Modeling Production-Chain Information

Jeffrey Dean Kelly
Honeywell Process Solutions

Data alone is not information. Use this model to transform data into vital information for decision making.

THE CHALLENGE OF UNDERSTANDING AND managing a diverse information system such as those found in the process industries, especially those of the production-chains, can be both difficult and rewarding. The focus of this article is to highlight the behavioral and structural nature of a production-chain's information system from the perspective of how it is planned, scheduled, simulated, monitored and reconciled; the other aspects of how it is controlled and executed in real or near real-time are briefly addressed. A production-chain is the make-side of the value-chain, whereas the supply-chain is the buy-side and the demand-chain is the sell-side. It is in the production-chain where raw materials are transformed into finished goods via many specialized and complex processes and where many economists believe true wealth is created. We begin by unfolding a clear distinction between the process industries and the discrete-parts manufacturing industries in order to convey in some detail why advances in discrete industry information systems are generally not comprehensive nor specific enough to deal with the distinctive characteristics of the process industry (*1*).

In the process industries, value-add is achieved through the operations of mixing, reacting and separating, whereas in the discrete industries value is added through activities or steps such as machining, trimming and assembling. Typical process industries are pharmaceuticals, food and beverage, oil refining, petrochemicals, bulk and specialty chemicals and polymers. Typical discrete industries are automotive, apparel, electronics, furniture and appliances. An important differentiator between the two industries is the breakdown of production into three key dimensions we classify as quantity, logic and quality. The discrete industries, although having the dimension of quality, deal mainly with the aspects of quantity and logic (known as logistics); where quantity deals with the amount or number of items and logic represents the operating rules and manufacturing procedural details. Quality, in our context, refers to the compositions, properties and conditions of the material being processed, although concomitantly, quality is also defined as the amount of variability in the performance of the company and in the traits, attributes or characteristics of the company's products (*2*). When superior raw materials are purchased and optimal manufacturing techniques are used, then quality discrete products with strict tolerances can be produced. However, in the process industries the raw materials are typically comprised of non-pure and relatively unknown species, such as the world crude-oils or intermediate stocks from different processes. Therefore, it should be apparent that an information system designed for the discrete industries may not be suitable to support the unique aspects and nuances of the process industries. This is a relatively well-known conjecture, given that computer integrated manufacturing (CIM) and manufacturing execution systems (MES), prevalent in the discrete industries, have not been well adapted to the process industries. It is also interesting to note that popular enterprise resource planning systems (ERP) and supply-chain management (SCM) software usually struggle to adapt to the business requirements of the process industries unless they are specifically designed and developed for that purpose.

A more advanced description of the differences between process and discrete industries can be enumerated by how they are planned and scheduled. Moreover, there are also strong differences when viewed from how these two industries are controlled and executed at the plant and shop floor. Discrete industries are characterized by bill-of-materials (BoM) planning and open-shop scheduling. BoM planning is also known as material requirements planning (MRP) and manufacturing resource planning (MRP II), which use a BoM explosion to determine the raw materials and work-in-progress quantities based on the demand orders for finished goods. BoMs are suitable to propagate the raw material requirements essentially because most discrete industries' flow-paths are convergent or Leontief. Open-shop scheduling refers to the well-known problems of project-shop, machine-shop, job-shop and flow-shop scheduling (*3*) and deal with the assignment of usually renewable resources (*e.g.,* machines, people) and the sequencing and timing of activities.

In contrast, the process industries are planned using either linear programs (LP) or nonlinear programs (NLP) to determine the raw material requirements, or lot-sizing, and in some cases, the operating condition requirements based on the quantity and quality constraints of the production-chain. These planning systems also aid in the finished product portfolio selection based on forecasted demand orders. The primary reason LPs are necessary is to handle the divergent flow-paths of the process industry production-chains. Divergent means there are co-products and byproducts that do not drive the top-line economics of the processing plant, even though they yield bottom-line credits. The main reason NLPs are required is to handle the ubiquitous bilinear quantity times quality relationships giving rise to the well-known pooling problem (*4, 5*).

The lot-sizing problem is also found in process industry scheduling, and when this occurs, the scheduling problems are classified as closed-shops (*6*) which also includes the set of resource-constrained project scheduling problems (RCPSP) and are inherently time-indexed (i.e., must have time-periods or time-slices). Lot-sizing is a key aspect of process industry scheduling given that intermediate storage, hold-up or inventory is finite and definitely not unlimited due to the usually substantial liquid and gas material quantities being processed. And, given the spatially and temporally aggregated plans that are either inaccurate or at too high a level to be of any use when scheduling (*i.e.*, longer-term material allocation estimates must be re-calculated due to the actual or impending equipment selections). Though inventory restrictions are real in the discrete industries, in the process industries overflow containment has much less flexibility given that the cost of transporting the excess stock to a nearby warehouse is rarely a viable option. Therefore lot, charge, batch and movement-sizing is a powerful degree-of-freedom or lever when scheduling around capacity bottlenecks. Open-shops are a subset of closed-shops because in open-shops there is only a trivial lot-sizing problem with respect to whether a job or project (*i.e.,* like a raw material or stock) should be scheduled on a particular machine (*7*).

The purpose of the information model

The main theme of all quality programs is to enable and steward the notion of plan, operate, monitor, analyze and improve (POMAI), design, measure, analyze, improve and control (DMAIC), or plan, execute, capture and analyze (PECA). This can be distilled down into three dimensions — plan, perform and perfect. Planning may include research, design, strategy and scheduling throughout the entire decision-making hierarchy of the overall business. Performing includes all of aspects of the production's execution, tactics, control, regulation and operation. Perfecting involves reflecting, analyzing, monitoring and improving the overall system. This then forms a continuous improvement feedback loop called the production-chain information model (PCIM) loop, which has strong ties to the basis of process control theory.

These three dimensions can be divided into two layers — one called the "production and analysis" layer, and the other the "process and execution" layer. The "production and analysis" layer includes two key functions, planning (*e.g.,* advanced planning and scheduling, material requirements planning and master production scheduling) and perfecting (*e.g.,* statistical process control, steady-state and dynamic process simulators, fault detection and identification (*8*), and statistical data reconciliation). The "process and execution" layer has one key function, performing, which includes real-time optimization (RTO), advanced process control (APS), proportional-integral-derivative (PID), relay ladder logic (RLL) and laboratory data system (LDS). LDS is included in this layer given that quality data, such as the melting index of a polymer or the concentration of a stream must be performed by either a field or control laboratory, whereby the data is then used in a higher-level (or supervisory) manual improvement cycle to control quality.

The purpose of the production-chain information system is to accommodate not just time-series and transactional data, but to provide an unprecedented level of integrated and timely production context. The key to achieving this is through a PCIM. We begin by theorizing that the solution to any problem requires a technology or innovation. Formalizing the definition of a problem can be articulated as the combination of a model plus data. For example, in the presentation of a scheduling solution to a production problem, with all of its quantity, logic and quality constraints and bounds looking out many time-periods into the future virtually hidden from the solution, the Gantt or time-chart shows the results or information of the scheduling optimizer in chronological time order. The solution does not, however, detail the underlying system details used to generate the schedule. Thus, to have information there first must be a problem to give the information meaning and purpose and then a methodology, skill, expertise, tool and/or knowledge to transform the problem into a solution and hence information. Stated mathematically we have that a *model + data + intelligence = information.*

It is the model and the intelligence that when combined, will create information, although data is still a vital term in the equation. Moreover, accurate information requires an accurate model, data and intelligence. Accurate intelligence means choosing the most appropriate technique or method to cast the model and data into useful information such as choosing the right SPC run-chart to detect the anomalies with the most negative impact on production. It should also be pointed out that the model and data provide the structural aspects of the problem, whereas intelligence provides the behavioral context. This simple relationship has a chemical analogy relating to the process of combustion and that is that *fuel + air + ignition = fire*. It can be argued that intelligence is the "spark" necessary to turn the model and data into information. Yet, there is another related piece of the analogy that fits well with the description of information and that is the fuel-to-air ratio. Too lean or too rich a mixture will not support combustion even with ample ignition. This too is the case with the mixture of model and data. Too much data and not enough model will only yield itself and will not be altered into something more useful.

In the above we talked about the notion of planning, performing and perfecting in the production-chain in context with the quality revolution. This then related to the idea that information requires a model, data and some technology or science to convert the model and data into information. All of this is good, however, there must be a benefit to the effort. The benefit side of the equation relates to the three known innovations in manufacturing laid out by (*9*). The first is interchangeable or replaceable parts pioneered during the U.S. Civil War for the fast and cheap production of rifles and munitions. The second is the growing and/or acquiring of a portfolio of products (or companies) to hedge and reduce the impact and risks of downturns in the economy or limited availability of raw materials by increasing ones diversification. And the third, which is the innovation we are currently implementing, is the notion of production for final demand or demand-driven production (DDP). Any production-chain information system must be tuned to the underlying business direction that through the effects of internet collaboration, high-speed communications for data and information flow and integrated value-chains, producing only those products that the market demands will reduce raw material and finished goods inventories, transportation costs, pricing fluctuations, advertising and promotions and increase customer service levels just to name a few (*10*).

Both a model and intelligence need to be injected in order to have information. Unfortunately and contrary to popular belief, there can never only be just one model to support the production-chain or business as there are typically many business functions that must be performed by the organization and so too are there many models that instill the essence of the business's primary objectives. Notwithstanding, there is a legitimate argument to have as few sources of data as possible, but mandating only one model is not sensible or practical. As for the intelligence component, because there are many people of varying talent and skill that are employed in an organization, so too are there many types of techniques and methods from which to transform model and data into information. For a production-chain, which has as its principal mandate to value-add raw materials into finished goods, the ability to plan and schedule the production, years, months, weeks and days into the future is a profound and powerful expertise.

Planning and scheduling

Planning and scheduling involves time and must be performed on a regular basis in order to mitigate against the effects of uncertainty, complexity and organization structure — this is known as hierarchical planning and scheduling (*11*). The key differentiator between planning and scheduling is what is known as time-buckets or time-periods. Big time-buckets allow one or more activities or operations to be performed on a process-unit within the same time-period and is known as planning. Small time-buckets permit at most one operation or task to be performed by a particular piece of equipment within one time-period and is known as scheduling. Planning is known as providing a course of action for strategic, tactical and operational decisions and in the process industries involves many non-linearities due to the quantity times quality relations as mentioned. Scheduling must provide the well-known functions of assigning resources to tasks, sequencing the tasks and then timing the tasks, but in the process industries, scheduling must also be able to size the flow or quantity amounts in order to alleviate the many finite capacity constraints that exist in the production-chain while assessing the impact of these decisions on the overall quality of the stocks being produced.

At the heart of all planning and scheduling systems are models. There are often one or more planning models and usually one or more scheduling models required to manage the complexity and scope of the problems. Each model may contain the same data and it is important to recognize that there really cannot be just one model to perform the necessary look-ahead functions. Planning is about aggregation of space and time and scheduling is about disaggregating the planning results and generating implementable schedules working the details ignored in the plan. A production-chain information system must be able to support many types of models for the various functions else useful information will not be the outcome. Some of the techniques required to effectively model complex production-chains can be found in Ref. 12. In that article, production modeling objects are categorized into three major items — units, operations and stocks. Units are the physical equipment necessary to support the technology to transform raw materials into finished goods and are renewable resources. Operations are the functional tasks or steps the units are engaged in and the nonrenewable stock resources represent the materials being value-added which may also include secondary supporting resources such as labor, catalyst, tools and utilities. Units and stocks are considered as the structural components of the production model and the operations are considered as the behavioral components. The combination of both the structural and behavioral elements of production can model any and all complex manufacturing processes including the discrete industries' production scenario.

Simulating, monitoring and reconciling

Simulating the manufacturing, monitoring its operation and reconciling its production are more passive types of functions than planning and scheduling, but can provide a powerful retrospective of the production from which to improve. There are typically two types of simulators used within the process industries, process and production simulators. Process simulation is a primary tool used in the design and to a lesser extent the monitoring of processes especially the well-known steady-state process simulators modeling the rigorous details of the physical properties, thermodynamics, reaction kinetics and fluid mechanics of the process. These simulators have now evolved into modeling the dynamics, or more specifically the rate of change of matter, energy and momentum over time and are used to train operators and to design and tune advanced control strategies. Process simulators have detailed models and require minimal data known as specifications to solve and generate for example liquid and vapor temperature profiles and separation composition gradients throughout the spatial context of the processing equipment. These models are too detailed to be used in planning and scheduling, although some attempts to use the first-order derivative information in the search for better plans has been employed with usually intangible and/or unverifiable benefits. However, the open-equation form of the process simulators are used in real-time optimizers to push the process-unit closer to its cheaper and potentially safer constraints. Current production simulators, on the other hand, are more focused on the quantity and quality aspects of the production and relegate or ignore the fine points of the process simulator models. Production simulators are not appropriate to design process-units nor are they suitable to monitor the vapor and liquid traffic inside a distillation or fractionation tower to detect jet flooding. Production simulators are very useful to balance the supply and demand quantity orders of the plant with respect to the storage and process-unit capacities and key quality stipulations. These simulators are also used to predict or trace the composition and properties of the stock flowing in and out of the different production stages for both past and future production as for instance crude-oil composition tracing. Accurate information regarding the quality of a tank's material can be used to make spot, retail or discretionary buys or sells of work-in-progress stocks or feed and productstocks. Nevertheless, production simulators rarely organize or help the end-user manage the logic or operating rules of the production except through embedded ad hoc and unsupportable heuristics.

We touched briefly on the monitoring function that simulators can support, but there is an entire discipline in statistics devoted to the monitoring of manufacturing processes known as statistical process control or statistical process monitoring (SPM). These techniques allow for the analysis of both univariate and multivariate data sets from characterizing the data's distribution to correlating independent variables with dependent variables found in latent variable regression techniques such as principle component analysis (PCA) and partial least squares (PLS) (*13*). The key difference between process and production simulators and SPC methods are that the SPC methods rely on empirical and data-driven regression models whereas simulators use the fundamental laws of conservation to predict the dependent variables. Both types of models are useful for detecting and identifying the faults, defects, outliers, anomalies or gross errors in the system being monitored where a combination of both methods can be very powerful. Another useful technique that combines both statistics and conservation laws is statistical data reconciliation. SDR is gaining momentum in the process industries especially in the production-chains to reconcile yesterday's production with its receipts and shipments of stock. At the core of SDR are the material, energy and momentum balances including within process stream balances for qualities such as composition and concentration. Reconciliation is appropriate to detect field and laboratory measurement gross errors by cross-checking the measurement data with respect to other measurements including transactional data for stock movements using the balance equations as extra measurements so to speak to enhance the redundancy of the system (*14*). This redundancy allows for the spatial location of a measurement failure to be located if statistically significant (*i.e.,* above a 95% confidence interval). **CEP**

Literature Cited

1. **Crama, Y., *et al*.,** "A Discussion of Production Planning Approaches in the Process Industries," CORE Discussion Paper, Catholic University of Louvain, CORE and IAG, Louvain-la-Neuve, Belgium (2001).
2. **Bodington, C. E., Ed.,** *Planning, Scheduling and Control Integration in the Process Industries*, McGraw-Hill Inc., San Francisco (1995).
3. **Pinedo, M.,** *Scheduling: Theory, Algorithms and Systems,* Prentice Hall, NJ (1995).
4. **Kelly, J. D., and J. L. Mann,** "Crude-Oil Blend Scheduling Optimization: An Application with Multi-million Dollar Benefits — Parts I and II," *Hydro. Proc.* (June/July 2003a).
5. **Kelly, J. D.,** "Formulating Production Planning Models," *Chem. Eng. Prog.* (Jan. 2004).
6. **Graves, S. C.,** "A Review of Production Scheduling," *Operations Research*, **29** (4), pp. 646–675 (1981).
7. **Kelly, J. D., and D. J. Adams,** "Plant-wide Scheduling in Petroleum Refineries: Optimizing Quantity, Logic and Quality Decisions," Society of Chemical Industries Meeting, London, U.K. (Nov. 2003c).
8. **Himmelblau, D. M.,** *Fault Detection and Diagnosis in Chemical and Petrochemical Processes,* Elsevier Scientific Publishing Co., Amsterdam (1978).
9. **Norman, A., *et. al*.,** "The Need for a Paradigm for Innovation," http://www.eco.utexas.edu/homepages/faculty/Norman/long/InnParadigm.html, Dept. of Economics, Univ. of Texas at Austin (Aug. 1999).
10. **Taylor, D. A.,** *Supply Chains: A Manager's Guide,* Addison-Wesley, Boston (2004).
11. **Bitran, G. R., and A. C. Hax,** "On the Design of Hierarchical Production Planning," *Decision Science*, **8** (28) (1977).
12. **Kelly, J. D.**, "Production Modeling for Multimodal Operations," *Chem. Eng. Prog.* (Feb. 2004).
13. **Nomikos, P. and J. F. MacGregor,** "Multivariate SPC Charts for Monitoring Batch Processes," *Technometrics*, **37**, pp. 41–59 (1995).
14. **Kelly, J. D.,** "The Necessity of Data Reconciliation: Some Practical Issues," NPRA Computer Conference, Chicago, IL, (Nov. 2000).

JEFFREY DEAN KELLY is a solutions architect for advanced planning and scheduling at Honeywell Process Solutions (300 Yorkland Blvd., Toronto, Ontario, Canada, M2J 1S1; Phone: (416) 502-5148; Fax: (416) 502-5130; E-mail: jeff.kelly@honeywell.com). Kelly is a chemical engineer and has a Master's degree in Advanced Process Control from McMaster University. He has worked as an advanced control engineer at both Shell Canada and Imperial Oil, including implementing real-time optimization programs and tactical planning and scheduling solutions in their refineries. He has installed plant-wide data reconciliation packages in several oil-refineries around the world and has written many publications on the subject.