible species. The Venn diagram in Figure 3 shows that only one species satisfies all seven reactions: but-2-enoyl. Presented with this evidence

Species label Chemical formula: $C_4H_5O_1$ Parameters for H(*T*), S(*T*) expressions

```
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
```

```
sc3h5cho + ho2 = sc3h5co + h2o2     1.000E+12 0.00  1.192E+04
sc3h5cho + ch3 = sc3h5co + ch4      3.980E+12 0.00  8.700E+03
sc3h5cho + o = sc3h5co + oh         7.180E+12 0.00  1.389E+03
sc3h5cho + o2 = sc3h5co + ho2       4.000E+13 0.00  3.760E+04
sc3h5cho + h = sc3h5co + h2         2.600E+12 0.00  2.600E+03
sc3h5cho + oh = sc3h5co + h2o       1.070E+13 0.00 -3.480E+02
sc3h5co = c3h5-s + co               8.600E+15 0.00  2.300E+04
```

Figure 2: *Thermodynamic and kinetic data for the species sC3H5CO.*

in a convenient tabular form, along with clues such as the difference in enthalpy of formation between the CHEMKIN polynomial and a Benson group additive estimate of the proposed species, the user would confirm this match. Once one species has been identified, it is used to help identify other species, much like solving a Sudoku puzzle. The human-computer workflow is shown in Figure 4.

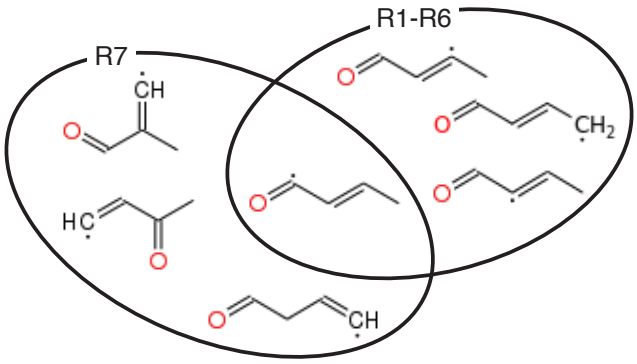

Figure 3: *Possible structures for sC3H5CO based on the evidence in Figure 2*

Our tool is built upon the new Python version of the open-source Reaction Mechanism Generator software, RMG-Py [10]. The modular nature of the RMG-Py software allows us to re-use many classes and methods for molecular graph matching, species generation, reaction templating, property estimation, and model generation. Currently the tool is somewhat difficult to get started, requiring manual intervention and often some debugging of CHEMKIN files (most published files have minor syntax errors), but once a model is running there is a simple web-based interface, allowing multiple users to review the evidence and confirm matches from any device with a web browser.

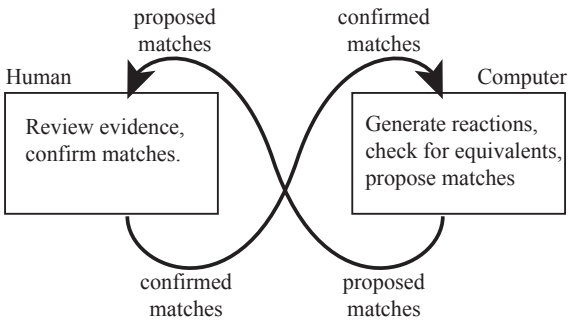

Figure 4: *The workflow benefits from the strengths of a human's intelligence and a computer's speed and working memory.*

2.2 Analysis of Combustion Symposium Proceedings

We set about to analyze The Proceedings of the Combustion Institute 34 (2013). The Reaction Kinetics section contained 60 papers. Of those, 33 mention or use CHEMKIN and 28 were published with some supplementary material, 17 of which include kinetic models of some sort. Of these 17: two are only partial lists of reactions rather than complete kinetic models, so were excluded from this analysis (one of these was provided in PDF format anyway, making it almost impossible to parse and reuse); one provided a FORTRAN subroutine for calculating the rate of generation of each species rather than a typical CHEMKIN file, so cannot be interpreted or compared; and two provided the log file from the CHEMKIN pre-processor 'ckinterp' instead of the input, which is perhaps more human-readable but cannot be easily re-used. We were able to obtain the original input files corresponding to one of the log files by contacting the authors, bringing the number of interpretable kinetic models up to 13, which are listed in Table 1.

Because the tool has not yet been updated with the latest features of RMG-Py, it is limited to reacting species containing C, H, and O atoms; apart from N_2 as a bath gas collider, species containing nitrogen are excluded from the current study.

Table 1: *Models being imported from the Proceedings of the Combustion Institute 34 (2013)*

First author	Page	Ref.	Notes
Somers	225	[11]	2-methyl furan oxidation
Labbe	259	[12]	Flame chemistry of tetrahydropyran
Matsugi	269	[13]	2- and 3-ring aromatics formation in the pyrolysis of toluene
Dagaut	289	[14]	Decalin oxidation and pyrolysis
Herbinet	297	[15]	Low temperature oxidation of benzene and toluene in mixture with n-decane
Husson	325	[16]	Ethylbenzene oxidation
Wang	335	[17]	Autoignition of 3-methylheptane
Malewicki	353	[18]	Pyrolysis and oxidation of iso-octane
Malewicki	361	[19]	Pyrolysis and oxidation of n-decane and n-dodecane
Liu	401	[20]	Reduced biodiesel surrogate
Darcy	411	[21]	Ignition of a mixture of n-heptane and n-propylbenzene
Sheen	527	[22]	Unsaturated hydrocarbons
Veloo	599	[23]	Propanal oxidation

3 Results and Discussion

The largest model we have completely imported is a recently revised gasoline-surrogate model from LLNL, containing 1731 species. Figure 5 shows our progress importing the 13 kinetic models from the Proceedings of the Combustion Institute 34 (2013). Once one or two large models are completed, it will become faster to import future models, because so many data are reused.

3.1 Naming

The importer tool greatly increased the speed of identifying certain species which would have been extremely difficult to decipher by hand. As shown in Table 2, different models have drastically different naming conventions. The CHEMKIN name A1 represents benzene in one model, but the benzyl radical in another. Strange names, such as b13de2m for isoprene, are not uncommon (b13de2m standing for 2-methyl-1,3-butadiene).

Table 2: *Different Names for Species in the Proceedings of the Combustion Institute 34 (2013)*

Species	SMILES	CHEMKIN Names
C_3H_5	`C=[C]C`	C3H5-T, ch3cch2, TC3H5
Benzene	`c1ccccc1`	A1, A, C6H6#, c6h6
Benzyl	`c1[c]cccc1`	A1J, A1, C6H5#, c6h5
Butyl	`[CH2]CCC`	PC4H9, R20C4H9, NC4H9, NC4H9P
2-Hexyl	`CCCC[CH]C`	R72C6H13, hex2yl, C6H13-2
Isoprene	`C=CC(=C)C`	b13de2m, IC5H8

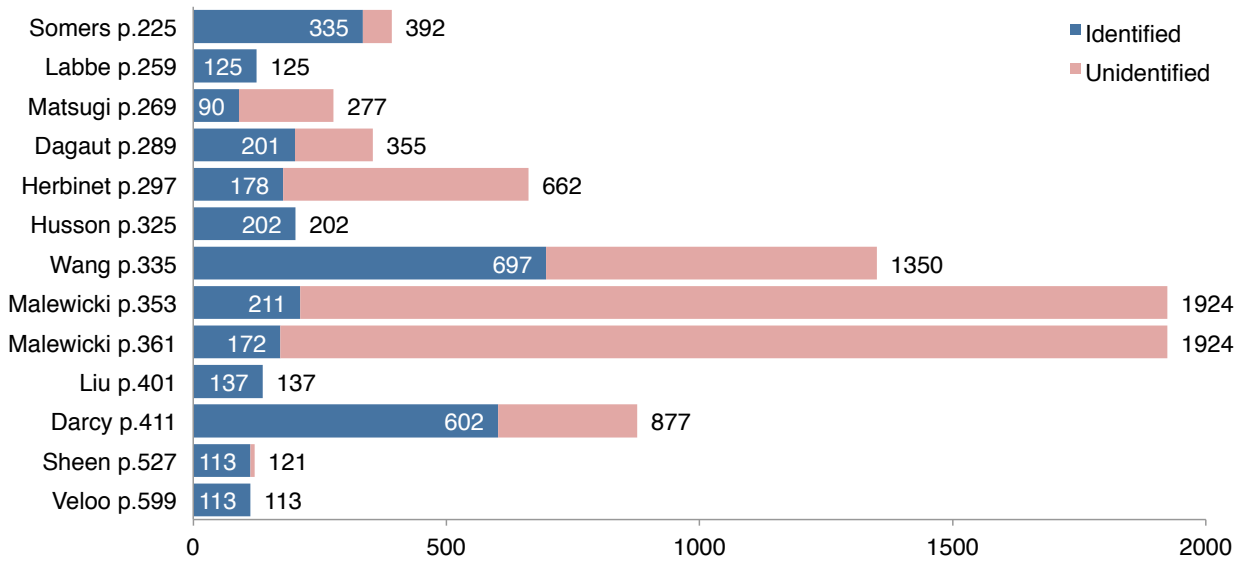

Figure 5: *Models being imported from the Proceedings of the Combustion Institute 34 (2013), showing number of species identified at time of writing, and total model size (C/H/O species only).*

3.2 Thermochemistry

To compare thermochemistry, we evaluate the standard enthalpy of formation at 298 K in kJ/mol. Of the 339 species identified in two or more models, 93 have perfect agreement and 132 (40%) have less than a 1 kJ/mol disagreement; however 62 (18%) have a range exceeding 10 kJ/mol. Some species with a particularly large range of enthalpies are listed in Table 3.

3.3 Kinetics

To compare kinetics, we evaluated each rate coefficient at 1000 K, and took the base-10 logarithm. Several interesting discrepancies are noticeable when comparing the reaction rates for different models.

Some rates for hydrogen abstractions by oxygen are clearly disagreed upon, varying by about thirty orders of magnitude. Most have some rates around 10^{-19} $cm^3/mol/s$ and others around 10^{11} $cm^3/mol/s$. Some reactions with large discrepancies of this sort are in Table 4.

$C_2H_5 + O_2 \rightleftharpoons C_2H_4 + HO_2$ occurs in nine models and has a range of 3 orders of magnitude, with $\log(k/(cm^3/mol/s))$ ranging from 8.2 to 11.

Though many reactions have a large disparity, it is important to note that 260 of 772 reactions have a range of less than 0.15 orders of magnitude, or 40%.

Table 3: *Species with large discrepancies in enthalpies of formation*

SMILES	Name	Num. models	Range of ΔH° (kJ/mol)
`[CH]=C=[CH]`	C_3H_2	10	96.6
`CC(=O)O[O]`	$C_2H_3O_3$	8	56.2
`[CH2]CC=O`	C_3H_5O	8	96.3
`CC=C[C]=O`	C_4H_5O	7	56.9
`[CH2]C#CC`	C_4H_5	6	89.0
`CC=C=O`	C_3H_4O	6	76.0
`C[C](C)C=O`	C_4H_7O	5	70.1
`[CH2]CCC=C`	C_5H_9	5	77.2
`C=CC[CH]C`	C_5H_9	5	66.1
`C#CC#CC#C`	C_6H_2	5	184.9

Table 4: *Reactions with large discrepancies in rate coefficient at 1000 K*

Reaction	Num. models
$IC_3H_7 + O_2 \rightleftharpoons C_3H_6 + HO_2$	7
$SC_4H_9 + O_2 \rightleftharpoons C_4H_8 + HO_2$	5
$NC_3H_7 + O_2 \rightleftharpoons C_3H_6 + HO_2$	4
$PC_4H_9 + O_2 \rightleftharpoons C_4H_8 + HO_2$	4
$AC_3H_5 + HO_2 \rightleftharpoons C_2H_3 + CH_2O + OH$	4
$C_6H_{13} + O_2 \rightleftharpoons C_6H_{12} + HO_2$	3
$C_5H_{11} + O_2 \longrightarrow C_5H_{10} + HO_2$	3

4 Conclusions

A tool for comparing and decoding CHEMKIN files has been demonstrated and successfully used to compare several large kinetic models. This tool will become more efficient as it is further utilized, resulting in easier model comparison and correction. We are developing a publicly accessible web interface to facilitate using the tool; email the authors for news in this regard.

5 Acknowledgements

We thank contributions from Esma Goktekin, Anthony J. Silvi, Pierre L. Bhoorasingh, Belinda L. Slakman, Farbia Seyedzadeh Khanshan, Sophie Podhurst, and Olusade Mabawonku. This material is based upon work supported by the National Science Foundation under Grant No. 1403171.

6 References

[1] S. M. Sarathy, C. K. Westbrook, M. Mehl, W. J. Pitz, C. Togbe, P. Dagaut, H. Wang, M. A. Oehlschlaeger, U. Niemann, K. Seshadri, P. S. Veloo, C. Ji, F. N. Egolfopoulos, T. Lu, Combust. Flame 158 (2011) 2338–2357.
[2] M. Frenklach, Proc. Combust. Inst. 31 (2007) 125–140.
[3] S. Gordon, B. J. McBride, Computer Program for Calculation of Complex Chemical Equilibrium Composition, Rocket Performance, Incident and Reflected Shocks and Chapman-Jouguet Detonations, Technical Report, Glenn Research Center, Cleveland, OH, 1976.
[4] R. J. Kee, J. A. Miller, T. H. Jefferson, CHEMKIN: A general-purpose, problem-independent, transportable, FORTRAN chemical kinetics code package, Technical Report SAND80-8003, 1980.
[5] M. Mehl, W. J. Pitz, C. K. Westbrook, H. J. Curran, Proc. Combust. Inst. 33 (2011) 193–200.
[6] S. Rolland, J. M. Simmie, Int. J. Chem. Kinet. 36 (2004) 467–471.
[7] S. Rolland, J. M. Simmie, Int. J. Chem. Kinet. 37 (2005) 119–125.
[8] A. Ratkiewicz, T. N. Truong, Energy 43 (2012) 64–72.
[9] E. Fisher, W. J. Pitz, H. J. Curran, C. K. Westbrook, Proc. Combust. Inst. 28 (2000) 1579–1586.
[10] W. H. Green, R. West, J. W. Allen, others, RMG-Py Reaction Mechanism Generator open-source software package (Python version), `http://greengroup.github.com/RMG-Py/` 2013.
[11] K. P. Somers, J. M. Simmie, F. Gillespie, U. Burke, J. Connolly, W. K. Metcalfe, F. Battin-Leclerc, P. Dirrenberger, O. Herbinet, P. A. Glaude, H. J. Curran, Proc. Combust. Inst. 34 (2013) 225–232.
[12] N. J. Labbe, V. Seshadri, T. Kasper, N. Hansen, P. Oßwald, P. R. Westmoreland, Proc. Combust. Inst. 34 (2013) 259–267.
[13] A. Matsugi, A. Miyoshi, Proc. Combust. Inst. 34 (2013) 269–277.
[14] P. Dagaut, A. Ristori, A. Frassoldati, T. Faravelli, G. Dayma, E. Ranzi, Proc. Combust. Inst. 34 (2013) 289–296.
[15] O. Herbinet, B. Husson, M. Ferrari, P.-A. Glaude, F. Battin-Leclerc, Proc. Combust. Inst. 34 (2013) 297–305.
[16] B. Husson, M. Ferrari, O. Herbinet, S. S. Ahmed, P.-A. Glaude, F. Battin-Leclerc, Proc. Combust. Inst. 34 (2013) 325–333.
[17] W. Wang, Z. Li, M. A. Oehlschlaeger, D. Healy, H. J. Curran, S. M. Sarathy, M. Mehl, W. J. Pitz, C. K. Westbrook, Proc. Combust. Inst. 34 (2013) 335–343.
[18] T. Malewicki, A. Comandini, K. Brezinsky, Proc. Combust. Inst. 34 (2013) 353–360.
[19] T. Malewicki, K. Brezinsky, Proc. Combust. Inst. 34 (2013) 361–368.
[20] W. Liu, R. Sivaramakrishnan, M. J. Davis, S. Som, D. E. Longman, T. F. Lu, Proc. Combust. Inst. 34 (2013) 401–409.
[21] D. Darcy, M. Mehl, J. M. Simmie, J. Würmel, W. K. Metcalfe, C. K. Westbrook, W. J. Pitz, H. J. Curran, Proc. Combust. Inst. 34 (2013) 411–418.
[22] D. A. Sheen, C. M. Rosado-Reyes, W. Tsang, Proc. Combust. Inst. 34 (2013) 527–536.
[23] P. S. Veloo, P. Dagaut, C. Togbe, G. Dayma, S. M. Sarathy, C. K. Westbrook, F. N. Egolfopoulos, Proc. Combust. Inst. 34 (2013) 599–606.