

Vision 2020: Computational Needs of the Chemical Industry

T.F. Edgar, University of Texas

D.A. Dixon, Pacific Northwest National Laboratory

G.V. Reklaitis, Purdue University

INTRODUCTION

There are a number of forces driving the U.S. chemical industry as it moves into the 21st Century including shareholder return, globalization, efficient use of capital, faster product development, minimizing environmental impact, improved return on investment, improved and more efficient use of research, and efficient use of people. As the chemical industry tries to achieve these goals, it is investigating the expanded use and application of new computational technologies employed in areas such as modeling, computational chemistry, design, control, instrumentation, and operations. The key technology driver over the past 20 years has been to the continuing advances in digital computing. The 100-fold increase in computer speed each decade has led to significant reductions in hardware cost for computers of all types and has increased the scope of applications in chemical engineering.

A forecast of future advances in process modeling, control, instrumentation, and optimization is a major part in the recently completed report, "Technology Vision 2020: Report of the U.S. Chemical Industry". This report was sponsored by five major societies and associations (AIChE, ACS, CCR, CMA, and SOCMA), and involved more than 200 business and technical leaders from industry, academia and government. It presents a roadmap for the next 20 years for the chemical and allied industries.

The collaboration among the five societies, as well as government agencies (DOE, NIST, NSF, and EPA), has spawned many additional workshops, generating more detailed R&D roadmaps on specific areas of chemical technology. Several workshops pertinent to this paper have been held during 1997 and 1998, covering the areas of instrumentation, control, operations,

and computational chemistry. Other Vision 2020 workshops have been held on subjects such as separations, catalysis, polymers, green chemistry and engineering, and computational fluid dynamics (See <http://www.chem.purdue.edu/v2020/>, the Vision 2020 web site for workshop reports).

This paper reviews the computational needs of the chemical industry as articulated in various Vision 2020 workshops. Subsequent sections of this paper deal with process engineering paradigm in 2020, computational chemistry and molecular modeling, process control and instrumentation, and process operations.

PROCESS ENGINEERING IN 2020

Increased computational speeds have spurred advances in a wide range of areas of transport phenomena, thermodynamics, reaction kinetics, and materials properties and behavior. Fundamental mathematical models are becoming available due to an improved understanding of microscopic and molecular behavior, which could ultimately lead to ab initio process design. This will enable design of a process to yield a product (e.g., a polymer) with a given set of target properties, predictable environmental impact, and minimum costs. Ideally one would want to be able to start with a set of material properties and then reverse-engineer the process chemistry and process design that gives those properties.

Historically the chemical industry has used the following sequential steps to achieve commercialization:

- (1) research and development
- (2) scaleup
- (3) design
- (4) optimization

Note that steps (1) and (2) generally involve several types of experimentation, such as laboratory discovery, followed by bench-scale experiments (often of a batch nature), and then operation of a continuous flow or batch pilot plant. It is at this level that models can be postulated and unknown parameters can be estimated in order to validate the models. Using these models a plant can be designed and then optimized. If the uncertainty in process design is high, pilot scale testing may involve several generations (sizes) of equipment. With the advent of molecular scale models for predicting component behavior, some laboratory testing can be obviated in lieu of simulation. This expands upon the traditional relationship of scientific theory and experiment to form a new development/design paradigm of process engineering (see Figure 1).

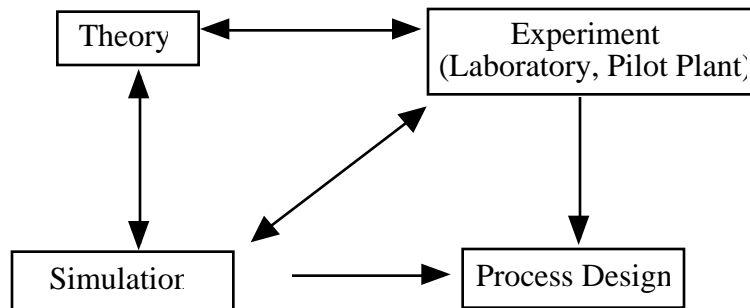


Figure 1. Process Engineering Paradigm for the 21st Century

The development of mathematical models that afford a seamless transition from microscopic to macroscopic levels (e.g., a commercial process) is a worthy goal, and much progress in this direction has occurred in the past ten years in areas such as computational fluid dynamics. However, due to computational limitations and to some extent academic specializations, process engineering research has devolved into four more or less distinct areas:

- (1) process design
- (2) structure property relationships
- (3) process control
- (4) process operations

In fact during the next two years one can attend separate research conferences in each of these areas, but only a few hardy souls will participate in cross-fertilizing the areas by attending multiple conferences. Consider the interaction of process design and control; process design decisions can be made that simultaneously optimize plant profitability and the controllability of the plant, rather than the traditional two-step approach of designing the most profitable plant and then considering how to control it in a subsequent design phase. In fact the different models, problem scope, and terminology used in each of these areas is an indicator that no lingua franca has emerged. Actually areas (1), (3), and (4) fall under a broad umbrella of systems technology, but until these three areas begin to use a common set of mathematical models, progress towards a more catholic view of process design will be impeded.

A molecular level understanding of chemical manufacturing processes would greatly enhance the ability of chemical engineers to optimize process design and operations as well as ensure adequate protection of the environment and safe operating conditions. Currently there is considerable uncertainty in thermodynamic and reaction models, so plants are normally over-designed (above required capacity) to allow for this uncertainty. Also plants are operated conservatively because of an inadequate understanding of dynamic process behavior and the dire consequences if an unsafe condition arises. Chemical reactors are at the heart of this issue, with uncertainties on kinetic mechanisms and rate constants and the effects of reactor geometry (such as catalyst beds) on heat and mass transfer. Clearly the availability of better microscopic mathematical models for macroscopic plant simulation will help the chemical industry operate more profitably and more reliably in the future.

Besides providing fundamental data for process simulations, computational chemistry plays an important role in the molecular design process beginning at the basic research level. By predicting accurate thermochemistry, one can quickly scope out the feasibility of reaction pathways as to whether a reaction is allowed or not. Computational chemistry can also reliably predict a wide range of spectroscopic properties to aid in the identification of chemical species, especially important reaction intermediates. Electronic structure calculations can also provide quantitative insights into bonding, orbital energies and form, facilitating the design of new molecules with the appropriate reactivity.

COMPUTATIONAL CHEMISTRY AND MOLECULAR MODELING

The computational chemistry subgroup of Vision 2020 under the sponsorship of the Council of Chemical Research has outlined a set of computational “Grand Challenges” or “technology bundles” that will have a dramatic impact on the practice of chemistry throughout the chemical enterprise, especially the chemical industry. The computational “Grand Challenges” are given in Table 1.

Table 1

**Computational “Grand Challenges” for Materials and Process Design
in the Chemical Enterprise**

- A. Reliable prediction of biological activity from chemical structure
 - B. Reliable prediction of environmental fate from chemical structure
 - C. Design of efficient catalysts for chemical processes
 - D. Design of efficient processes in chemical plants from an understanding of microscopic molecular behavior
 - E. Design of a material with a given set of target properties
-

“Grand challenge A” or “Bundle A” in Table 1 has received recent emphasis because this area includes drug design. However, the biological activity due to a specific chemical is needed for other areas such as agricultural pesticide design and predictive toxicology. The potential for toxic impact of any chemical must be addressed before a chemical is manufactured, sold to the public, or released to the environment. Furthermore, the toxic behavior must be evaluated not only for human health issues but also for its potential ecological impact on plants and animals. Examining chemical toxicity is presently an extremely expensive process which can take a number of years of detailed tests. Such evaluations usually occur late in development and the

inability to anticipate the evaluation of toxicological testing can place large R&D investments at risk. Also, the possibility exists that unanticipated toxicological problems with intermediates and by-products can create liabilities. The cost of toxicology testing is generally too high to complete testing early in the development process. Thus reliable, cost-effective means for predicting toxicological behavior would be of great benefit to the industry.

Grand Challenge B in Table 1 is focused on the need to predict the fate of any compound that is released into the environment. For example, even if a compound is not toxic, a degradation product may show toxic behavior. Besides being toxic to various organisms, chemicals released into the environment can affect it in other ways. A difficulty in dealing with the environmental impact of a chemical is that the temporal and spatial scales cover many orders of magnitude from picoseconds to 100,000 years in time and from angstroms to thousands of kilometers in distance. Furthermore, the chemistry can be extremely complex and the chemistry that occurs on different scales may be coupled. For example, chemical reactions that occur on a surface may be influenced by not only the local site but sites that are not near but affect the local electronic structure or the surrounding medium.

Grand Challenges C and D in Table 1 are tightly coupled but are separated here because different computational aspects may be needed to address these areas. Catalysis and catalytic processes are involved in manufacturing most petroleum and chemical products and account for nearly 20% of the U.S. GDP. Improved catalysts would increase efficiency leading to reduced energy requirements, while increasing product selectivity and concomitantly decreasing wastes and emissions. Considerable effort has been devoted to the ab initio design of catalysts but such work is difficult because of the types of atoms involved (often transition metals) and of the fact that extended surfaces are often involved. Besides the complexity of the materials themselves, an additional requirement is the need for accurate results. Although computational results can often provide insight into how a catalyst works, the true design of a catalyst will require the ability to predict accurate thermodynamic and kinetic results. For example, a factor of two to four in catalyst efficiency can determine the economic feasibility of a process. Such accuracies mean that thermodynamic quantities should be predicted to within 0.1 to 0.2 kcal/mol and rate constants to within ~15%, certainly difficult, if not impossible, by today's standards. Even for

the nominally simple area of acid/base catalysis, many additional features may have to be included in the model, for example, the effects of solvation.

Another example of complexity is found in zeolites where the sheer size of the active region makes modeling studies difficult. Modeling of the surfaces present in heterogeneous catalysts is even more challenging because of the large numbers of atoms involved and the wide range of potential reactive sites. If the catalyst contains transition metals, the modeling task is difficult because of the problems in the treatment of the electronic structure of such systems with single configuration wave functions in a molecular orbital framework.

A molecular level understanding of chemical manufacturing processes would greatly aid the development of steady state and dynamic models of these processes. As discussed in subsequent sections, process modeling is extensively practiced by the chemical industry in order to optimize chemical processes. However, one needs to be able to develop a model of the process and then predict not only thermochemical and thermophysical properties but also accurate rate constants as input data for the process simulation. Another critical set of data needed for the models are thermophysical properties. These include such simple quantities as boiling points and also include more complex phenomena such as vapor/liquid equilibria phase (VLE) diagrams, diffusion, liquid densities, and the prediction of critical points. The complexity of process simulations depends on whether a static or dynamic simulation is used and whether effects such as fluid flow and mass transfer are included. Examples of complex phenomena that are just now being considered include the effects of turbulence and chaotic dynamics on the reactor system. A key role of computational chemistry is to provide input parameters of increasing accuracy and reliability to the process simulations.

Grand Challenge E in Table 1 is extremely difficult to treat at the present time. Given a structure, we can often predict at some level what the properties of the material are likely to be. The accuracy of the results and the methods used to treat them depend critically on the complexity of the structure as well as the availability of information on similar structures. For example, various QSPR (**q**uantitative **s**tructure **p**roperty **r**elationship) models are available for the prediction of polymer properties. However, the inverse engineering design problem,

designing structures given a set of desired properties, is far more difficult. The market may demand or need a new material with a specific set of properties yet given the properties it is extremely difficult to know which monomers to put together to make a polymer and what molecular weight the polymer should have. Today the inverse design problem is attacked empirically by the synthetic chemist with his/her wealth of knowledge based on intuition and on experience. A significant amount of work is already underway to develop the "Holy Grail" of materials design; namely, effective and powerful "reverse engineering" software to solve the problem of going backwards from a set of desired properties to realistic chemical structures and material morphologies which may have these properties. These efforts are usually based on artificial intelligence techniques and have, so far, only had limited success. Much work needs to be done before this approach reaches the point of being used routinely and with confidence by the chemical industry.

The achievement the goals outlined in Table 1 will require significant advances in a number of science and technology areas. A summary of the important scientific research areas needed to accomplish the goals outlined in Table 1 are given in Table 2 and a summary of technical issues that need to be addressed are given in Table 3. Below we highlight some of these issues.

Table 2

Research areas for Implementation of Grand Challenges^a

1. Accurate methods for calculating thermochemical and thermophysical properties, spectroscopy, and kinetics (A,B,C,D,E)
2. Efficient methods for generating accurate potential functions for molecular mechanics-based methods (A,B,C,D,E)
3. Improved methods for molecular dynamics simulations at long times for large ensembles (A,B,C,D,E)
4. Improved methods for including quantum effects (A,B,C,D,E)
5. Improved methods for including environmental effects such as solvent effects (A,B,C,D,E)
6. Efficient and accurate computational methods for treating solid state structures (B,C,D,E)
7. Improved optimization strategies for the determination of large, complex structures such as predicting protein structure from sequence (A,B,C,D,E)
8. Accurate methods for treating the upscaling problem: molecular ---> microscopic ---> mesoscopic ---> macroscopic (A,B,C,D,E)
9. New techniques for materials design and bulk property prediction (E)
10. New methods for predictive toxicology (A,B).
11. Integration of computational fluid dynamics (including lattice-Boltzmann approaches) with physics, chemistry, and biology to predict the behavior of reacting flows at different spatial and temporal scales (B,D)^b

^a Impact on Grand Challenge from Table 1 given in parentheses.

^b Additional research area for implementation of Grand Challenges

Table 3

Technology needs for Implementation of Grand Challenges^a

1. High performance, scaleable, portable computer codes for advanced (massively parallel) computer architectures (A,B,C,D,E)
2. Improved problem solving environments (PSE's) to make computational tools more widely accessible (A,B,C,D,E)
3. Improved database and data analysis technologies (A,B,C,D,E)
4. Computer-aided synthesis methods with a focus on materials (E)
5. Computer architectures, operating systems, and networks (A,B,C,D,E)

^a Impact on Grand Challenge from Table 1 given in parentheses.

^b Technology issue for implementation of Grand Challenges

There are a number of methods for obtaining accurate molecular properties. One can now push the thermochemical accuracy to about 0.5 kcal/mol if effects such as the proper zero point energy, core/valence effects, and relativistic effects are considered. Predicting kinetics can be considered as an extension of thermochemical calculations if one uses variational transition state theory. Instead of just needing an optimized geometry and calculated second derivatives at one point on the potential energy surface, this information is required at up to hundreds of points. It is necessary to incorporate solvent effects in order to predict reaction rate constants in solution. The prediction of rate constants is critical for process and environmental models. Predicted rate constants (computational kinetics) have already found use in such complex systems as atmospheric chemistry, design of chemical vapor deposition (CVD) reactors, chemical plant design, and combustion models. Spectroscopic predictions are increasing in their accuracy but it is still difficult to predict NMR chemical shifts to better than a few ppm, vibrational frequencies to a few cm^{-1} or electronic transitions to a few tenths of an electron volt for a broad range of complex chemicals.

There is a real need for accurate methods for predicting accurate thermophysics for gases and liquids. For gases, certain properties can be predicted with reasonable reliability based on the interaction potentials of molecular dimers and transport theory. For liquids, such properties can be predicted by using molecular dynamics and grand canonical Monte Carlo (GCMC) simulations. The GCMC simulations are quite reliable for some properties for some compounds but they are very dependent on the quality of the empirical potential functions. Such predictions, today, are much less reliable for mixtures or if ions are present.

The whole area of potential functions needs to be carefully addressed. Potential functions are needed for all atomistic simulations, e.g., molecular dynamics and energy minimizations of materials, polymers, solutions, and proteins; Monte Carlo methods; and Brownian dynamics. However, reliable potential functions are not available for all atoms and all bond types as well as for a wide range of properties such as polarization due to the medium. At present, it is very time-consuming to construct potential functions. A robust, automated potential function generator for producing a polarizable force field for all atom types needs to be developed. It needs to be able to incorporate both the results of quantum mechanical calculations as well as the empirical data.

There is a critical need to be able to take atomistic simulations such as molecular dynamics to much longer time scales. At present, it is routinely possible to study atomistic systems (or systems represented as interacting atoms, such as proteins and polymeric systems) for periods of the order of nanoseconds. However, much longer time scales are needed for the study of such problems as phase transitions, rare events, kinetics, and long time protein dynamics for protein folding. Even today, long runs on current computing systems create as-yet-unresolved data issues due to massive amounts of data generated. For example, a single timestep of a million-atom simulation easily manipulates tens of megabytes of data. While a reasonable strategy for short simulations of small systems is to dump configurations every 10th or 50th timestep for later analysis, this is clearly not an option for large-scale simulations over long time frames. Methodologies for implementing and modifying data analysis "on-the-fly" must be developed and refined.

The question of reaching macroscopic time scales from molecular dynamics simulations cannot be solved solely by increases in hardware capacity, since there are fundamental limitations on how many time steps can be executed per second on a computer, whether parallel or serial. One can scale the size of the problem with increasing numbers of processors but not to longer times. To cover macroscopic time scales measured in seconds while following molecular dynamics time steps of 10^{-15} seconds requires the execution of on the order of 10^{15} time steps. Even with five orders of magnitude increases in clock cycles, the required computations will take days. Between now and 2020, clock rates will undoubtedly increase but not by this magnitude. Hence, the long time problem in molecular dynamics will not be solved purely by hardware improvements. The key is the development of theoretically-sound, time-coarsening methodologies which will permit dynamics-based methods to traverse long time scales. Brownian dynamics with molecular-dynamics-sampled interactions and dynamic Monte Carlo methods are promising possibilities for this purpose.

A technology issue that will have an enormous impact on computational chemistry is that of computer architectures, operating systems and networks. The highest performance today is pushing 0.75 Tflops of sustainable performance on highly tuned code. The biggest technical issue is how to deal with nonuniform memory access (NUMA) and the associated latency for data transfer between memory on distributed processors. The latest step for large scale computers are massively parallel computing systems based on symmetric multi-processors (SMP). The goal is 10's of petaflop performance by 2020. This will be achieved by improvements in the speeds of individual chips which have been doubling every 18 months, although the cost of building plants to produce them may lead to a lengthening of the time to double processor speed. There will need to be significant improvements in switches as well as in memory speeds and I/O devices (disks) will need to be much faster and cheaper. There is a real need for significant advances in application software for usable teraflop to petaflop performance to be achieved as well as improvements in operating systems (OS). One major issue will be the need for single-threaded OS's that are fault-tolerant as the reliability of any single processor means that some will fail on a given day. It is the issue of operating systems, especially for large-scale batch computing, that is likely to hold up the ability to broadly address the computational Grand Challenge issues raised above.

In summary, rapid advances on many fronts suggest that we will be able to address the complex computational Grand challenges outlined above. This will fundamentally change how we will do chemistry in the future in research, development and in production. Getting there will not be simple and will require novel approaches including the use of teams from a range of disciplines to develop the software, manage the computer systems and perform the research.

PROCESS CONTROL AND INSTRUMENTATION

The process control and instrumentation issues identified in Vision 2020 include changes in the way plants operate, computer hardware improvements, the merging of models for design, operations, and control, development of new sensors, integration of measurement and control, and developments in advanced control. In the factory of the future, the industrial environment where process control is carried out will be different than it is today. In fact, some forward-thinking companies believe that the operator in the factory of the future may need to be an engineer as done in Europe. Because of greater integration of the plant equipment, tighter quality specification, and more emphasis on maximum profitability while maintaining safe operating conditions, the importance of process control will increase. Very sophisticated computer-based tools will be at the disposal of plant personnel. Controllers will be self-tuning, operating conditions will be optimized frequently, fault detection algorithms will deal with abnormal events, total plant control will be implemented using a hierarchical (distributed) multivariable strategy, and expert systems will help the plant engineer make intelligent decisions (those he or she can be trusted to make). Plant data will be analyzed continuously, reconciled using material and energy balances and nonlinear programming, and unmeasured variables will be reconstructed using parameter estimation techniques. Digital instrumentation will be more reliable, will be self-calibrating, and composition measurements which were heretofore not available will be measured on-line. There are many industrial plants that have already incorporated several of these ideas, but no plant has reached the highest level of sophistication over the total spectrum of control activities.

We are now beginning to see a new stage in the evolution of plant information and control architectures. The last 20 years of progress in computer control has been spurred by acceptance across a wide spectrum of vendors of the distributed control hub system for process control, which was pioneered during the 1970s by Honeywell. A distributed control system (DCS) employs a hierarchy of computers, with a single microcomputer controlling 8 to 16 individual control loops. More detailed calculations are performed using workstations, which receive information from the lower-level devices. Set points, often determined by real-time optimization, are sent from the higher level to the lower level. With the focus now on enterprise integration, automation vendors are now implementing Windows NT as the new solution for process control, utilizing personal computers in a client-server architecture rather than the hub-centric approach used for the past 20 years. This promotes an open application environment (open control systems) and makes accessible the wide variety of pc object-oriented software tools (e.g., browsers) that are now available.

The demand for smart field devices is rising rapidly. It is desirable to be able to query a remote instrument and determine if the instrument is functioning properly. Of course digital-based rather than analog instruments have the key advantage that signals can be transmitted digitally (even by wireless) without the normal degradation experienced with analog instruments. In addition, smart instruments have the ability to perform self-calibration and fault detection/diagnosis. Smart valves include PID (proportional-integral-derivative) control resident in the instrument, which can permit the central computers to do more advanced process control and information management. It is projected that installations of smart instruments can reduce instrumentation costs by up to 30% over conventional approaches. There has been much recent activity in defining standards for the digital, multidrop (connection) communications protocol between sensors, actuators, and controllers. In the U.S. the concept is called fieldbus control, and vendors and users have been working together to develop and test interoperability standards via several commercial implementations.

When data become readily available at a central point, it will be easier to apply advanced advisory systems (e.g., expert systems) to monitor the plant for performance as well as detect and diagnose faults. Recent efforts have built on the traditional single variable statistical process

control (SPC) approach and extended it to multivariable problems (many process variables and sensors) using multivariate statistics and such tools as principal component analysis. These techniques can be used for sensor validation to determine if a given sensor has failed or exhibits bias, drift, or lack of precision.

In the area of process modeling industrial groups are beginning to examine whether it is possible to achieve a seamless transition between models used for flowsheet design and simulation and models used for control. The CAPE-OPEN industrial consortium in Europe and other groups in the U.S. are working towards an open architecture for commercial simulators to achieve "plug and play" using company-specific software such as physical property packages. The extension of these steady-state flowsheet simulators to handle dynamic cases is now becoming an active area (e.g., linking Aspenplus to Speedup). The goal is to have models for real-time control that run at 50 to 500 times real-time, but this will require increased computational efficiency and perhaps application of parallel computing.

A new generation of model-based control theory has emerged during the past decade that is tailored to the successful operation of modern plants, addressing the "difficult" process characteristics encountered in chemical plants shown in Table 4.

Table 4

**Process Characteristics That Must Be
Treated by Advanced Control**

- Time Delays
 - Nonminimum Phase
 - Disturbances
 - Unmeasured Variables
 - Noise
 - Time-Varying Parameters
 - Nonlinearities
 - Constraints
 - Multivariable Interactions
-

These advanced algorithms include model predictive control (MPC), robust control, and adaptive control, where a mathematical model is explicit in developing a control strategy.

In MPC control actions are obtained from on-line optimization (usually by solving a quadratic program or QP), which handles process variable constraints. MPC also unifies treatment of load and set-point changes via the use of disturbance models and the Kalman filter. MPC can be extended to handle nonlinear models, as shown in Figure 2.

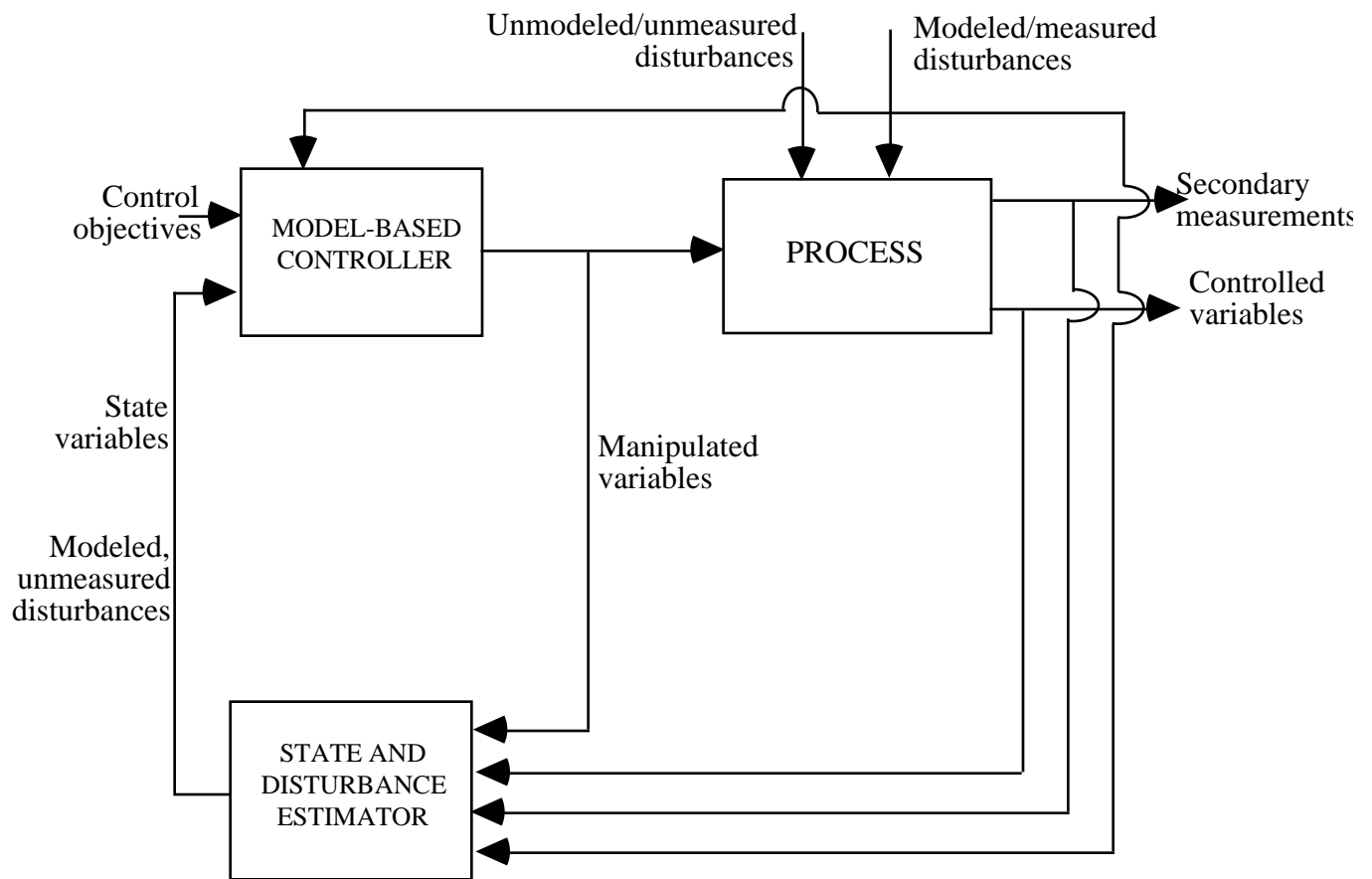


Figure 2. Generalized Block Diagram for Model Predictive Control

The success of Model Predictive Control (MPC) in solving large multivariable industrial control problems is impressive. Model predictive control of units with as many as 10 inputs and 10 outputs is already established in industrial practice. Computing power is not causing a critical bottleneck in process control, but larger MPC implementations and faster sample rates will

probably accompany faster computing. Improved algorithms could easily have more impact than the improved hardware for the next several years. MPC will appear at the lowest level in the DCS, which will reduce the number of PID loops implemented.

Adaptive control implies that the controller parameters should be adapted in real-time to yield optimal performance at all times; this is often done by comparing model predictions with on-line plant data and updating the process model parameters. The use of nonlinear models and controllers is underway in some applications. Some of the new versions of MPC are incorporating model adaptation, but so far adaptive control has not had much impact. This is due to problems in keeping such loops operational, largely because of the sensitivity of multivariable adaptive controllers to model mismatch.

Recent announcements by software vendors indicate that the combination of process simulation, optimization, and control into one software package will be a near-term reality, i.e., a set of consistent models across R&D, engineering, and production stages, with increased emphasis on rigorous dynamic models and the best control solutions. Software users will be able to optimize plant-wide operations using real-time data and current economic objectives. Future software will determine the location and cause of operating problems and provides a unified framework for data reconciliation and parameter estimation in real-time.

There are still many questions to be answered regarding the connection between modeling and control. This includes the explicit modeling information needed to achieve a particular level of control performance, the fundamental limitations on control performance even for perfect models, and the tradeoffs between modeling accuracy, control performance, and stability.

Process Measurement and Control Workshop

In recognition of the needs and challenges in the areas of process measurement and control, a workshop entitled “Process Measurement and Control: Industry Needs” was convened in New Orleans, March 6-8, 1998. Publications from the workshop will appear in Vol. 23, Issue No. 2 (1999) of *Computers and Chemical Engineering* in the future. The goals of the workshop were:

- (1) To survey the current state-of-the-art in academic research and industrial practice in the areas of measurement and control, particularly as they apply to the chemical and processing industries. The extent of integration of measurements with control is a particular focus of the survey;
- (2) To identify major impediments to further progress in the field and the adoption of these methods by industry; and
- (3) To determine highly promising new directions for methodological developments and application areas.

The workshop emphasized future development and application in eight areas:

- Molecular Characterization and Separations
- Nonlinear Model Predictive Control
- Information and Data Handling
- Controller Performance Monitoring
- Sensors
- Estimation and Inferential Control
- Microfabricated Instrumentation Systems
- Adaptive Control and Identification

See <http://fourier.che.udel.edu/~doyle/V2020/Index.html> for further information on workshop findings.

As an example of a specific roadmap the second topic (nonlinear model predictive control or NMPC) has been mainly of academic interest so far, with a few industrial applications involving neural nets. What is needed is an analysis tool to determine the appropriate technology (NMPC vs. MPC) based on the process description, performance objective, and operating region. There is also a desire to represent complex physical systems so that they are more amenable to optimization-based (and model-based) control methods. The improved modeling paradigms should address model reduction techniques, low-order physical modeling approaches, maintenance of complex models, and how common model attributes contribute pathological features to the corresponding optimization problem. Hybrid modeling, which combines fundamental and empirical models, and methodologies for development of nonlinear models (e.g., input sequence design, model structure selection, parameter adaptation) deserve attention. More details are contained in the URL for this workshop.

Chemical Instrumentation

Chemical analysis is a critically important enabling technology essential to every phase of chemical science, product and process development, and manufacturing control. Advances in chemical measurement over the past two decades have greatly accelerated progress in chemical science, biotechnology, materials science, and process engineering. Chemical measurements also play a key role in numerous related industries, such as pharmaceutical, pulp and paper and food processing. During recent years, impressive advances have been made in the resolution, sensitivity, and specificity of chemical analysis. The conduct of analytical chemistry has been transformed by advances in high-field superconducting magnets, multiple-wavelength lasers, multiplexed array detectors, atomic-force microscopes, scanning spectral analysis, and the integration of computers with instrumentation. These methods have been extended to the detection and spectral characterization of molecular structure at the atomic level.

A Vision 2020 workshop was held in March, 1997, to assess future directions for R&D in chemical instrumentation. Research needs identified included:

- transfer of analytical laboratory capabilities into plants, incorporating ease of maintenance and support, utilizing new technology and molecular scale devices
- improved real-time characterization of polymers (molecular weight distribution, branching)
- improved structure/property/processing modeling capability, especially macromolecular products such as biomolecules and biopolymers
- physical/chemical characterization of solids and slurries
- on-line characterization of biotechnological processes
- new approaches for sampling and system interlinks to control and information systems
- self-calibrating and self-diagnostic (smart) sensors
- identification of processes needing microfabricated instruments and development of corresponding models/control systems
- integration of data from multiple sensors for environmental compliance, product development, and process control, including soft sensors.
- advanced measurement techniques to support combinatorial chemistry in catalysis and drug discovery

For more details see the URL www.nist.gov/cstl/hndocs/ExternalTechnologyBundles.html.

PROCESS OPERATIONS

Three of the four technology thrust areas of the Vision 2020 document, namely supply chain management, information, systems, and manufacturing and operations, address the business and manufacturing functions of the chemical enterprise. This clearly reflects the importance that efficient production and distribution of chemical products has on the economic viability of the enterprise now and over the next 25 years. In this section, we highlight the role that technical computing and information systems play as technology enablers for effective operation and present the most important challenges and needs which must be addressed in the

future. The discussion of research issues draws on a workshop on "R&D Needs in Systems Technologies for Process Operations" which was convened in July 1998. For full details of the workshop report the reader is invited to consult the Vision 2020 web site identified earlier in this paper.

In the present context, process operations refers to the management and use of human, capital, material, energy, and information resources to produce desired chemical products safely, flexibly, reliably, cost effectively, and responsibly for the environment and community. The traditional scope of operations encompasses the plant and its associated decision levels, as shown in Figure 3.

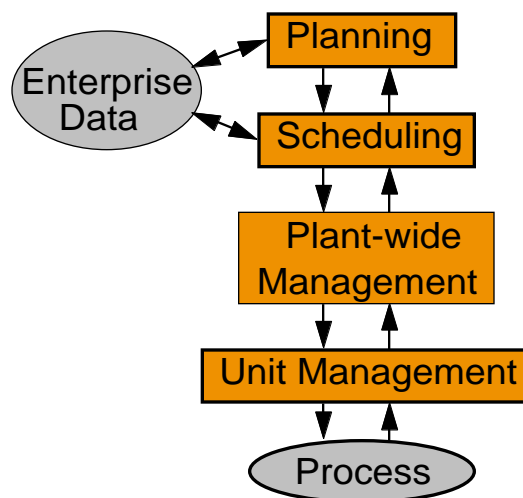


Figure 3. Plant Decision Hierarchy

The key information sources for the plant operational decision hierarchy are the enterprise data, consisting of commercial and financial information, and the process itself. The unit management level includes the process control, monitoring and diagnosis, and on-line data acquisition functions. The plant wide management level serves to coordinate the network of process units and to provide cost-effective setpoints via real-time optimization. The scheduling decision layer addresses time varying capacity and manpower utilization decisions, while the planning level sets production goals which meet supply and logistics constraints. Ideally there is bi-directional communication between levels, with higher levels setting goals for lower levels

and the lower levels communicating constraints and performance information to the higher levels. In practice the information flow tends to be top down, invariably resulting in mismatches between goals and their realization.

In recent years this traditional view of operations has been expanded to include the interactions between suppliers, multiple plant sites, distribution sites and transportation networks, and customers. The planning and management of this expanded network, referred to as the supply chain, poses challenging decision problems because of the wide temporal scale and dynamics of the events which must be considered, the broad spatial distribution and dimensions of the entities which must be managed, and the high degree of uncertainty which it experiences because of changing market factors and variable facilities uptimes and productivity. Clearly the supply chain is a dynamic system of high complexity. Nonetheless, the vision proposed for the operational domain is that in 2020 the success of a chemical enterprise will depend upon how effectively it generates value by dynamically optimizing the deployment of its supply chain resources. The seven factors critical to the achievement of the vision are:

- speed to market: time from piloting to the market place
- efficient operation: in terms of operational cost and asset utilization
- health, environment and safety: factors affecting workers and the community
- quality workforce: management, engineering, and operational staff which is process and business literate
- technology infrastructure: processes, instrumentation, and equipment as well as information systems and technical computing
- quality and product integrity: work process for producing the product right the first time
- functional integration: bi-directional linkage of all decision levels of the supply chain

The challenges and needs which must be met under each of these factors in order to allow the vision of the dynamically optimized supply chain to be realized certainly do require innovations which extend beyond developments in information and computing technology alone.

However, it is clear that the infrastructure for storing and sharing information and technical computing tools which exploit that information constitute the key enabling technology. The information that must be stored and shared includes transactional information, resources costs and availabilities, plant status information, models, and model solutions. This diversity of information types must be effectively organized and must be sharable using reliable high speed networks. The enabling technical computing components include model building methods and tools; solution algorithms using numerical, symbolic, and logic based methods; visualization and interpretation methods; interfaces for use and training; and integration of all of these components into usable decision support tools.

Present Status

At the present time the essential elements of information technology to support operations are at hand both in terms of data infrastructure and network connectivity. Commercial data base management systems and transactional systems are common in the industry. Plant information systems and historians are in widespread use and enterprise-wide data base system installations are growing explosively. UNIX or Windows NT based networks are common and internet and web-based applications are growing rapidly. Despite this growth there is as yet only limited integration of business and manufacturing data and only limited tools to facilitate effective use of this data. Indeed the general consensus is that corporations are drowning in a sea of data. The challenge is to extract information and knowledge and thus to derive value from this data.

The present status of technical computing of relevance to process operations can best be characterized as a patchwork of areas of different levels of development. At the planning level, multi-time period linear programming tools capable of handling large scale systems are well developed and have been in use, especially in the petroleum/petrochemical sector since the 1970's. Real-time, plant-wide optimization applications using steady state process models are growing rapidly in the petrochemical domain, although some of the statistical and computational formulations and algorithms remain under active development. The methodology for scheduling of multipurpose batch and continuous production facilities has been under

investigation since the late 1970's, initially using rule-based and heuristic randomized search methods and, more recently, using optimization based (mixed integer linear programming) methods. Application of the latter in industry is as yet limited but growing. Successful solution of problems involving over hundred equipment items and several hundred distinct production tasks have been reported, although the deployment of the technology still requires high levels of expertise and effort. As noted in the section on Process Control and Instrumentation, tools for abnormal situation management are as yet in their infancy, although significant industry led developments are in progress. Linear model predictive control has been practiced in the field since the early 1980's although the theoretical supports for the methodology were developed later. Plant data rectification has been practiced since the mid-1980's but typically applications have been confined to linear models and simple statistical descriptions of the errors in the measurements.

Challenges

The long term challenges for the application of computing technology can be divided into four major areas:

- conversion of data into knowledge
- support tools for the process
- support tools for the business
- training methodologies

The development of tools which would facilitate conversion of the extensive data contained in enterprise information systems into actionable information and ultimately knowledge is of highest priority. Some of the capabilities which need to be pursued include soft-sensors, data rectification techniques, trend analysis and monitoring methods, and data visualization techniques. Soft-sensors are critical to simplifying the detection of erroneous measurements by localizing the detection logic. Data rectification refers to the process of condensing and correcting redundant and inaccurate or erroneous process data so as to obtain the most likely status of the plant. Trend analysis and monitoring refers to the process of using process

knowledge and models to identify and characterize process trends so as to provide timely predictions of when and what corrective action needs to be taken. Data visualization is an essential element for facilitating understanding of process behavior and tendencies.

The decision support tools for the process include stream-lined modeling methodology, multi-view systems for abnormal situation management, nonlinear and adaptive model predictive control, and process optimization using dynamic, and especially hybrid models. Model building is generally perceived to be a key stumbling block because of the level of expertise required to both formulate process models and to implement them using contemporary tools. The goal is to make model building and management rapid and reliable and to create environments in which the models associated with the various levels of the operational decision hierarchy will be consistent and unified. The role of abnormal situation management systems is to identify plant trends, to diagnose likely causes and consequences, and to provide intelligent advisory to plant personnel. While components which address portions of this entire process have been under investigation for the past decade, full integration of the various qualitative and quantitative support tools remains to be realized. Needed developments in process control have been discussed in an earlier session and hence will not be reiterated here, except to note that control of batch and other intentionally dynamic processes needs to be given considerably more attention. Finally, the optimization of models consisting of differential algebraic systems and especially differential algebraic systems with discrete elements is essential to the realization of the vision for process operations. The latter type of so-called hybrid systems is particularly relevant to process which involve batch and semicontinuous operations.

The overall goal of these decision support methodologies for the process is to realize the integrated model-centered paradigm for process operation shown in Figure 4. Under this paradigm all of the decision levels of the operational hierarchy are fully integrated through the shared use of consistent robust models. Models serve as the central repository of process knowledge. Information flows from the lower levels to the higher levels to ensure that decisions fully consistent with the status and capacity of the production resources are made.

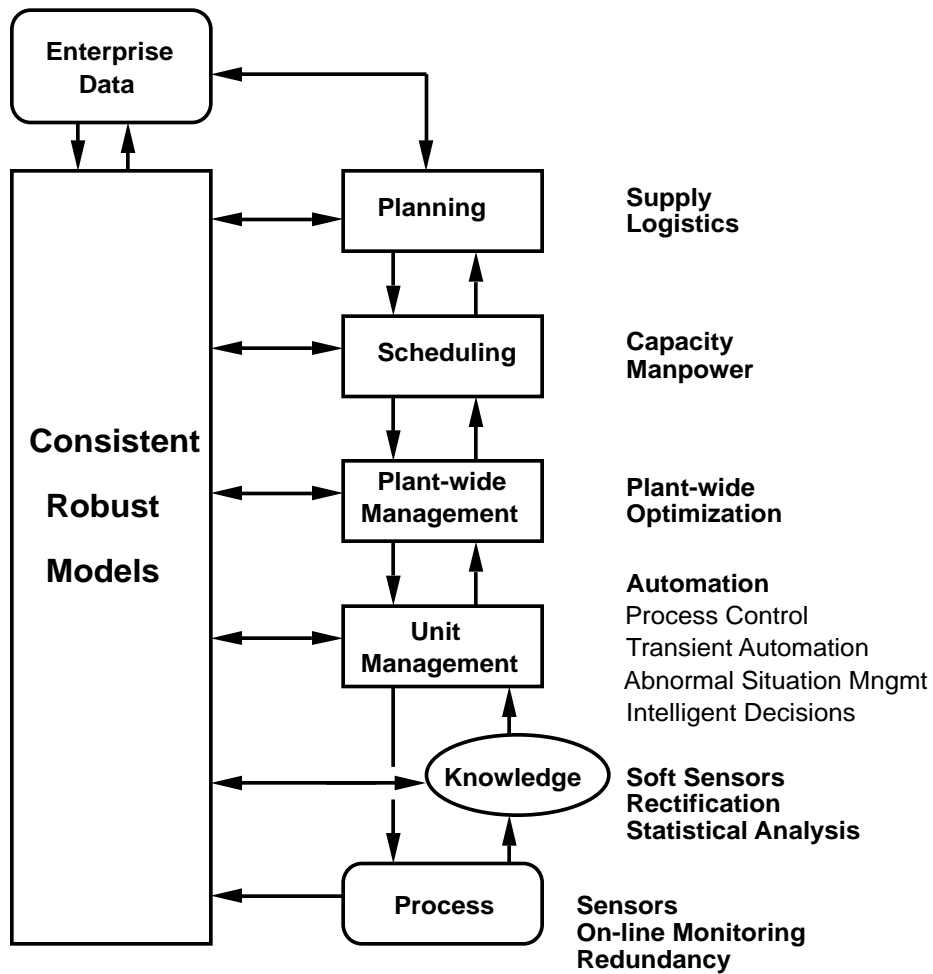


Figure 4. Integrated Model-Centered Operation

The third area of need is in the development of tools to support the overall business decision processes. The objective is to expand the envelope beyond the process itself and to encompass the business processes that are essential to driving manufacturing and the entire supply chain . The tools include improved sales and market forecasting methodologies, supply and logistics planning techniques, methodologies for quantitative risk assessment, optimization based plant scheduling methods, business modeling frameworks, and approaches to dynamic supply chain optimization. Optimization-based scheduling requires the solution of very high dimensionality models expressed in terms of discrete 0-1 variables. The key need is to be able to solve scheduling problems with hundreds of thousands of such variables reliably and quickly. Such capabilities need to be extended to allow treatment of models which encompass the entire supply

chain and to quantitatively address business issues such as resource and capital planning associated with the supply chain, siting of new products, and the impact of mergers and acquisition on the supply chain.

Finally, in order to realize the benefits of the developments in the other three areas, it is necessary, indeed essential, to create training methodologies for the workforce. These computer-based training methodologies must make efficient use of students' time, recognize differences in levels of expertise, and employ extensive visualization tools, including virtual reality components. Methods must also be developed to aid process staff in the understanding of models and the meaning of the solutions resulting from the various decision support tools which are based on these models. Such understanding is critical both to the initial adoption of such models and to the continuous improvement process since it is only from understanding the constraints of the existing operation and their implications that cost effective improvements can be systematically generated.

In conclusion, the process operations specific information systems and technical computing developments outlined above are essential to the realization of the goal of the dynamically optimized supply chain. Continuing increases in computing power, network bandwidth, and availability of faster and cheaper memory will no doubt facilitate achievement of this goal. However, the scope and complexity of the underlying decision problems require methodological developments which offer effective gains that are orders of magnitude beyond the likely increases in raw computing power and communication bandwidth. Process-oriented technical computing really does play the pivotal role in the future of process operations.