

From Macroscopic World to Microscopic World through Mazes of Process Graphs and from Microscopic World to Mesoscopic World through Drunkards' Paths

**Computing in Chemical Engineering Award
Lecture**

**Copyright © by L. T. Fan 2004
All rights reserved.**

**Department of Chemical Engineering
Kansas State University
Manhattan, Kansas 66506, U.S.A.
Phone: (785) 532-5584
Fax: (785) 532-7372
E-mail: fan@cheme.ksu.edu**

**Computing and Systems Technology (CAST) Banquet
November 18, 2003
AIChE Annual Meeting
San Francisco, CA, November 16 – 21, 2003**

Abstract

This is a narrative of my 2003 Computing in Engineering Award lecture. It essentially comprises two parts: graph-theoretical approach to process-network synthesis and stochastic analysis and modeling of random phenomena in process systems. These areas are two foci of my research and teaching endeavors in recent years.

Table of Contents

	Page
Abstract	i
Introduction	4
From Macroscopic World to Microscopic World through Mazes of Process	
Graphs	5
Pre-process-graphs: early 1950's ~ late 1980's	5
Meeting of minds: 1989 ~ '90	6
First milestone: early 1990's	6
Second milestone: mid 1990's	9
Third milestone: mid 1990's ~ late '90's	11
Fourth milestone: late 1990's ~ early 2000	11
Fifth milestone: early 2000	12
Current status and future prospect of P-graphs	12
From Microscopic World to Mesoscopic World through Drunkards' Paths	13
Awakening at dawn: 1950's	13
Encounter with Markov: mid 1960's ~ early 1970's	13
Encounter with Prigogine: early 1980's	15
Encounter with van Kampen: mid 1980's	16
Current status and future prospect of stochastic analysis and modeling	17
Concluding Remarks	18
Acknowledgements	18
Attachment 1. List of Publications on P-graphs by L. T. Fan and Collaborators	
Attachment 2. List of Publications on Stochastic Analysis and Modeling by L. T. Fan and Collaborators	
Attachment 3. Abbreviated or Condensed View Graphs of Three Papers on P-graph Presented at the Annual Meeting of AIChE, San Francisco, CA, November 16-21, 2003	

Attachment 4. Abbreviated or Condensed View Graphs of Two Papers on Stochastic Analysis and Modeling, One Recently Published and the Other Presented at the Annual Meeting of AIChE, San Francisco, CA, November 16-21, 2003

Introduction

The award citation states, “For broad and outstanding contributions to the analysis, synthesis, and control of process and material systems.” This citation implies that my contributions have encompassed many areas of “Computing and Systems Technologies” in Chemical Engineering in the nearly 50 years of my professional career. These areas include: 1. Computer-Aided Process Synthesis, Design, and Optimization; 2. Process Identification, Dynamics and Control; 3. Stochastic, Statistical, Fractal, and Chaos Analyses; and 4. Modeling, Simulation and Numerical Solution. The list of my publications submitted in support of my nomination contains 5 books, 1 book chapter, and 374 refereed journal articles in these areas. I am at the ripe “old (?) or young (?)” age of 74, and am half-retired; however, I continue to work at about the same pace as when I was a full-time faculty member of 40 years, chairing the department for exactly 30 years.

The title of my lecture reflects the sub-areas of two areas, specifically the first and third, on which my collaborators and I have been focusing in recent years. The graph-theoretic approach originally established by my collaborators and me for the optimal synthesis of process networks has been extended eventually to the identification of catalytic-reaction or metabolic pathways through mimicking their synthesis in nature. Obviously, the process networks are macroscopic, and the catalytic-reaction or metabolic pathways are microscopic, thus, the front part of the title. While our group’s effort on stochastic analysis and modeling was originally prompted by my interest in the motion of molecular species and the reactions among them and while some meaningful results were obtained, we came to the realization that it would be far more fruitful to concentrate our effort mainly on the motion and behavior of gas bubbles, liquid droplets or solid particles and the interactions among them. Clearly, the molecular species are microscopic, and the gas bubbles, liquid droplets or solid particles are mesoscopic, thus, the back part of the title. It is worth pointing out that these mesoscopic objects are ubiquitous in a wide range of process systems and in nature. Attachment 1 lists our publications pertaining to the front part; and Attachment 2, the back part.

From Macroscopic World to Microscopic World through Mazes of Process Graphs

This section chronicles in terms of several periods of my involvement in the evolution of the graph-theoretic approach to process-network synthesis on the basis of process graphs, or P-graphs in short. Brief descriptions are given of some memorable events and breakthroughs that occurred during these periods. Also given are basic definitions and concise discourses of concepts whenever deemed appropriate.

Pre-process-graphs: early 1950's ~ late 1980's

Ever since I received my B.S. in 1952, I have been engaged in the conceptual design and optimal synthesis of processes and process plants, first in industry in “Ilha Formosa” (a beautiful island, i.e., Taiwan) and then in a U.S. government laboratory. This was followed by my long academic and consulting career continually engaged in the conceptual design and optimal synthesis of processes and process plants. Until the late 1980's, I have resorted to every conceivable means or method available for design and optimization, including heuristics, dynamic programming, structural parameter, super-structure with structural parameters aided by various search techniques, and conventional graphs. It has been gratifying to know that some of the processes conceived or conceptualized by me have been industrially implemented and that some books as well as a number of articles have been published based on the works of my collaborators and me. In spite of these successes, I was constantly frustrated by our group's incapability to extend our work to large-scale processes, characterized by complex networks, i.e., flowsheets. At the same time, I was aware that this incapability was attributable to the lack of efficient computational tools; non-linearity of mathematical models of processes or process plants; and combinatorial complexity, which is exponential, in synthesizing large process flowsheets. What we needed was a robust and highly efficient algorithm for process-network synthesis to overcome these obstacles.

Meeting of minds: 1989 ~ '90

My plea for the aforementioned robust and highly efficient algorithm for process-network synthesis was answered when I met Dr. F. Friedler at the North American-German Workshop on Chemical Engineering Mathematics, held in Göttingen, Germany, July 18 to 23 of 1989. I was there as a lecturer; and Dr. Friedler together with his wife, Dr. K. Tarjan, a mathematician, were participants. At that time, Dr. Friedler was a young researcher affiliated with The Institute of Technical Chemistry of the Hungarian Academy of Sciences located in Veszprém; he was already a full-fledged mathematician-computer scientist. According to Dr. Friedler, he closely followed my work on process optimization and synthesis. Hence, he could feel my frustration in my inability to deal with large-scale systems. Meanwhile, since he was neither a chemist nor a chemical engineer, he himself had been struggling mightily to establish a formal framework for the algorithm to execute optimal process-network synthesis in terms of a set of axioms couched in the parlance of chemical, or material, transformation. His collaboration with his colleagues at the Institute was not sufficiently fruitful; his attempt to initiate joint research, also in this regard, with some of Europe's leading process-systems engineering groups came to naught. After Dr. Friedler's presentation at the Workshop and an ensuing short discussion between us, we sensed each other's need and immediately decided to collaborate. Both he and his wife spent a year with me as research associates. The rest is history - at least up to now.

First milestone: early 1990's

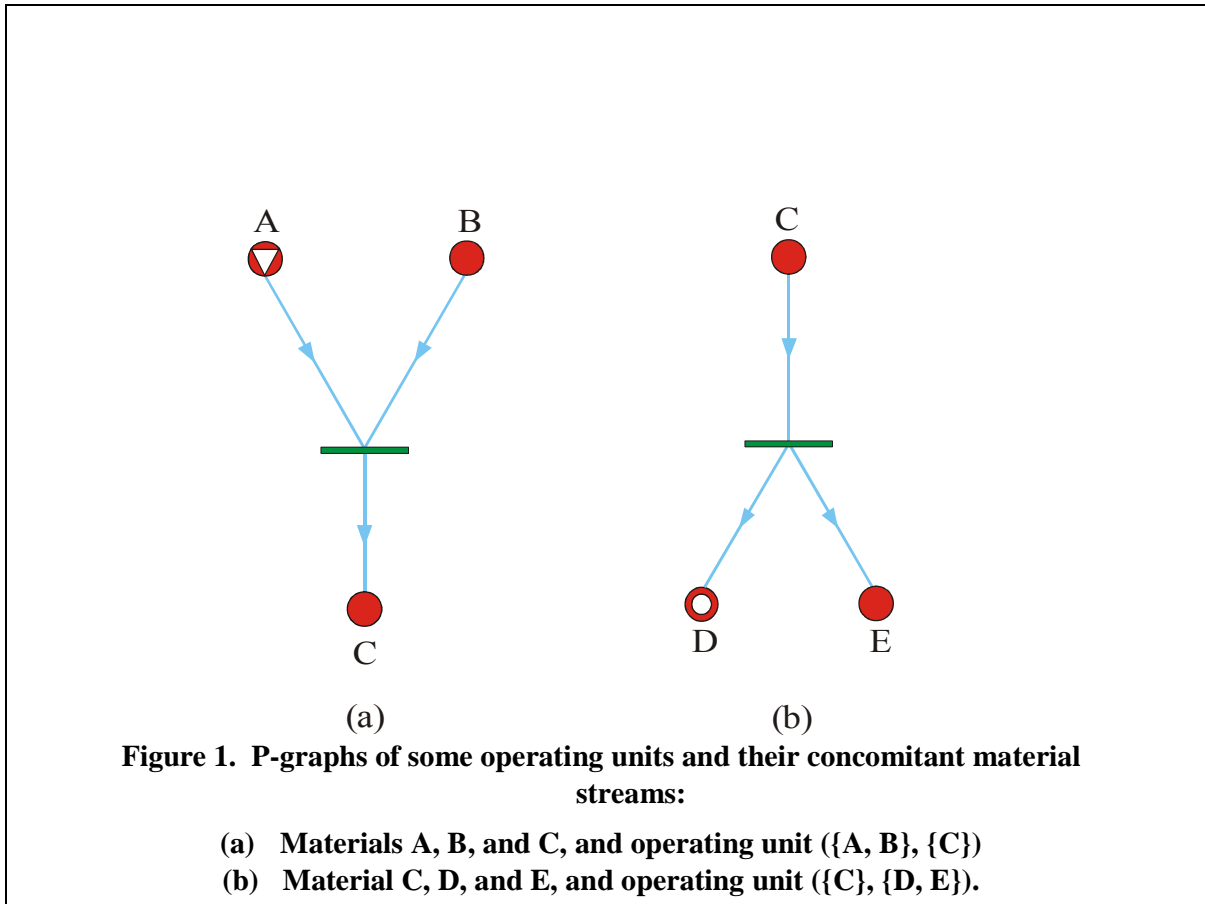
The first milestone of our journey toward the formalization and popularization of P-graphs for process-network synthesis was reached sometime in the early '90's when several memorable breakthroughs occurred. These breakthroughs included the presentations of a series of papers at major national and international technical

conferences; the review of the first major manuscript for journal publication by the late Professor Rippin and his daughter, a mathematician; and the publication of the first two refereed journal articles, “Combinatorial Algorithms for Process Synthesis, *Computers Chem. Eng.*, **16**, S313-320 (1992),” and “Theoretic Approach to Process Synthesis: Axioms and Theorems, *Chem. Eng. Sci.*, **47**, 1973-1988 (1992),” both by Friedler, Tarjan, Huang, and Fan (see Attachment 1). At that time, Prof. Rippin was one of the two chief co-editors of *Computers in Chemical Engineering*. His most favorable and encouraging comments on our manuscript were indeed gratifying; naturally, receiving such comments made us euphoric.

At this juncture, a short discussion on graphs in general and P-graphs in particular is in order. Graphs constitute a natural mathematical language or logical tool for describing and representing networks. Examples of such networks are gas or oil pipelines, waterways or irrigation channels, process flowsheets, highways, railroads, telephone lines, family trees, social relationships, and organizational structures. Some of these networks are physically visible, and some are not. The conventional graphs are represented by nodes, o , and arcs, $-$ or \rightarrow . Monopartite and bipartite graphs are typical conventional graphs; the former contains one kind of nodes, and the latter, two kinds of nodes.

Then, what are P-graphs? They are unique bipartite graphs depicted in Figure 1. Obviously, the question arises as to the rationale for proposing P-graphs or to their need. At the very early stage of development, it was demonstrated unequivocally that the conventional graphs, either monopartite or bipartite, are incapable of uniquely representing process networks: they are not sufficiently rich syntactically and semantically. Hence, a special class of bipartite graphs is sorely needed for this purpose. In this regard, it is worth noting that in recent years we have witnessed a proliferation of special classes of graphs in various fields, including call graphs, social graphs, and highway graphs. According to Hayes (*Graph Theory in Practice: Part I, American Scientist*, January-February, 2000), “... The next step is to develop a mathematical model of the structure, which typically takes the form of an algorithm for generating graphs with the same statistical properties. Such models of very large graphs will be the subject of ...” The P-graphs constitute one such special class of graphs.

At the outset of developing P-graphs, we identified the basic trait of process networks: they are all involved in the changes of compositions and locations of materials, or collectively, material transformation. Then, what are the same statistical properties of these graphs that need to be identified to generate a mathematical model of the structure in the form of an algorithm for constructing graphs of the process networks? Fortunately,



the mass-conservation law and stoichiometric principle are two “perfect” statistical properties shared by all the process networks or their graph representations; such “perfect” statistics naturally leads to rigorous axioms. It is not difficult to imagine that an algorithm or algorithms on the basis of such rigorous axioms would be robust. Specifically, we have established five axioms, which in turn, have given rise to the three algorithms, algorithm MSG for maximal-structure generation; algorithm SSG for solution structure generation; and algorithm ABB for accelerated branch-and-bound for generating the optimal and near optimal solutions. The maximal structure is the rigorously defined super-structure without redundancy.

The profound robustness and efficacy of the three algorithms, MSG, SSG, and ABB, have their roots in the rigorously, or exactly, stated set of five axioms. Nevertheless, their extraordinarily computational efficiency is mainly, but not exclusively, attributable to the drastic reduction in the search space resulting from the construction of the maximal structure in polynomial steps with algorithm MSG. This is illustrated in Figure 2, based on the optimal synthesis of an industrial-scale process containing 35 operating units, which are functional units, performing material transformation. This example appears in the first several publications of ours. Note that $(2^n - 1)$ networks are possibly generated from n operating units, the majority of which is, almost always, combinatorially infeasible.

Second milestone: mid 1990's

We reached the second milestone of our journey around mid '90's when our P-graph-based approach to process-network synthesis received favorable comments by Professor Sargent in a report of the Center for Process Systems Engineering of Imperial College. The report was kindly sent to me by Prof. Sargent. Our approach was also endorsed by Prof. Sargent in the Rippin memorial issue of Computers Chem. Eng. (Vol.

22, No. 1-2, 1998). Quoting him directly, “The task performance models also set conditions (both qualitative and quantitative)..., as considered by Friedler et al. (1992)... Armed with these rules and conditions we can devise an algorithm which systematically generates all feasible state-task-networks.” From reading both the report and paper, I

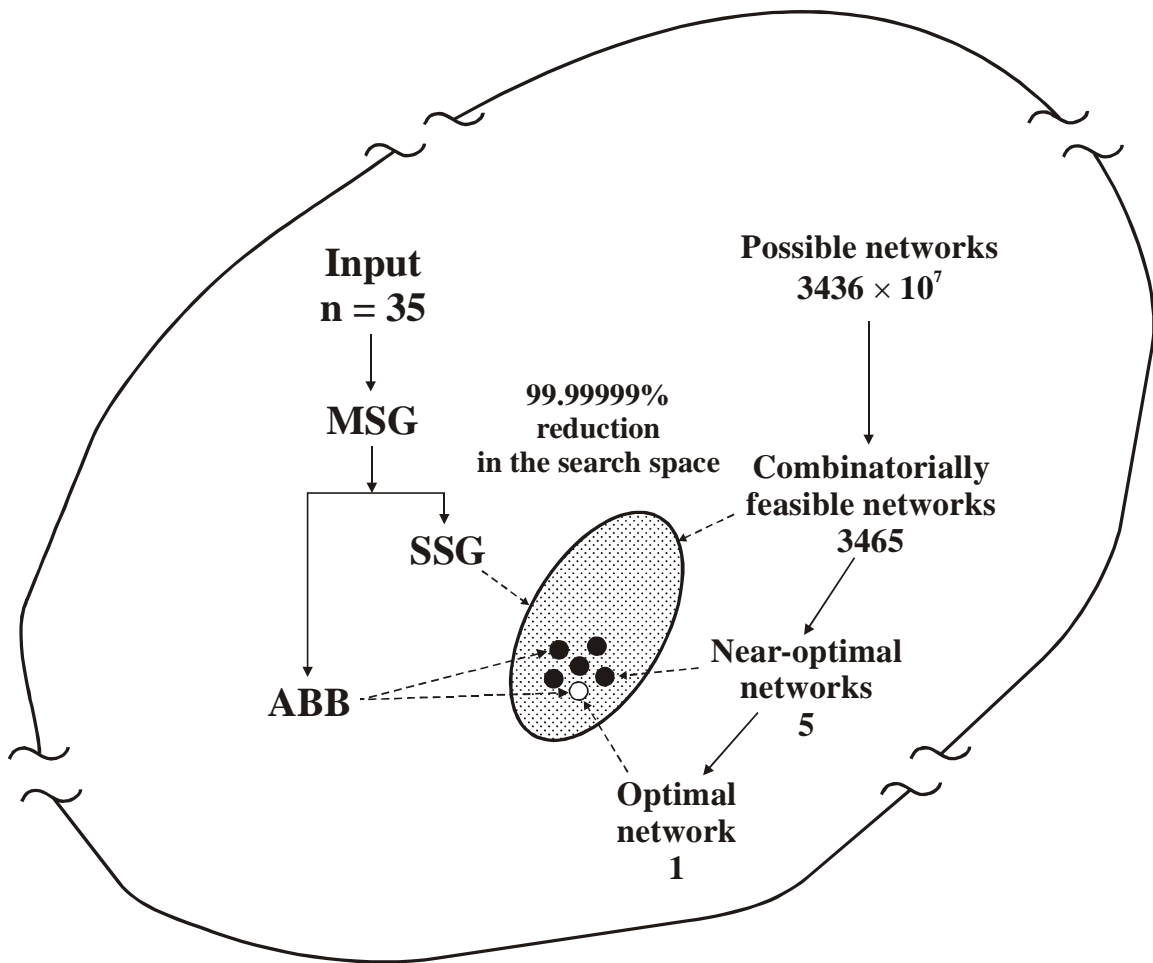


Figure 2. Reduction in the search space of $(2^n - 1)$.

sensed that his favorable inclination to our approach was largely influenced by Prof. Rippin whose name is prominently mentioned in connection with the first milestone.

Third milestone: mid 1990's ~ late '90's

The third milestone of our journey was reached sometime between mid and late '90's. Around that time, we came to the realization that P-graphs can be potentially adapted for or extended to the synthesis of mesoscopic processes and the identification of molecular networks by mimicking their syntheses in nature. The former led to the publication of a series of papers on azeotropic distillation, the first of which is the journal article entitled "Identifying Operating Units for the Design and Synthesis of Azeotropic-Distillation Systems," by Feng, Fan, Friedler, and Seib, appearing in *Industrial Engineering Chemistry Research* in year 2000 (see Attachment 1). The latter led to the preparation of a series of papers on the identification of catalytic-reaction and metabolic pathways, of which the combinatorial complexity of their syntheses is $(3^n - 1)$ instead of $(2^n - 1)$. The first of the papers in this series entitled, "Combinatorial Framework for the Systematic Generation of Reaction Pathways," was presented at the AIChE Annual Meeting held in Dallas, Texas in 1999, following which was the publication of three journal articles (see Attachment 1).

Fourth milestone: late 1990's ~ early 2000

The fourth milestone of our journey was reached essentially around year 2000. The most noteworthy event during this period was the endorsement of our P-graph-based approach by Dr. George Keller in his Institute Lecture at the 1999 AIChE Annual Meeting. Subsequently, the major portion of his Lecture was published in CEP (Volume 96, No. 1, 2000). Quoting directly from the published article, "...the P-graph may be the fastest computationally, as well as the method most likely to find a truly optimal solution."

Since the late '90's, the pace of adoption of P-graphs by other researchers has accelerated, as evident from their citations in presentations and publications including books and journal articles. Surprisingly but understandably, we were hardly asked to review any of these books and articles prior to their publication. As a matter of fact, we took it as another sign of the acceptance of P-graphs.

Fifth milestone: early 2000

The fifth milestone of our journey was reached in early 2000 when Prof. Klaus Timmerhaus identified our P-graph-based approach as the method of choice for algorithmic flowsheeting. This has resulted in the inclusion of a 30-page section entitled, ALGORITHMIC FLOWSHEET GENERATION, in the 5-th edition of "Plant Design and Economics for Chemical Engineers," by Peters, Timmerhaus, and West. This premier textbook on plant design is published by McGraw-Hill.

Around this time, we published a landmark paper in Volume 24 of Computers in Chemical Engineering. This paper, co-authored by Brendel, Friedler and Fan, provides additional proof of the rigorousness and superiority of our P-graph-based algorithmic method for process synthesis over other algorithmic methods for process synthesis that initiate the procedure with the construction of the super-structure as is the case with our method (see Attachment 1).

Current status and future prospect of P-graphs

Now the point is reached for me to reveal what is transpiring currently and what will transpire in the near future regarding our work on P-graphs. These include: optimal syntheses of various downstream processing systems for biochemical production of chemicals from grains and other natural resources; complex azeotropic-distillation system synthesis; design of alternative synthetic routes; profit-potential estimation; separation-network synthesis incorporating separators effected by different separation methods; heat-integrated separation-network synthesis; identification of catalytic-reaction mechanisms; and metabolic-pathway identification and metabolic-flux analysis, which would greatly facilitate drug discovery. The most important of all, however, are further

generalization of the algorithms and substantial updating of concomitant software. Also envisioned is computational acceleration through parallel and grid computing. Clearly, these efforts entail additional funding from private and/or public sources.

Given in Attachment 3 to illustrate our current and future works related to P-graphs are the abbreviated or condensed view graphs of the 3 papers that were presented at this 2003 Annual Meeting of AIChE. These illustrations are arranged in the order of presentation.

From Microscopic World to Mesoscopic World through Drunkards' Paths

“Drunkards’ paths” are often used in popular expositions of random walks that probably belong to the simplest class of stochastic processes. Collectively, stochastic processes constitute a rigorous branch of mathematics or mathematical statistics. It is concerned with random phenomena occurring over time or space according to a certain mathematical property defined by a distribution of the random variable. What distinguishes any stochastic model from the corresponding deterministic or continuum model is its capability to represent rigorously not only the gross, or mean, behavior of the phenomenon or process of interest but also its inherent fluctuations. This capability of revealing inherent or characteristic fluctuations is absent in the deterministic or continuum model.

This section is structured similar to the preceding section. The headings of subsections of the preceding section are mainly in terms of milestones. In contrast, those in this section are mainly in terms of major personalities whose contributions inspired us or guided our work during my journey.

Awakening at dawn: 1950's

Since my undergraduate days, I have been a student of the analysis, design, fabrication and operation of continuous flow chemical reactors in either the tubular or

stirred-vessel configuration. It was not difficult for me, like everyone else, to experience or observe that the flow of reacting molecules through any of such reactors is not ideal: These molecules either disperse in the direction of flow or imperfectly mixed. This naturally led to my intense interest in the notion of residence-time distribution.

In an attempt to understand the residence-time distribution, I started to immerse myself in the subject whenever time permitted. My effort became fruitful towards the end of the '50's. A large part of the contributions of my collaborators and me is contained in the well-received monograph published by Marcel-Dekker in 1966, "Models for Flow Systems and Chemical Reactors," written with my closest friend and classmate, the late C. Y. Wen.

The molecules flowing through a reactor are microscopic and discrete entities; they move or behave independently and randomly. It is not surprising, therefore, that qualitative discourses of residence-time distributions in the publications, including my own, are full of statistical, or stochastic, jargon; yet, the quantitative treatments of residence-time distributions are entirely deterministic involving much "hand-waving" arguments. For me, this was and has been for a long time intellectually untenable. I knew deep down that the residence-time distribution could be rigorously treated based on statistics or stochastic processes. This was very clear to me even in the late '50's: I became aware that the frequently-used alternative name to the residence-time distribution is the age distribution belonging to the parlance of actuaries, practitioners of biostatics or stochastic processes.

During the second half of the '50's, my participation in the various process operations related to solid particles, liquid droplets or gas bubbles noticeably quickened. To me, those mesoscopic discrete entities almost always randomly "dance" in process vessels. The majority, if not all, of the mathematical models of processes involving these entities that were available then, however, were deterministic in nature. I attempted to remedy this situation by emulating the methods of statistical mechanics without much success. Eventually, I reached the conclusion that the methods of stochastic processes would be most appropriate because of the time-dependency of these processes.

Encounter with Markov: mid 1960's ~ early 1970's

During this period, some of my graduate students, research associates and I self-studied the classical textbooks on stochastic processes; including those by Feller, Cox and Miller, and Chiang. Moreover, we frequently consulted and extensively collaborated with faculty members in the statistics department. The core of our study was on the Markov processes comprising Markov chain (time discrete), Markov process (time continuous), and diffusion process (state and time continuous), the last of which leads to the well-known Fokker-Planck equations. All these Markov processes invoke the renowned Markov assumption. This assumption contends that the probability of occurrence of the event of interest at present depends only on its occurrence in the preceding time period. In other words,

$$\begin{aligned} & \Pr[X(t_n) \leq x_n \mid X(t_1) = x_1, X(t_2) = x_2, \dots, X(t_{n-1}) = x_{n-1}] \\ &= \Pr[X(t_n) \leq x_n \mid X(t_{n-1}) = x_{n-1}] \end{aligned}$$

We were fairly successful in stochastically modeling a number of processes or phenomena involving discrete microscopic or mesoscopic entities by invoking the Markov assumption. On the other hand, we were somewhat disappointed that all the methods or techniques of stochastic processes we learned from the aforementioned classical textbooks were applicable only to linear processes. Yet, many of the significant problems requiring stochastic treatment but remaining unsolved were nonlinear in nature.

Encounter with Prigogine: early 1980's

I have been attracted to the papers by Prigogine, who recently passed away, since the mid '60's because of my interest in Non-equilibrium Thermodynamics. In fact, my student and I contributed a short article on the subject to IEC Research although it was concerned only with the linear version of Non-equilibrium Thermodynamics by Onsager. I found almost all of the monographs and papers written by Prigogine to be extremely difficult to tackle. It was indeed discouraging; I attributed it to my mental incapability.

It was a relief when a short story about Prigogine in a reputable publication caught my eye one day while I was "wasting" my time browsing "randomly" through books and journals in the university library. The gist of one passage said that book

companies requested Prigogine to write monographs only with his students or research associates who suffered from his notoriously convoluted writings and thus became proficient in interpreting or delineating them. One such monograph is, “Self-organization in Non-Equilibrium Systems,” by Nicolis and Prigogine, published in 1977. Reading through the monograph, especially one passage in it, thrilled me to no end. It essentially states that Prigogine’s epoch-making analysis and modeling of self-organizing phenomena far removed from equilibria is attributable largely to the fact that he used non-linear, birth-death processes, a class of time-continuous Markov processes. It results in the non-linear master equations that are solvable. Later, the contribution of Prigogine earned him a Nobel Prize in Chemistry. Nevertheless, I could never “master” the solution procedure adopted by him. Despite this dilemma, I was encouraged: “Anything good enough for Prigogine is good enough for me.”

Encounter with van Kampen: mid 1980’s

My dilemma vanished when we discovered monographs and journal articles by one of the leading theoretical physicists, van Kampen. The most important among his contributions is a monograph, “Stochastic Processes in Physics and Chemistry,” published in 1982. I consider this monograph, without question, as the bible of non-linear master equations, probability-balance equations or gain-loss equations derived from the birth-death processes. The complexity arising from the non-linearity is circumvented by a rational approximate method, i.e., system-size expansion.

Mastering the master-equation approach immeasurably enhanced our groups’ productivity. It has led to the publication of series of papers, each ranging from two to more than a dozen, mostly dealing with mesoscopic entities or systems on various subjects such as chemically reacting systems, solids mixing, grinding, fluidization, crystallization, filtration, biochemical processes, interphase mass transport, and residence-time distribution. These papers are listed in Attachment 2. Naturally, we called on the system-size expansion whenever necessary to deal with non-linearity. Among the publications, I would like to single out the following two articles pertaining to the last-mentioned subject, residence-time distribution.

“The Surface-Renewal Theory of Interphase Transport: A Stochastic Treatment,”
Chem. Eng. Sci., **48**, 3971-3982 (1993), by Fan, Shen, and Chou.

“Stochastic Modeling of Transient Residence-Time Distributions during Start-
Up,” Chem. Eng. Sci., **50**, 211-221 (1995), also by Fan, Shen, and Chou.

The first of the two is concerned with the rigorous, non-handwaving treatment of the celebrated interphase transport theory of Danckwerts. The residence-time distribution at the interface comes into play in this theory.

The readers might recall that I alluded to my frustration with the deterministic or continuum treatment of residence-time distribution, which ultimately motivated me to start my journey through “Drunkards’ Paths” charted by the rigorous mathematics of Stochastic Processes. I am firmly committed to continue my work in stochastic analysis and modeling. To follow a drunkard’s path would be much more enjoyable than to walk through a rigid path.

Current status and future prospect of stochastic analysis and modeling

To illustrate what our group is currently doing, one of our recent papers (see Attachment 2) and a paper presented at the 2003 Annual Meeting of AIChE are listed below.

“Stochastic Modeling of Thermal Death Kinetics of a Cell Population: Revisited.”

“Stochastic Modeling and Simulation of the Formation of Carbon Molecular Sieves by Carbon Deposition.”

The abbreviated or condensed viewgraphs of these papers are given in Attachment 4. The papers are based on linear master equations; they are being actively extended to various non-linear cases. Moreover, we are preparing monographs on stochastic analysis

and modeling of particulate systems, chemically reacting systems, and biochemical systems.

The future of our work will entail the extension to such subjects as nanoparticle formation, HIV infection, tumor growth, air pollution, membrane separation, and many others. The entire community of Chemical Engineering is getting into nanotechnology. Why not us also? After all, nanoparticles are discrete and mesoscopic and are bound to behave stochastically in various environments or circumstances.

Concluding Remarks

At the outset, I unabashedly declared my age, namely, 74. Hence, the most appropriate concluding remark would be, “Old professors never die; they just (asymptotically) fade away.” Obviously, I borrowed heavily from General Douglas MacArthur.

Acknowledgements

I would like to express my profound appreciation to all my current and former students, assistants, associates, collaborators, and teachers; all my current and former colleagues and staff in the department; all organizations and agencies in and out of the University that supported my research; all attendants who were bewildered by my “entangled” and “random” talks; and last, but not least, all my family members, especially my wife, Eva, who has accompanied me for 45 years in the journey through the “Mazes of Process Graphs” and “Drunkards’ Paths.”