European Symposium on Computer Aided Process Engineering – 15
L. Puigjaner and A. Espuña (Editors)

New approaches for representing, analyzing and visualizing complex kinetic mechanisms

Ioannis P. Androulakis[a,b,*]
[a]Biomedical Engineering Department
[b]Chemical & Biochemical Engineering Department
Rutgers, The State university of New Jersey
Piscataway, NJ 08854, USA

Abstract

Complex kinetic representations involving thousands of reacting species and tens of thousands of reactions are currently required for the rational analysis of modern combustion systems. In order to represent, analyze and visualize effectively the ignition processes advanced computational techniques will be required. Recently, we introduced a novel concept that captured the principal elemental transformations in complex kinetic environments in the form of graphs. In order to arrive at a compact representation of the information content of these graphs novel concepts from machine learning, such as feature selection in time series and hashing, are implemented. These approaches allow the projection of the totality of the information contained in the graph describing the chemical transformations onto a single scalar. The temporally evolving graphs are treated as streaming data and locality-preserving hashing allows the unique assignment of a scalar "motif" value to each such graph. Analysis of those motifs allows the quick identification of "clusters" of identical reaction graphs that correspond to regimes with similar kinetic histories. The approach is illustrated with highly complex kinetic mechanisms describing pentane autoignition. It is demonstrated how this novel representation allows the quick identification of regions where similar temporal history of the chemical transformations is experienced.

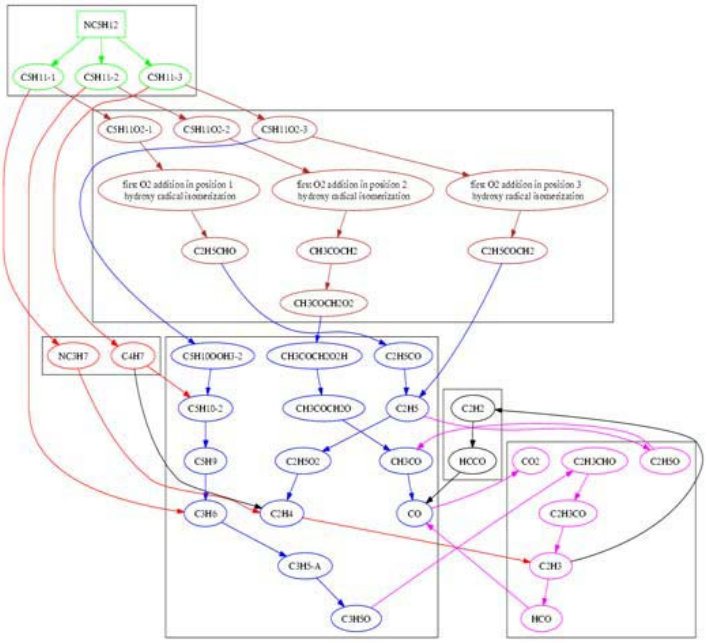
Keywords: complex kinetic mechanisms, combustion, visualization

1. The complexities of combustion

Hydrocarbon oxidation, one of the leading chemical transformations in transportation applications, is a process easily stated, yet very complicated to understand in detail and even more intricate to properly model. Macroscopic models of combustion, such as the so-called Shell Model, have been used for the longest time in order to predict the rate of heat release or the rate of consumption/production of the main fuel/product molecules (Crua et al., 2004; Halstead et al., 1977). However, due to environmental concerns combustion models are revisited to address critical issues such as predictive response

* Author to whom correspondence should be addressed: yannis@rci.rutgers.edu

over extended ranges of operation, and increased accuracy requirements at the ppm level. It has been realized that macroscopic characteristics are inadequate and therefore a detailed description of combustion chemistry is needed (Gardiner, 1996). A distinguishing feature of the combustion of many hydrocarbons is the "negative temperature coefficient", NTC (Pilling ad Hancock, 1997). At very low temperatures, as the temperature is raised the ignition delay of a fuel is decreased. However, for many fuels, at a certain temperature the ignition delay reverses its coarse and becomes longer as the temperature is raised. At some higher temperature the ignition delay decreases again. In order to capture all the intricate details of fuel-like molecules (Griffiths, 1995) mechanisms of significant size are required, potentially involving thousands of species and tens of thousands of reactions (Curran et al., 1998). On of the key complications is that different temperature regimes are known to activate different reaction pathways.



Multi-point ignition in homogeneous charge compression ignition engines is a great manifestation of the impact of chemistry induced due to spatial inhomogeneities in composition, temperature and pressure. The effects of non-homogeneous fuel-air mixtures are key to understanding the nature of multi-point ignition and further conceptualize potential embodiments that enhance or control the simultaneous ignition of the mixture, (Echekki, 2003).

2. Flux diagrams of complex kinetic mechanisms

For the description of complex chemical transformations, we recently introduced the concept of time-integrated element flux analysis that allows the on-line generation of

graphs that monitor the elemental transformations. The concept was successfully applied in the analysis, reduction and generation of complex kinetic mechanisms (Androulakis et al., 2004). For a selected element, the algorithm traces the transformation between all reactive species and source-sink pairs are defined as a function of time. Hence, we construct a complicated graph with nodes the reactive species and arcs weighted proportional to the rate of element flux from a source to a sink.

Figure 1 summarizes the time evolution of the reaction pathways during the oxidation of pentane from a low initial temperature. The "brown" arcs identify the low temperature initiation pathways, as temperature increases the NTC regime is reached indicated by "blue", subsequent T increase activates high temperature oxidation pathways indicated by the "purple" and "red" arcs. The coupling of the time-integrated element flux analysis and graph visualization software (Gasner et al., 2002) allows the visual inspection of the reaction pathways. If visualized in a dynamic mode the graph changes the arc weights as a function of time and depending on the pathways that are active. The oxygen addition reaction to the fuel radicals is indicative of low temperature chemistry, NTC pathways become significant as T increases and eventually, at high T, the direct fuel decomposition reactions dominate.

3. A novel representation of reaction flux diagrams

The novelty of our approach is that we do not simply define a static graph representation of a chemical mechanism but rather the dynamic evolution of the reaction pathways, thus enabling the monitoring of the activation/deactivation of critical pathways as a function of time. This graph representation of the evolution of the kinetic transformation is a very powerful representation in studying the spatio-temporal sensitivities of a reaction process. Ultimately, we wish to identify the impact of local

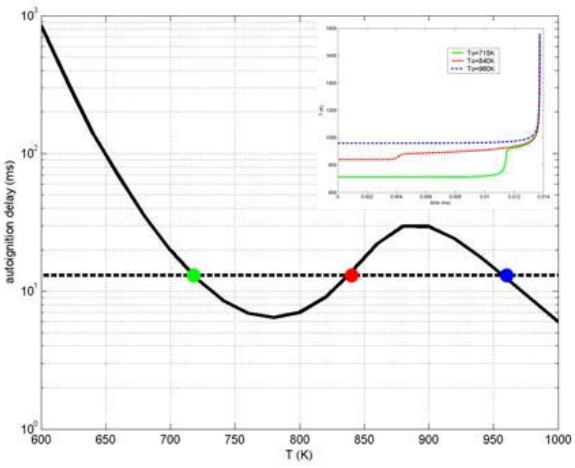


Figure 2: n-Pentane autoignition delay

conditions (flow field, geometry) on the chemistry. Having a succinct representation of the chemistry is a very critical step. Although the flux diagrams do capture such

dependencies, they are still too cumbersome to work with. What is therefore needed is a novel way of capturing the totality of the graph representation information, via the definition of some novel informative attribute, in order to enable the analysis of spatio-temporal emerging patterns. In this paper we propose a novel graph representations based on recent advances in the area of proximity preserving hashing.

Visualizing graph evolution and finding frequent patterns in graphs is an area that is attracting significant interest (Abello and Kom, 2000). However, in these cases the structure and connectivity of the graph is analyzed. In studying kinetics the connectivity is always the same (superset of all physically possible source-sink species based on the mechanism) but the strength of the connectivity (weight of an arc which is

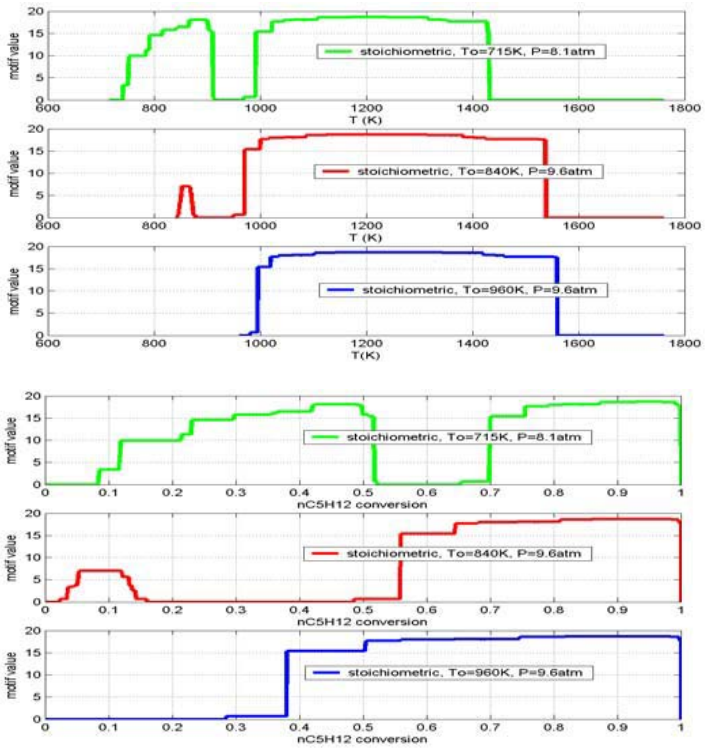


Figure 3: Motif evolution as a function of temperature and fuel conversion

"proportional" to the elemental transformation from one species to another) is evolving over time. Therefore, in order to capture the information content of the graph we need to model the spatio-temporal evolution of the connecting weights. The generated graphs can potentially have an enormous number of arcs denoting mass transformations between species. The PI has previously illustrated (Androulakis et al., 2004) how to identify a, still large, core of important arcs for monitoring. Therefore, at every point in space and time, during the course of a calculation, the element flux analysis will generate a number of arcs with their corresponding weights. The issue therefore becomes how to "compare" large number of such "graphs" efficiently and effectively. For a mechanism with N species the maximum number of graph entries is N*N. Each

graph, evaluated at time t, can be easily transformed from a matrix to a vector representation as in $V[(i-1)*N+j,t]=arc(i,j,t); i,j=1,...,N; t=1,...$. Therefore, each graph in time is assigned a point in a multidimensional space. Monitoring the graphs in time corresponds to a problem of analyzing points in this multidimensional space.

What is proposed in this work is to explore concepts from proximity preserving hashing (Indyk et al., 1997). Hashing is a concept widely used in computer science that assigns an identifier to a sequence of integers (Knuth, 1997). The concept of "piece-wise aggregate approximation" (Lin et al., 2002) is used to transform the arcs of the element flux graphs, in the form of the vector V defined earlier, into a sequence of integers. These are subsequently hashed to a unique identifier (motif) with the property that "nearby graphs", i.e. graphs with similar arc weight values, hash to similar values. The procedure is based on the methodology recently proposed by (Lin et al., 2002) for the analysis of multi-dimensional time series. In short, the temporal evolution of the graphs arcs are transformed according to the following steps:

1. Graph elements are normalized to N(0,1) using the standard transformation $(x-\mu)/\sigma$
2. The normalized temporal values of the graph weights are now considered to be elements of a multidimensional time series. The series are approximated by piece-wise linear components of length “w” and each segment is represented by the corresponding time-window average
3. Given that the normalized distributions have Gaussian characteristics, breakpoints of equal probability can be derived. For instance: cumulative probability in the intervals $(-\infty, -0.43)$, $(-0.43, 0.43)$ and $(0.43,+\infty)$ are all equal. Each segment is assigned a unique symbol. By doings so the original expression signal has been transformed into a sequence of symbols of given alphabet (α)
4. Once the graph elements have been transformed to a unique sequence of symbols, the sequence is “hashed” based on a locality-preserving using function. A starting point will be the hash function provided in (Lin et al., 2002): $h(C,w,\alpha)=1+\sum_{i=1}^{w}(ord(\hat{c}_i-1)*\alpha^{i-1}$. Locality preserving hashing is among the leading topics in computer science due to its significant relevance in searching multi-dimensional spaces.

The principle is demonstrated with a highly complicated pentane oxidation mechanism (Curran et al., 1998). The mechanism has been designed to properly account for low temperature oxidation, negative temperature coefficient regime and high temperature oxidation. The mechanism captures the nonlinear dependence of autoinignition delay as function of the initial temperature, *Figure 2*. Three initial conditions (715K, 840K, and 960K) are selected for analysis because they all generate the same autoignition delay but follow different induction chemistries. For each case we generate the element flux graphs and use the arc weights to identify the temporal evolution of the hash value assigned to the graph. The results of *Figure 3* are quite interesting. The reduced representation using a single scalar correctly identifies that "different" induction chemistries are active in these three cases. Furthermore, once the oxidation process has progressed all three cases should eventually give rise to similar high temperature

oxidation chemistry, corresponding to motif value of about 20. This novel representation of the chemistry has addressed a critical issue: how to represent and compare in a very compact manner the evolution of chemistry in order to identify regimes that follow similar evolution in terms of the active reaction pathways. Unlike previous approaches we monitor in space and in time not simply individual species (Koegler, 2001) but a metric that captures the chemistry history. It is important to realize that monitoring species concentrations would have generated a very different, and incorrect, picture. As seen from the results of *Figure 3* the fuel conversion levels are quite different when the same type of "high temperature chemistry" is reached in each case. By monitoring species profiles one may have reached the conclusion that because different local conditions gave rise to ignition they may correspond to distinct mechanisms. In reality the local conditions are such that the same chemistry, as defined by the important reaction pathways, is activated. This is independent of the level of concentration of any radical or molecular species. The approach proposed defines a comprehensive metric that captures the totality of the chemical transformations in a significantly reduced domain.

4. Conclusions and future work

Significant computational issues remain to be addressed, as we would like to optimize the search and improve the representation. The expected size of the graph describing the mass exchange among various species in the mechanism is expected to be extremely large. Alternative graph analysis, proximity preserving hashing and soft computing methods will be explored for representing and characterizing the evolution of the kinetic transformations. Advances in nonlinear projections and self-organizing maps for reducing the dimensionality and enabling visualization and mining of time varying large data sets will be investigated.

References

Abello, J., and J. Korn, 2000, Visualizing massive multi-graphs, Proc. IEEE Symp. Information Visualization - InfoVis'00

Androulakis, I.P., J.M. Grenda, and J.W. Bozzelli, 2004, AIChE J., 11,2956

Crua, C., D.A. Kennaird, S.S. Sazhin, and M.R. Heikal, 2004, Int. J. Engine Res., 5

Curran, H.J., W.J. Pitz, C.K. Westbrook, C.V. Callahan and F.L. Dryer, 1998, 27th Symp. Comb., 379-385

Echekki, T., and J.H. Chen, 2003, Combustion and Flame, 134:169-192

Gardiner, W.C., 1996, Gas-phase combustion chemistry, Springer-Verlag

Gasner, E., E. Koutsofios, and G. North, 2002, Drawing graphs with dot - dot User's manual

Griffiths, J.F., 1995, Prog. Energy Combust. Sci., 21,27

Halstead, M.P., L.J. Kirch, and C.P. Quinn, 1977, Combust. Flame, 30,46

Indyk, P., R. Motwami, P. Raghavan and S. Vempala, 1996, Proceedings of the 29th Annual ACM Symposium on Theory of Computing, 618-625

Knuth, D.E., 1997, The art of computer programming, Addison-Wesley

Koegler, W.S., 2001, Case study: Application of feature tracking to analysis of autoignition simulation data, IEEE Visualization Conference, San Diego, CA

Lin, J., E. Keogh, S. Lonardi, and P. Patel, 2002, SIGKDD'02, Edmonton, Canada

Pilling, M.J., and G. Hancock (editors), 1997, *Low temperature combustion and autoignition*, Elsevier