Actuated Wingsuit for Controlled, Self-Propelled Flight

Human flight has been an incredible driving force behind human innovation. From early epics such as the story of Icarus to Leonardo da Vinci's designs of hang gliders and helicopters, and from the Wright brothers' first planes to our first trip to the moon, human flight has accompanied and inspired us throughout much of our scientific and technological progress. The field has made spectacular advances since its inception and now permeates modern systems of everyday life, from aerial photography to air ambulances and commercial travel. For many of us, however, human flight has now become a necessity rather than a source of inspiration, evoking visions of crowded cabins with tiny windows and cramped seating rather than majestic thoughts of wind-beneath-your-wings airborne travel.

This applied research project recaptures the spirit and ambition of the man-on-the-moon-type projects that have so successfully fueled human inventiveness and innovation in the past. Simultaneously, the project tackles a perceived lack of large-scale, multidisciplinary projects that capture the public's imagination.

The Challenge: Unconstrained Human Flight

The goal of this project is to achieve unconstrained human flight by building on existing wingsuit technology (see image above) and by leveraging research in lightweight structures and propulsion systems, nonequilibrium aerodynamics, and algorithmic methods for the control of highly dynamic systems (see image on next page). The result will be an actuated wingsuit that can be actively controlled by the flyer.

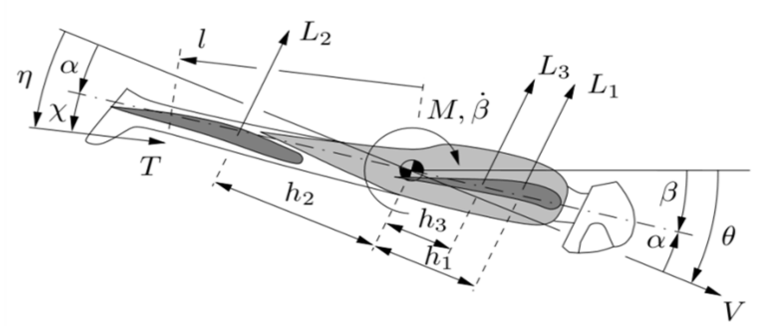
As with previous endeavors in human flight, this project requires a multidisciplinary effort by researchers in mechanical and electrical engineering, materials sciences, controls, human-machine interaction, and related disciplines. Similar to the principal investigator's previous projects, the current effort allows students from semester projects and those from bachelor, master, and PhD thesis programs to become involved in the challenge. Achieving unconstrained human flight is a highly multidisciplinary challenge, requiring competencies—and offering learning opportunities—that span the entire R&D cycle, from the derivation of theoretical results to their experimental validation, practical implementation, and revision.

Existing wingsuits are purely passive designs that allow a pilot to achieve glide ratios of approximately 2.5 (2500 m of horizontal travel for every 1000 m of vertical descent); in comparison, flying squirrels achieve glide ratios of at most 2.0.

***Contributors:** Raffaello D'Andrea, Mathias Wyss, and Markus Waibel, ETH Zurich, Switzerland*

Research Project Under Way

To date, few academic studies on unconstrained human flight have been conducted; however, a recent study by our group (see citation below) highlights the potential for academic work in this field and identifies several specific areas for research contributions. Both data and algorithmic results of current investigations are freely shared for further analysis, promoting multidisciplinary research in the area. The project follows a rigorous scientific approach with results presented at major international conferences and published in high-profile journals.



$$\dot{V} = \frac{T\cos\eta - D}{m} + g\sin\theta$$

$$\dot{\theta} = \frac{1}{V}\left(g\cos\theta - \frac{L + T\sin\eta}{m}\right)$$

$$\ddot{\beta} = \frac{M}{I}$$

$$L_k = \rho V^2 c_{L_k}, \quad c_{L_k} = s_k\alpha_k + q_k$$

$$\alpha_1 = \alpha - \frac{h_1\dot{\alpha}}{V}, \quad \alpha_2 = \alpha + \frac{h_2\dot{\alpha}}{V}, \quad \alpha_3 = \alpha - \frac{h_3\dot{\alpha}}{V}$$

$$c_m := -s_1h_1 + s_2h_2 - s_3h_3$$

$$c_{md} := s_1h_1^2 + s_2h_2^2 + s_3h_3^2$$

$$M = \rho V^2\left(c_m\alpha + \frac{c_{md}}{V}\dot{\alpha} + (-h_1q_1 + h_2q_2 - h_3q_3)\right)$$

$$L = \rho V^2 c_L$$

$$D = \rho V^2 c_D = \rho V^2\left(c_p + \frac{c_L^2}{c_i}\right)$$

$$c_L = 1.17\alpha + 0.39$$

$$\begin{bmatrix} \ddot{\beta} \\ \dot{\beta} \\ \dot{V} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & 0 & a_3 \\ 1 & 0 & 0 & 0 \\ 0 & a_4 & a_5 & a_6 \\ 0 & a_7 & a_8 & a_9 \end{bmatrix} \begin{bmatrix} \dot{\beta} \\ \Delta\beta \\ \Delta V \\ \Delta\theta \end{bmatrix}$$

A first-principles model (physics and mathematics) of how one flies a wingsuit (from G. Robson and R. D'Andrea, Longitudinal stability analysis of a jet-powered wingsuit, Proc. AIAA Guidance, Navigation, and Control Conference, San Antonio, Texas, 2010)

Join us!

The project is headquartered in Zurich, Switzerland, which offers a unique combination of world-class research facilities at ETH Zurich, mountain geography with suitable launch sites and easy access, and an active community of skydivers and wingsuit flyers. Efforts by the project team are leveraged through a community of researchers at ETH Zurich, pilots, seasoned wingsuit flyers, and technologists whose aim is to allow humans to experience unconstrained flight.

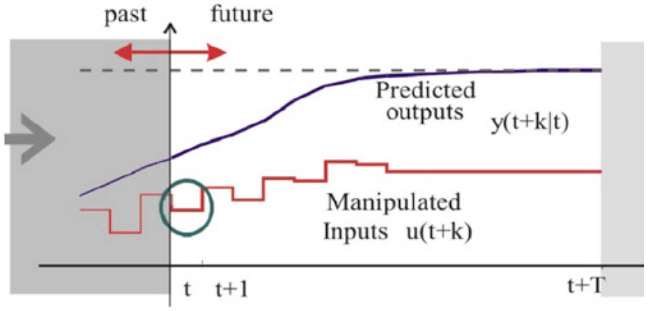
To join this challenge, visit http://raffaello.name/dynamic-works/actuated-wingsuits and contact info@raffaello.name.

Addressing Automotive Industry Needs with Model Predictive Control

The automotive industry is facing significant hurdles as it strives for dramatic improvements in fuel economy, reduced emissions, vehicle safety, and overall positive driving experience, including automated driving. Advanced control technology is recognized as a key enabler for overcoming these challenges; however, further advances in control research will be needed before solutions can be commercialized.

The automotive domain poses demanding control system requirements: control loops need to be able to operate in milliseconds, the computational infrastructure is limited to an embedded controller, and stability, robustness, and performance must be maintained over millions of individual vehicles and for hundreds of thousands of kilometers driven under vastly different climate and operating conditions.

Recent advances in the theory, algorithms, and synthesis methods of model predictive control (MPC) have attracted considerable interest from the automotive industry. Although production applications are rare (if any), this attractive and intuitive method has shown considerable promise in applications ranging from R&D prototypes to fully functional production-like vehicles. Significant results have been achieved, but numerous opportunities exist for further research and development.



Three Examples of Automotive Challenges for Advanced Control

Cornering and Stability Control

Advanced control facilitates effective active front steering (AFS) and optimal coordination with differential braking (DB) for superior vehicle cornering and driver assistance. Such control can greatly enhance handling and stability, especially in adverse weather conditions when operating the vehicle/tires in extreme nonlinear regions.

Idle Speed Control

Idle speed control (ISC) is one of the most basic and representative automotive control problems and is still one of the most important aspects of engine operation. The main objective is to keep the engine speed as low as possible for superior fuel economy while preventing engine stalls. Critical factors of ISC are limited actuator authority and time delays in the control channels.

Energy Management for Hybrid Vehicles

In hybrid electric vehicles (HEVs), the energy stored in the battery can be used by electric motors to supplement the engine. With the aid of an advanced controller, the HEV energy management system decides optimal power distribution under practical operating constraints.

MPC Basics

MPC operates by repeatedly solving a constrained optimal control problem initialized at the current estimate of the system state (see figure above). The formulation incorporates a system model, operational constraints, and a user-defined cost function. The use of the current state in the repeated optimization results in feedback that increases robustness with respect to open-loop optimal control.

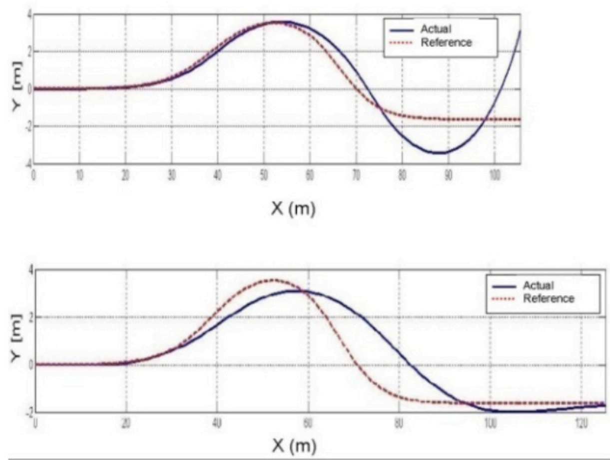
Advances in MPC technology, increased computational power of electronic control units, and increasing performance, safety, and emission requirements have attracted interest from the automotive industry. Key relevant advantages of MPC are the capability of handling constraints on inputs and states, the intuitive design, even for multivariable systems, and the ability to define control objectives and relative priorities by cost function.

Contributors: Davor Hrovat and H. Eric Tseng, Ford Research and Innovation Center, USA, and Stefano Di Cairano, Mitsubishi Electric Research Laboratory, USA

Enhanced Driving on Snow

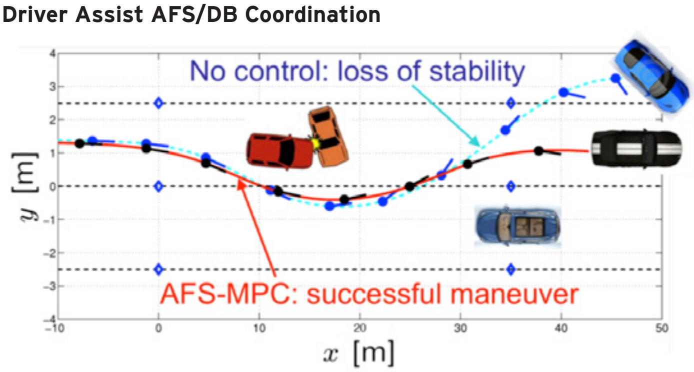
MPC can be used to exploit a model of the lateral vehicle dynamics to select optimal actions for steering and braking, hence achieving superior yaw rate tracking and overall vehicle cornering performance. Research prototype MPC controllers have been developed and tested on slippery surfaces. Results from road tests for double lane change (DLC) maneuvers on snow surfaces are shown below.

Autonomous AFS



The reference (red) and actual (blue) DLC trajectories on snow at 55 kph are shown for a fixed PID gain steering robot designed for asphalt surfaces (top) and for an MPC system that incorporates vehicle stability state constraints (bottom). Neither car has electronic stability control. The MPC anticipates that aggressive steering could lead to loss of traction, thus trading off its initial tracking error for future gains and successful completion of the DLC maneuver.

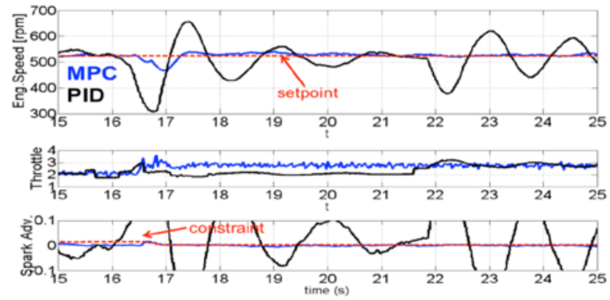
Driver Assist AFS/DB Coordination



DLC with (black and white car) and without (blue car) MPC-based driver assistance that coordinates active front steering and differential braking. The test executed on a snowy road at approximately 50 kph shows that MPC helps the driver to enlarge the car operating envelope. (In both cars, electronic stability control was disabled for this test.)

Disturbance Rejection for Idle Speed Control

MPC is attractive for ISC because it is capable of dealing with limited actuator authority and time delays. In tests, a multi-input, multi-output (MIMO) MPC controller outperformed conventional controllers by optimally coordinating the constrained control channels and handling the delays, and anticipating loads through the predictive model.



Disturbance rejection of MPC-ISC (blue) and conventional PID-based ISC (black) on a V8 engine. In this test, the AC compressor turned on after the power steering pump was already fully engaged. As a result, the spark authority was heavily reduced to avoid knocking, but MPC maintained control by adjusting the action on the throttle.

Power Smoothing for HEV Batteries

Battery management controllers have been developed for various HEV configurations. For a series HEV configuration, the MPC controller exploits the battery power to smooth engine transients, thus achieving more efficient operation. The MPC smooths transients while guaranteeing battery state-of-charge and respecting power constraints with superior tradeoff between steady-state and transient efficiency. The MPC controller was implemented in a fully functional series HEV prototype and evaluated in standard fuel economy (FE) tests, showing sizable FE advantages.

Future Directions

MPC has shown significant potential in several automotive applications, but several challenges remain before its use in production vehicles is widespread. These include further mitigation of the computational effort, especially for nonlinear optimization, easy-to-use tuning methods, as well as theoretical development and tools for stability, robustness, and performance guarantees.

For more information: D. Hrovat, S. Di Cairano, and H.E. Tseng, I.V. Kolmanovsky, The development of model predictive control in automotive industry: A survey, Proc. IEEE International Conference on Control Applications, 2012.

Avoiding Pilot-Induced Oscillations in Energy-Efficient Aircraft Designs

The aviation industry is vital to global economic well-being. In the U.S. alone, civil aviation provides more than one million jobs, a trade value of more than $75 billion, and a total contribution to the economy of almost $300 billion. Nevertheless, the aviation industry also has a negative impact on the environment and energy usage. In the U.S., air travel fuel use is 7% of fuel consumed for transportation, and jet fuel produces 65 million metric tons of CO_2 per year, or 4% of CO_2 emissions from energy usage nationwide.

To improve fuel efficiency and reduce environmental impact in the aviation sector, a variety of next-generation, energy-efficient aircraft design concepts are being explored. Many of these design concepts, however, rely on relaxed static or dynamic stability, which will likely lead to a resurgence in vehicle stability and control problems—particularly pilot-induced oscillation (PIO). Research in flight control of next-generation, energy-efficient aircraft to avoid PIO will be critical in enabling these new aircraft design concepts to operate safely in the future.

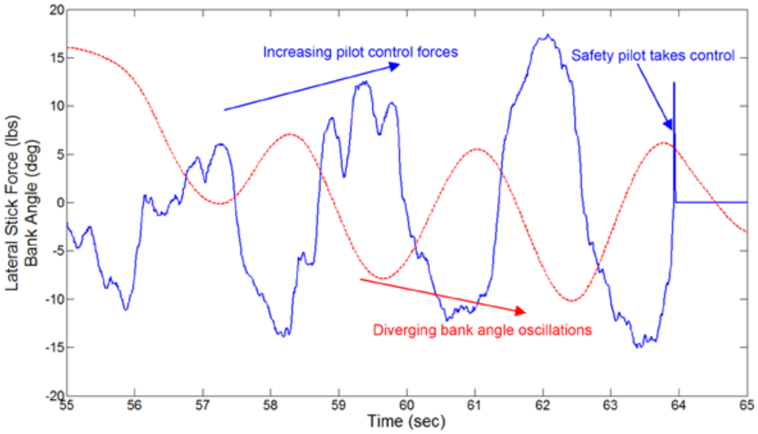
Aircraft design concepts for improved energy efficiency and environmental impact (Source: NASA)

Pilot-Induced Oscillation

A pilot-induced oscillation is a sustained or uncontrollable, inadvertent oscillation resulting from the pilot's efforts to control the aircraft. While PIOs can be easily identified during post-flight data analysis, often pilots do not know they are in a PIO—from their perspective the aircraft appears to have "broken."

When approaching instability, linear system performance degrades in a manner that is predictable to a pilot. As nonlinearities are introduced, however, gradual degradations can be replaced by sudden changes in aircraft behavior, resulting in the so-called "flying qualities cliff." With few warning signs provided by the aircraft as one approaches such a cliff, loss of control can easily occur. A common nonlinearity that is a major factor in PIO is control surface rate limiting. This phenomenon can introduce a delayed response. When the plane does not respond to the cockpit controls as expected, the pilot may move the controls more aggressively. The aircraft will ultimately overrespond, causing the pilot to reverse the control input and overreact again because of the delay. As this continues and develops fully into a PIO, the airplane response is essentially opposite of the pilot's command—for example, as the pilot commands a left bank, the airplane is in a right bank.

An example flight test PIO is shown at right. The pilot is attempting a precision landing with an aircraft response that is dominated by a rate-limited control surface response. The rate limit nonlinearity results in a PIO that increases with each cycle as the pilot attempts the final runway centerline capture. Note that the peak oscillation of the aircraft response (red) is opposite of the peak oscillation of the pilot command (blue) and that both are increasing in amplitude until the safety pilot takes control. (Source: STI)



Contributors: Diana M. Acosta, NASA Ames Research Center, USA; Yildiray Yildiz, Bilkent University, Turkey; David H. Klyde, Systems Technology Incorporated, USA

PIO Susceptibility of Energy-Efficient Aircraft Design Concepts

Pilot-induced oscillations have plagued aircraft since the beginning of aviation. Even the Wright brothers believed stability and control was their most difficult challenge, and analysis has shown that their aircraft was susceptible to PIOs. To avoid PIO tendencies, the commercial aviation industry has adopted common aircraft design conventions—for example, for the fuselage shape; wing and tail size, shape, and location; and propulsion unit location. These conventions result in highly stable aircraft that compromise energy efficiency.

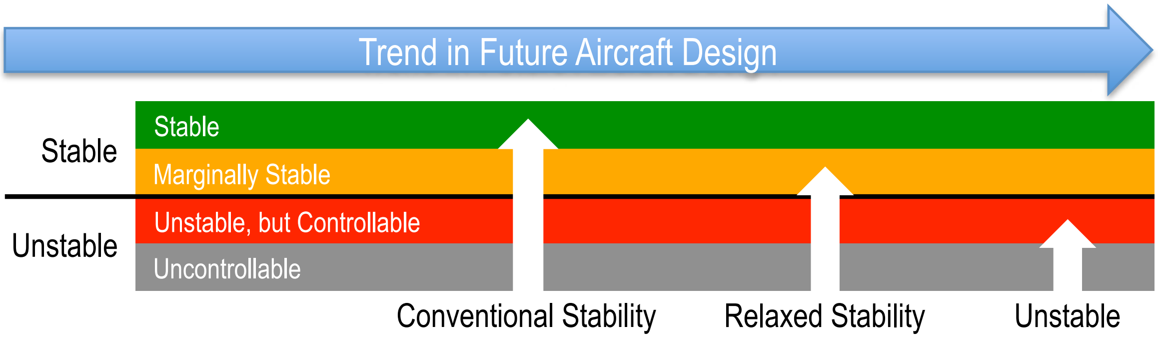
Future aircraft design concepts are beginning to deviate from many of today's common design conventions. Studies have identified relaxed static stability as a key technology for reducing fuel burn and cruise emissions. One design concept for a blended wing body (BWB) aircraft goes a step further with an unstable aircraft design augmented with closed-loop control to maintain stability, similar to modern fighter aircraft.

As the trend in aircraft design leads to marginally stable or unstable but controllable airframes, high levels of control power and feedback control augmentation are required to improve flying qualities and maintain closed-loop stability. In particular, best practices in PIO prevention recommend that an aircraft's actuation system exhibit sufficient rates and transient capability so as to avoid rate saturation of the surfaces. This poses a challenge for next-generation transport aircraft. For the BWB configuration, producing the power required to move the large control surfaces at a rate required for stability and control of the vehicle is technologically challenging.

Subject to traditional design practices, the strive toward energy efficiency and environmental compatibility in combination with the complexity of new designs will inevitably increase the susceptibility of future aircraft to PIO events. Technology is needed to mitigate the effects of PIO factors and allow aircraft to meet their potential in energy efficiency and environmental compatibility without abiding by constraining design practices.

Future Capabilities for PIO Avoidance and Recovery

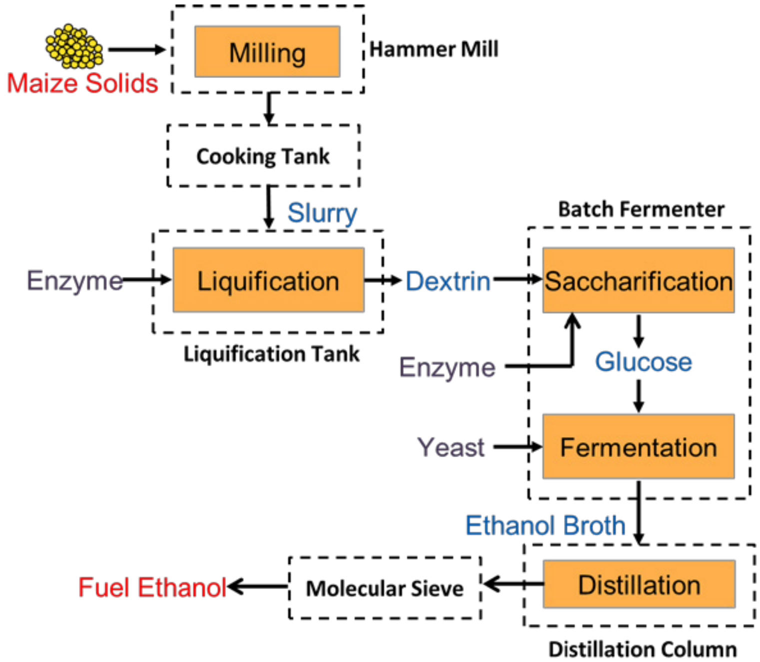
- Sensors will need to measure data pertinent to PIOs.
- Using the data collected from the sensors, estimation methods will need to identify or predict the onset of unfavorable dynamics (i.e., the approaching flying qualities cliff).
- Control effectors will need to provide sufficient control power with a fast response while being lightweight, producing little drag, and not requiring significant actuator power.
- Pilot interfaces, including visual and aural displays and cockpit controls, will need to inform the pilot of the situation and recommend an appropriate course of action.
- Control laws will need to determine appropriate actions for the pilot or a safe-mode autopilot and/or compensation for the flight control system.
- Flight control computers will need to be fast enough to complete computations without introducing computational time delay.



Future commercial aircraft are expected to rely increasingly on relaxed stability and unstable designs to improve fuel economy and reduce environmental impact.

Batch Control and Trajectory Optimization in Fuel Ethanol Production

Ethanol produced from fermentation of biomass-derived sugar is increasingly used as a transportation fuel, either neat or in petrol blends. As the world's largest biofuel producer, the U.S. produced more than 57 billion liters of ethanol in 2012, and the number will increase to 136 billion liters by 2022. Most of the ethanol in the U.S. is produced from maize-based plants, and more than 90% of the plants make use of the dry mill process, which involves four steps: milling, liquification, simultaneous saccharification and fermentation (SSF), and distillation. The SSF process, which is conducted in a batch fermenter, is considered the most important part of the entire production process. SSF breaks down dextrin into dextrose and converts dextrose into ethanol.



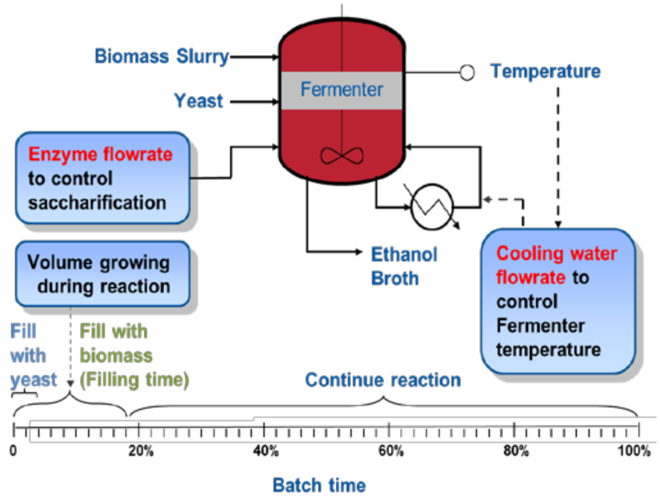
Flow chart of dry mill process

Current Progress

Fuel ethanol production is a $43 billion industry that contributes more than $8 billion in taxes to state, federal, and local governments. Researchers generally acknowledge that there is a gap between the theoretical maximum yield and the yield achieved in process plants. Therefore, even a low-percentage yield increase would result in hundreds of millions of dollars in profits. The most economic approach to increasing ethanol yield is to optimize operation of the SSF process during a batch.

Challenges

1. Modeling: A better understanding of the process represented by mathematical models is required. If models are derived from first principles, they usually consist of coupled ordinary differential equations (ODEs) resulting from component and energy balances.
2. Optimization: A major objective of batch control is for the manipulated variables to follow reference trajectories that maximize a performance index (e.g., the ethanol yield at the end of the run). However, there is no steady state and thus there are no constant setpoints over the course of a batch. Hence, a rigorous optimization strategy is required for control purposes.



Overview of SSF process and timeline of batch operation

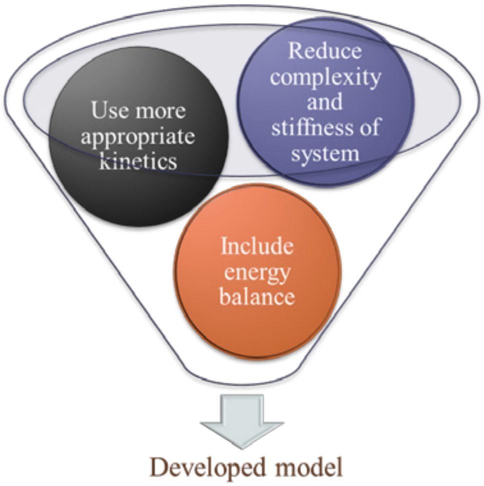
Contributors: Wei Dai and Juergen Hahn, Rensselaer Polytechnic Institute, USA

Modeling of SSF Process

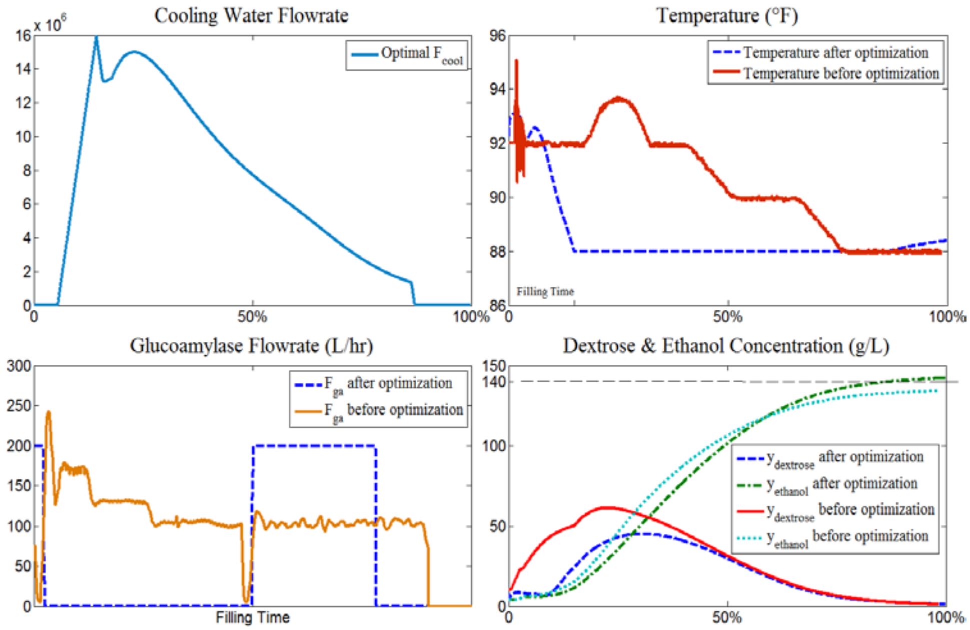
Existing models of the SSF process consist of dynamic balances of components such as the concentrations of yeast, dextrose, ethanol, and other substances. Three aspects have been investigated to improve the process model.

Trajectory Optimization

Maximizing the final ethanol yield is the key driver for researchers and engineers to find better control solutions for this process. This objective can be achieved by numerically computing the optimal setpoint trajectory of the controlled variables. A model composed of nine ODEs (eight component balances and one energy balance) and five path constraints is used here for illustration purposes. A simultaneous approach, which discretizes both the manipulated variables and the model, is used for the solution of the optimization problem.



Three aspects to improve a model



Optimal input profiles and simulation results after optimization

Future Work

More input variables, such as length of the fermentation process, could be considered for trajectory optimization.

Additionally, complex models that incorporate intermediate and branch reactions could be considered for trajectory optimization, whereas a simpler model could be used for fast model-based online control.

Results and Discussion

The optimal temperature profile has a high temperature during the filling phase and favors a low temperature for the remainder of the batch. The optimal enzyme addition profile indicates that glucoamylase should be injected into the fermenter toward the end of the filling phase rather than continuously during the entire phase. The dextrose concentration after optimization is well controlled within a reasonable range, and the ethanol concentration is increased after optimization by as much as 7%. Furthermore, relaxing the lower bound of the temperature constraint increases ethanol production by 11%. However, considering that a discrepancy always exists between the model and a real plant, these research findings will have to be validated in a plant setting.

For more information: W. Dai, D.P. Word, and J. Hahn, Modeling and dynamic optimization of fuel-grade ethanol fermentation using fed-batch process, Control Engineering Practice, 2013.

Biological Oscillators

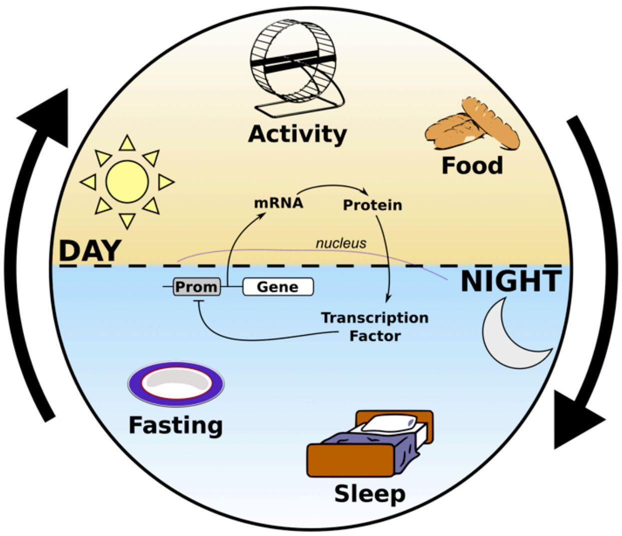
Periodic fluctuations in biological processes are found at all levels of life and are frequently the result of changes to gene expression. These rhythms play key roles in a variety of important processes, including circadian regulation, metabolism, embryo development, neuron firing, and cardiac rhythms. Oscillating gene regulatory networks function as finely tuned dynamic systems in which time-delayed negative feedback gives rise to sustained rhythms. Such rhythms display robustness to biological noise and evolutionary mutations while remaining acutely sensitive to such environmental cues as light or temperature.

The essential characteristics of biological oscillators can be represented by biophysical networks with many interacting species. Through the application of modeling and simulation tightly linked to experiment, systems biology provides a way to study such biophysical networks and hence to understand the mechanistic foundations of biological oscillators. Systems biology employs systematic measurement technologies such as genomics, bioinformatics, and proteomics to quantitatively measure the behavior of groups of interacting components in biophysical networks and harnesses mathematical and computational models to describe and predict dynamical behaviors.

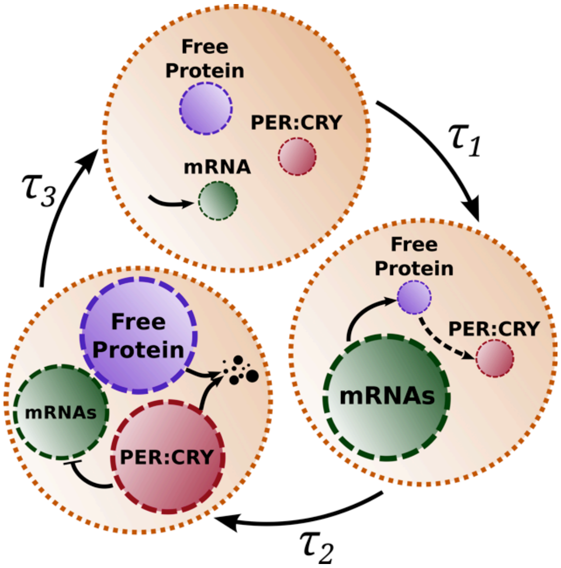
Why Is Systems and Control Theory Relevant?

Regulation, tracking, interactions, adaptation, robustness, communication, signaling, sensitivity, identification, dynamics, stability/instability, and causality are all concepts that are crucial in biological oscillators and have counterparts in the systems and control domain. Systems and control theory can be harnessed for

- Modeling of biological oscillators,
- Understanding the mechanisms of robustness and sensitivity,
- Reverse engineering of biological oscillators, and
- Control (e.g., in pharmaceutical regulation) and design (e.g., in synthetic biology) of rhythmic biological processes.



Genetic regulatory networks help coordinate important oscillatory behaviors, including circadian rhythms, shown here. Rhythmic light/dark cues optimize metabolic pathways for expected energy intake and demand.

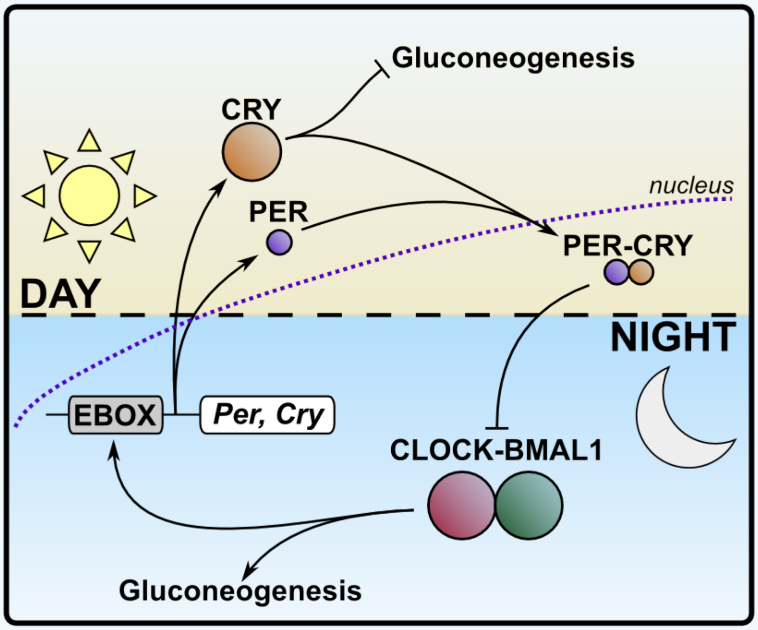


Sustained oscillations result from multiple sources of time delays. Transcription, translation, and degradation of repressive complexes (PER-CRY in mammalian circadian rhythms) all contribute to the oscillatory period.

Contributor: Francis J. Doyle III, University of California at Santa Barbara, USA

Example: Circadian Control of Mammalian Metabolic Pathways

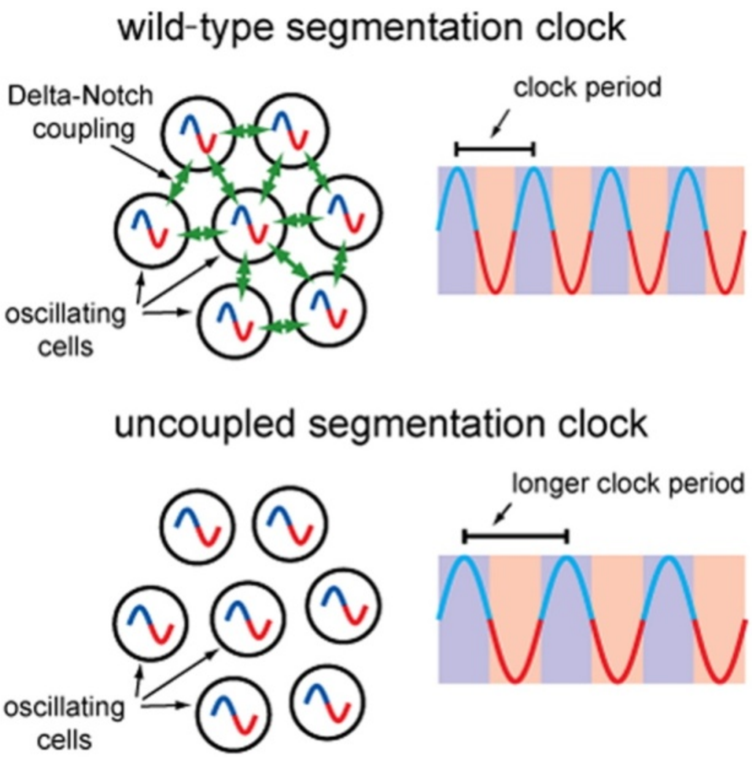
In mammals, circadian rhythms control the activity of many key metabolic pathways. For example, in the liver, the production of new glucose from energy reserves is repressed during the day, when meals are common, and activated at night, when they are scarce. However, in our current 24-hour society, circadian disturbances such as jet lag and shift work are also manifested as metabolic disorders. Without a keen understanding of the processes involved, appropriate behavioral or pharmaceutical therapies for long-term circadian disorders are difficult to find. By applying systems and control methodologies to models of circadian rhythms, the additional metabolic burden of irregular light schedules could possibly be alleviated.



Circadian rhythms control many aspects of mammalian metabolism, including the creation of new glucose in the liver (gluconeogenesis).

Example: How Is Network Oscillation Behavior Established From Cell Autonomous Oscillations?

Most biological rhythms are generated by a population of cellular oscillators coupled through intercellular signaling. One such system where this occurs is the segmentation clock shown at right, which governs the synchronized development of vertebrate embryos. Recent experimental evidence shows that the collective period of all cells may differ significantly from the autonomous period in the presence of intercellular delays. Although this phenomenon has been investigated using delay-coupled phase oscillators, a better understanding of the biological mechanisms that govern coupled oscillators is critical to connecting genetic regulatory networks to tissue-level oscillatory behavior.



Coupling independent oscillators can result in a collective period different than that of each cell. (Source: L. Herrgen et al., Curr. Biol. 20, 1244, 2010)

City Labs for Intelligent Road Transportation Systems

New Testbeds for Research in Transportation Control

Real-time management of city services, including traffic, energy, security, and information, is a new approach to the development of "smart cities." Low-cost and easy-to-deploy sensors and wireless communication protocols are enabling control system technologies to play a much-enhanced role in this area.

City-scale testbeds are now becoming available for research and experimentation, especially for intelligent transportation systems. Such "City Labs" are spearheading efforts to study and evaluate control systems in realistic arenas accessible to researchers and engineers. The ultimate goal is the implementation of new control systems to improve the daily lives of drivers and passengers, help traffic operators optimize the network, and reduce energy consumption and environmental impact.

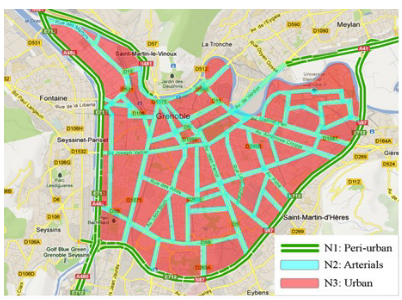
Current Traffic Management Approaches are Fragmented, Not Holistic

Nowadays, traffic problems are typically addressed at the level of a single vehicle or subsystem (e.g., in a specific arterial corridor or a part of an urban road). The current control and resource optimization strategies are inefficient when considering traffic at the global network level. Today's fragmented and uncoordinated approach is a significant obstacle for improving urban mobility and energy efficiency.

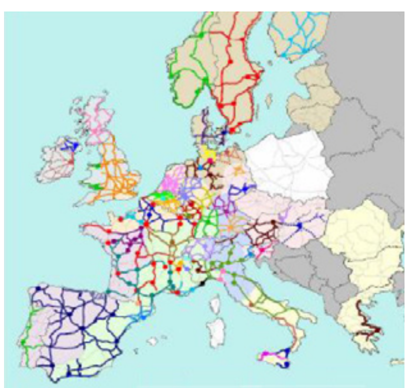
The Challenge of Heterogeneity in Road Transportation Systems

Intelligent transportation requires the modeling, analysis, and control of the transportation system as a whole. The diversity of elements in the system must be taken into account, including

- Vehicle classes (private cars, utility vehicles, trucks, buses);
- User groups (private, professional, public);
- Road networks (highways, arterial, urban);
- Transportation modes (high-speed roads, low-speed roads, bus lanes, tramways); and
- Implementation technologies (sensors, software, protocols).



Three main traffic networks in the city of Grenoble. The "peri-urban," "arterial," and "urban" networks are managed by different traffic authorities with little coordination or integration. Better control coordination with a holistic view is critical for optimal operation. (Source: NeCS team)



Main traffic management domains of the EU highway system. Optimal route planning at the EU level requires coordination of domain-level traffic management policies and better sharing of information. Technologies used in different networks and countries are heterogeneous and require greater integration. (Source: Easyway)

Contributor: Carlos Canudas de Wit, CNRS/GIPSA-Lab, France

From: ***The Impact of Control Technology,*** **2nd ed., T. Samad and A.M. Annaswamy (eds.), 2014. Available at www.ieeecss.org.**

Control Challenges for Intelligent Road Transportation

- *Exploiting new data sources:* Integration of various sensor technologies with different characteristics (radars, mobile phones, Bluetooth, video, magnetometers, etc.)
- *Secure and privacy-preserving data sharing:* Control with communication constrained by secure real-time information sharing and privacy-preserving data aggregation
- *Mathematical models:* New traffic models accounting for multiple modes, combination of micro and macro models, two-dimensional flows, and large network graphs
- *Model-based travel time forecasting:* Online prediction algorithms along multimodal networks
- *Coordinated control among subsystems:* Control methodologies and architectures for operating the road network as a whole in metropolitan areas
- *Optimal routing for dynamic traffic networks:* Online optimal planning algorithms accounting for traffic flow congestion and modern vehicle-to-network communication policies
- *Resilient traffic control:* Control strategies that account for the vulnerabilities introduced by subnetwork interconnections—resilience against malicious attacks on actuators (e.g., lights) and sensors

The Grenoble South Ring

This stretch of the Grenoble perimeter highway is 10.5 km long and includes 10 on-ramps and 6 off-ramps. Some 90,000 vehicles per day (5% trucks) travel on this road, taking 7 to 50 minutes for the trip. Sensing and actuation equipment includes

- 130 wireless magnetic sensors (flows and velocities),
- Four junctions with in-ramp queue measurements,
- Seven variable speed limit electronic panels (70-90 km/h), and
- Ramp metering (to be implemented).

Sensor data is collected every 15 seconds and transferred to a server with a maximum latency of one sample period (15 seconds). Variable speed limits can be actuated directly from the traffic operation center.

Grenoble Traffic Lab (GTL)

In collaboration with national traffic operators and local public authorities, a City Lab for traffic management has been launched in Grenoble. GTL collects information from the road-traffic infrastructure in real time, with minimum latency and fast sampling periods. Main components include

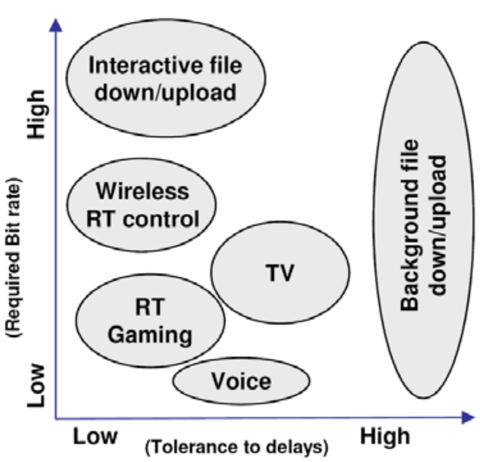
- A dense wireless sensor network providing macroscopic traffic variables,
- Real-time data collection and archiving for traffic classification and demand prediction,
- Traffic forecasting algorithms for traveling-time prediction up to 45 minutes ahead,
- Control algorithms for access control and variable speed limits, and
- A showroom and a micro simulator to validate algorithms and user interaction.

GTL objectives include transfer of research results to industry with the collaboration of multiple stakeholders.

The Grenoble Traffic Lab is supported by INRIA, CNRS, and University Joseph Fourier. For more information, visit http://necs.inrialpes.fr/.

Control Challenges in Mobile Telecommunications

Mobile telecommunications technology is having an unprecedented impact on human society. Currently, there are more than 6.8 billion cellular subscribers worldwide, and more than 4 million new phones are sold per day! Global revenue from subscriptions exceeds $6.5 trillion annually. New services are also appearing, including TV, web browsing, tethering, and real-time gaming. As in all areas of technology, the successful operation of modern telecommunication systems depends in part on highly sophisticated real-time control. The opportunities for advanced control are enormous, but the area poses new and interesting challenges. For example, the control problems in telecommunications have their own distinctive characteristics, including varying demands on data rate and delay latency (see figure below). Also, the control is necessarily carried out over the telecommunication channel itself, giving rise to networked control issues.



Bit rate/delay issues for mobile services (RT: real time)

Control in Telecommunications

Control appears at various levels in mobile telecommunications:

Power control: Used to adjust the signal-to-interference ratios (SIRs) of users so they are maintained at an appropriate level at the base station. These loops operate with significant delay, use coarsely quantized control signals, which limits the slew rate, and are subject to fast and large channel gain variations.

Link adaptation: Used to optimize performance by controlling the transmission rate jointly with the transmit power. Link adaptation uses quantized measurements related to signal quality to select transmission rates and modulation.

Scheduling: To maximize certain performance measures, 3G, 4G, and 5G cellular systems schedule users in the downlink and the uplink in real time; 4G and 5G systems schedule users in both frequency and time to capitalize on favorable instantaneous channel conditions.

Backhaul control: When bit rates over the air-interface increase, so does the need to control the data flow upstream of the radio link to minimize round-trip delays and ensure that data is always available for transmission to scheduled users.

Multipoint transmission and reception: Techniques such as coordinated multipoint transmission utilize noncollocated antennas. Data synchronization and inter-transmit-point control loops are then needed to align the powers and the delays for the various antennas.

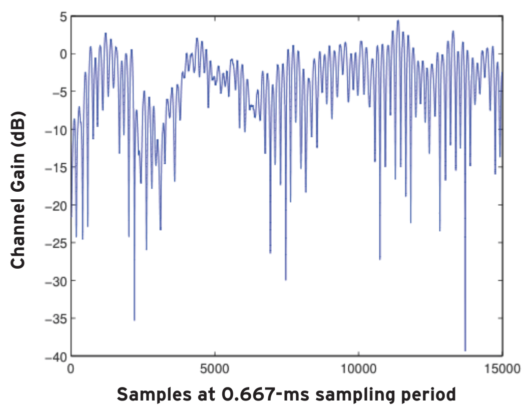
Telecommunications in Control

Not only is control central to modern mobile telecommunication systems, the reverse is also true; that is, the next-generation control systems are likely to be wireless-based due to flexible connectivity and reduced costs.

In the ongoing definition of 5G, not only bit rate but also delay is at the focal point. Examples of applications under consideration include:

- Vehicular control, e.g., collision avoidance and platoon driving
- Haptic control, e.g., advanced gaming and remote surgery

Wireless round-trip delays of the order of 1 ms are needed to enable these emerging applications.



Channel gain variation at 3 km/h

Contributors: Graham Goodwin, Mauricio Cea, and Katrina Lau, University of Newcastle, Australia, and Torbjörn Wigren, Ericsson AB, Sweden

Challenges

Challenges associated with telecommunication control problems include:

Power control

- Heavily quantized signals (1 or 2 bits)
- Delays
- Lost control signals
- Highly variable channel fading
- Significant nonlinearities
- Multivariable interactions

Scheduling

- Large and variable delays
- High uncertainty in channel gains

Link adaptation

- Coarsely quantized feedback
- Delays
- Discrete control action

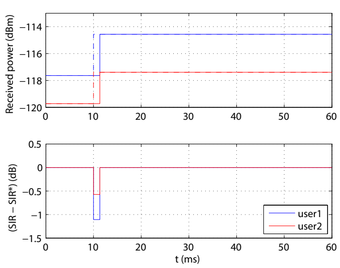
Opportunities for Advanced Control

The nature of the control challenges in telecommunications provides exciting opportunities for sophisticated control tools; however, the application of these tools in the telecommunications context raises new, and as yet not fully solved, challenges.

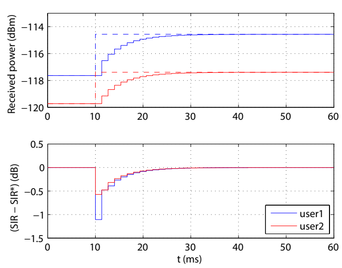
- New ideas in networked control are needed for the power control loops. This is difficult because only 1 bit (3G) or 2 bits (4G) can be sent per sample and bits may be lost.
- New scheduling algorithms are needed that exploit the dynamics and inherent constraints of the scheduling loop. High (stochastic) uncertainty, variable delays, high complexity, short sampling periods, and the need for low latency make this extremely challenging.
- Novel implementations of nonlinear filtering could be applied to load estimation and for prediction of channel fading, grant utilization, and intercell interference. However, high state dimension, severe nonlinearities, and fast sampling rates all pose challenges.
- New insights into decentralized control are needed to implement solutions. The stochastic nature of the problem and high demands on quality of service for users make this challenging.

An Example of the Use of Advanced Control in Mobile Telecommunications

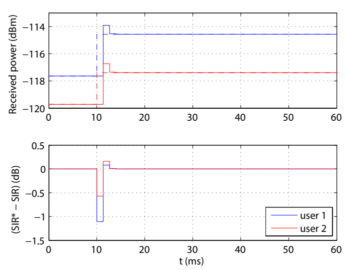
The following figures relate to inner-loop power control for 3G systems.



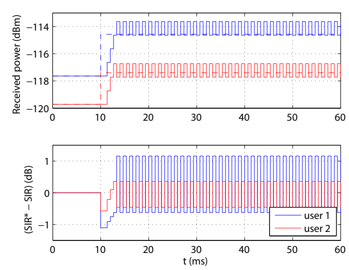
(a) One user at a time



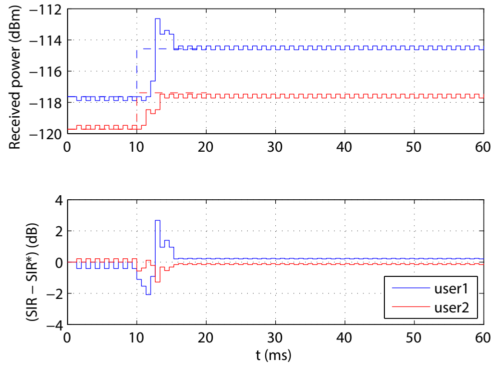
(b) Impact of multivariable interaction



(c) Impact of decoupling



(d) Impact of quantization



(e) Impact of nonlinear adaptive control with 1-bit quantizer

Figure (a) shows a minimum-variance inner-loop power controller for one user at a time with no control quantization.

Figure (b) shows that multivariable interactions significantly degrade the regulation performance when multiple (in this case, two) users are considered.

Figure (c) shows the impact of using a nonlinear decoupling algorithm. Note that the performance is now very similar to that achieved for a single user.

Figure (d) shows the impact of control signal quantization (to 1 bit), which undermines the gains achieved by decoupling.

Figure (e) shows that the decoupling performance is largely recovered by using a sophisticated nonlinear adaptive controller that optimally compensates for the 1-bit constraint.

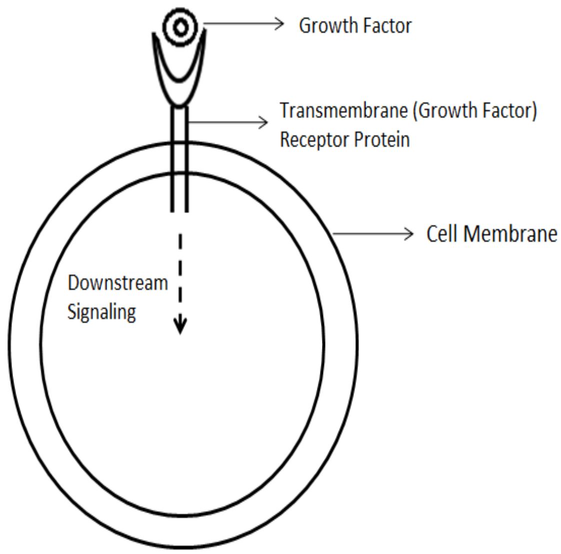
Control Engineering for Cancer Therapy

Cancer encompasses various diseases associated with loss of control in the mechanisms that regulate the cell numbers in a multicellular organism. It is usually caused by malfunctions in the cellular signaling pathways. Malfunctions occur in different ways and at different locations in a pathway. Consequently, therapy design should first identify the location and type of malfunction and then arrive at a suitable drug combination. Given the dynamics, feedback, and other complexities involved, systems and control approaches can be instrumental in both the identification and therapy aspects of cancer treatment.

Introduction to Molecular Biology and Cancer

Multicellular organisms such as humans are made up of about 100 trillion cells. A cell is the basic unit of life, and nothing smaller than a cell can be considered to be truly living. Each cell is like a massive factory where thousands of reactions are performed every second inside compartmentalized organelles. By far the largest organelle in a human cell is the *nucleus*, which contains the genetic information written using the four-letter language of DNA.

Cell division is under tight control and depends on signaling mechanisms from neighbors. Furthermore, in the absence of survival signals from neighbors, a cell will activate an intracellular suicide mechanism, *apoptosis*, and eliminate itself. It is this dynamic equilibrium between controlled cell proliferation and cell death that maintains the tissue architecture in adult multicellular organisms. When this dynamic equilibrium is disrupted, it leads to the formation of tumors, which are initially benign. Subsequently, these tumors can become *malignant* or *cancerous* by acquiring the ability to invade surrounding tissue. *Metastases* can occur as these tumors develop the ability to spread to distant sites via the blood or lymphatic system.



Under normal conditions, growth factors (or mitogens) external to a cell come and bind their respective transmembrane receptors. This binding leads to a signal transduction cascade inside the cell, as illustrated above, which ultimately results in the activation of genes involved in cell proliferation. Aberrant behavior such as mutations in some of the genes in the signal-transduction cascade can cause the cell proliferation genes to be activated even when the external growth factor stimulus is missing, and this is one of the mechanisms by which uncontrolled cell proliferation and possibly cancer can develop. Another mechanism by which cancer can develop is through the mutational inactivation of genes that serve as molecular brakes on cell division.

Genetic Regulatory Networks and Pathways

Genes (and other biological molecules such as proteins) interact with each other in a *multivariate* fashion. Historically, biologists have focused on experimentally studying the *marginal* cause-effect interactions between a small number of biological molecules, leading to what is called *biological pathway information*. This piecemeal approach, primarily studied using simpler organisms, has been very successful in unraveling the sequences of steps involved in metabolic processes; however, it has failed to completely elucidate the intricate cellular signaling that is associated with higher organisms such as humans. With the advent of high-throughput technologies such as microarrays (which can simultaneously provide measurements of the activity status of thousands of genes), several approaches have recently been proposed for modeling the multivariate interactions between genes, leading to what are called *genetic regulatory networks*. The study of these networks has been carried out using differential equations, Bayesian networks, Boolean networks, and their stochastic generalizations, the so-called probabilistic Boolean networks (PBNs). PBNs can be equivalently represented as homogenous Markov chains. By introducing external treatment as a control variable in the PBN, we obtain a *controlled Markov chain* or a *Markov decision process*. By formulating cancer treatment as the problem of moving the stationary distribution of a genetic regulatory network from an undesirable state to a desirable one, and trading off the costs involved, one can formulate an optimal control problem that can be solved using dynamic programming and its variants. One challenge with this approach is the huge amount of data needed to reliably infer a genetic regulatory network.

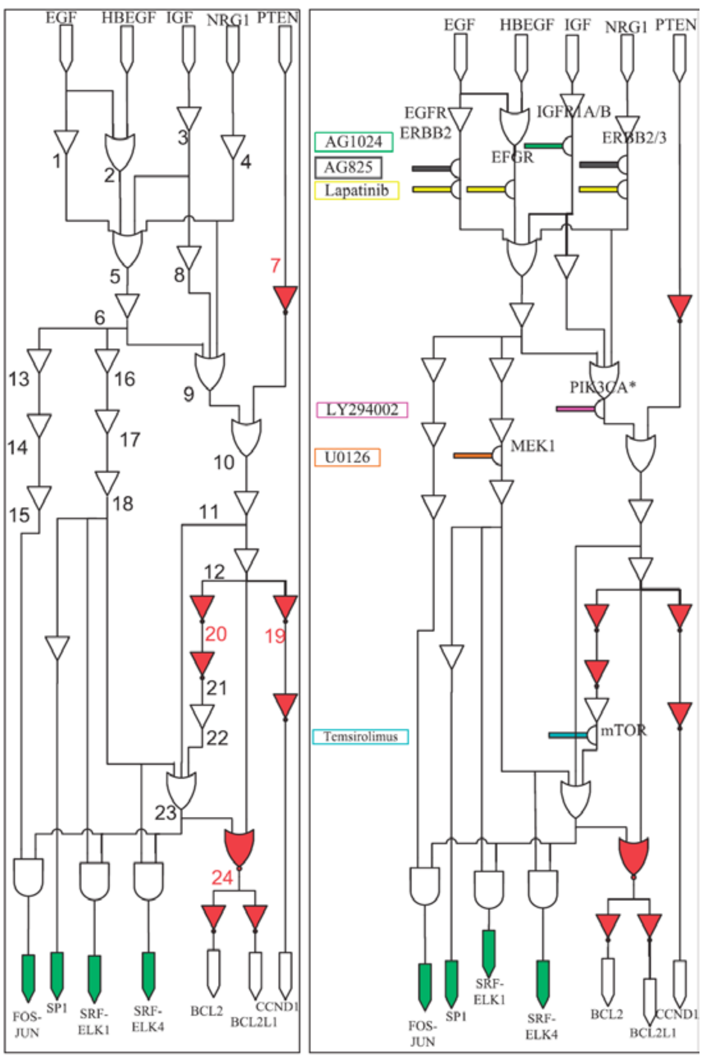
Contributor: Aniruddha Datta, Texas A&M University, USA

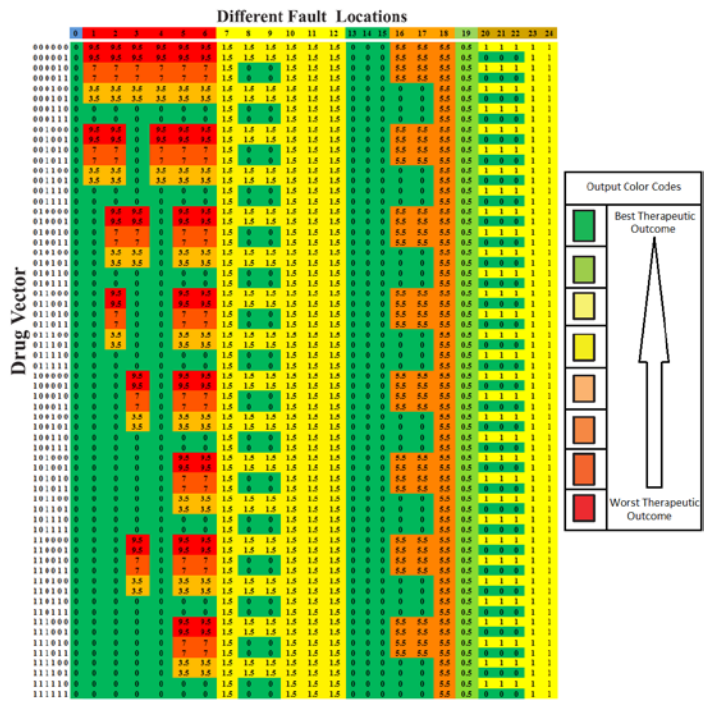
Combination Therapy Design Based on Pathway Information

Biological pathway information, despite its limitations, can also be useful in therapy design. In cases where feedback loops are absent, the pathway can be modeled as a digital circuit using logic gates. Computer simulation of the digital circuit can aid in identifying where in the pathway signal breakdown can occur using only input/output information. Furthermore, the effect of different anticancer drugs, whose main mechanism of action is to cut off the downstream signaling, can be superimposed on the digital circuit at the known appropriate points of intervention. Thereafter, this circuit can be used to make predictions about the efficacy of different drug combinations. See the figure below for an example.

Future Challenges in Experimental Validation

The predictions regarding combination therapy for cancer merit experimental validation, perhaps using cancer cell lines. However, experimental validation must deal with several complexities, such as (i) possible inaccuracies in the pathway model; (ii) the presence of feedback loops that have not been accounted for; (iii) the presence of multiple faults; and (iv) the heterogeneity of cancer tissue. Addressing each of these problems is a research issue in its own right and could significantly contribute to cancer treatment. Here, it is encouraging to note that issues of this type, such as uncertainty and robustness, have been extensively studied in engineering disciplines such as control theory, although adapting the ideas to the current context will still be a challenge.





With the circuit above, the efficacy of drug combinations can be predicted—shown here for six drugs and 24 possibilities for signaling breakdown (represented by the numbered gates in the left figure).

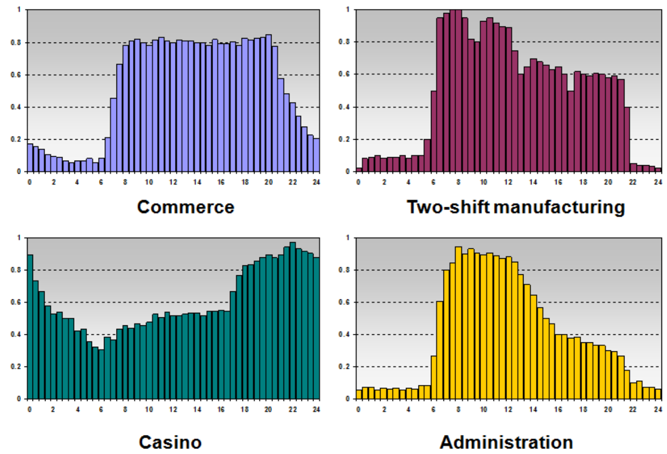
Left: Digital circuit model of the growth factor signaling pathway. The input signals are the growth factors and a brake on cell division; outputs are proteins/genes reporting on cell proliferation and programmed cell death. Right: Effect of anticancer drugs overlaid on the circuit.

For more information: A. Datta and E.R. Dougherty, Introduction to Genomic Signal Processing with Control, CRC Press, 2007; R. Layek et al., Cancer therapy design based on pathway logic, Bioinformatics, vol. 27, no. 4, pp. 548-555, 2011; M. Vidyasagar, Control System Synthesis: A Factorization Approach, MIT Press, Cambridge, MA, 1985.

Control for Energy-Efficient Buildings

Globally, the building sector is responsible for 40% of annual energy consumption and up to 30% of all energy-related greenhouse gas emissions; hence the interest in increasing energy efficiency in buildings. Heating, ventilation, and air conditioning (HVAC) is the principal building system of interest, but there are others: lighting, active façade systems, renewable generation sources, and storage.

Real-time control and optimization can help building owners and tenants minimize energy consumption and costs based on inputs from occupants, local utilities, and weather conditions. Challenges for implementation of advanced control solutions include the heterogeneity and complexity of typical building environments. Recent developments in building automation systems are addressing these and other challenges.



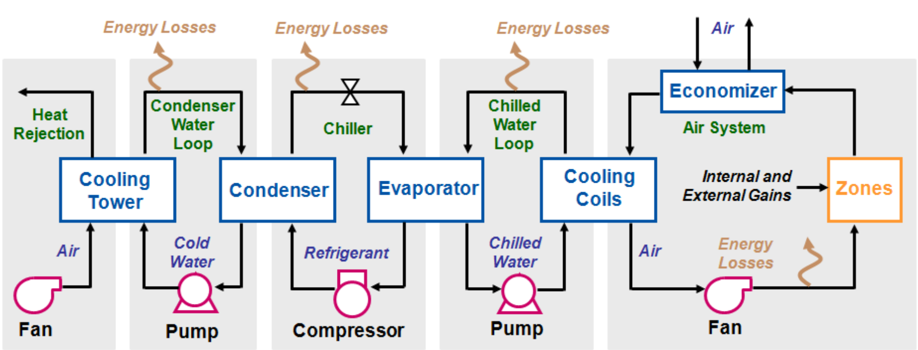
Daily consumption profiles: Every building has a unique consumption pattern

Trends in Building Automation

The cloud and data analytics. Cloud computing enables the retention of more detailed data about the facility as well as integration of automation and other business data. This in turn enables more powerful building analytics, which can better inform facility managers about likely equipment faults, deviations from expected energy use, or underperforming controllers.

Intelligent devices. Building controllers increasingly embed intelligent software and computational power, which enable delivery of enhanced functionality. Smarter devices can enable automated reconfiguration or parameter tuning in response to changes in the environment. Also, such devices will be able to share information with other devices, automatically synchronize, and support deployment of distributed optimization concepts.

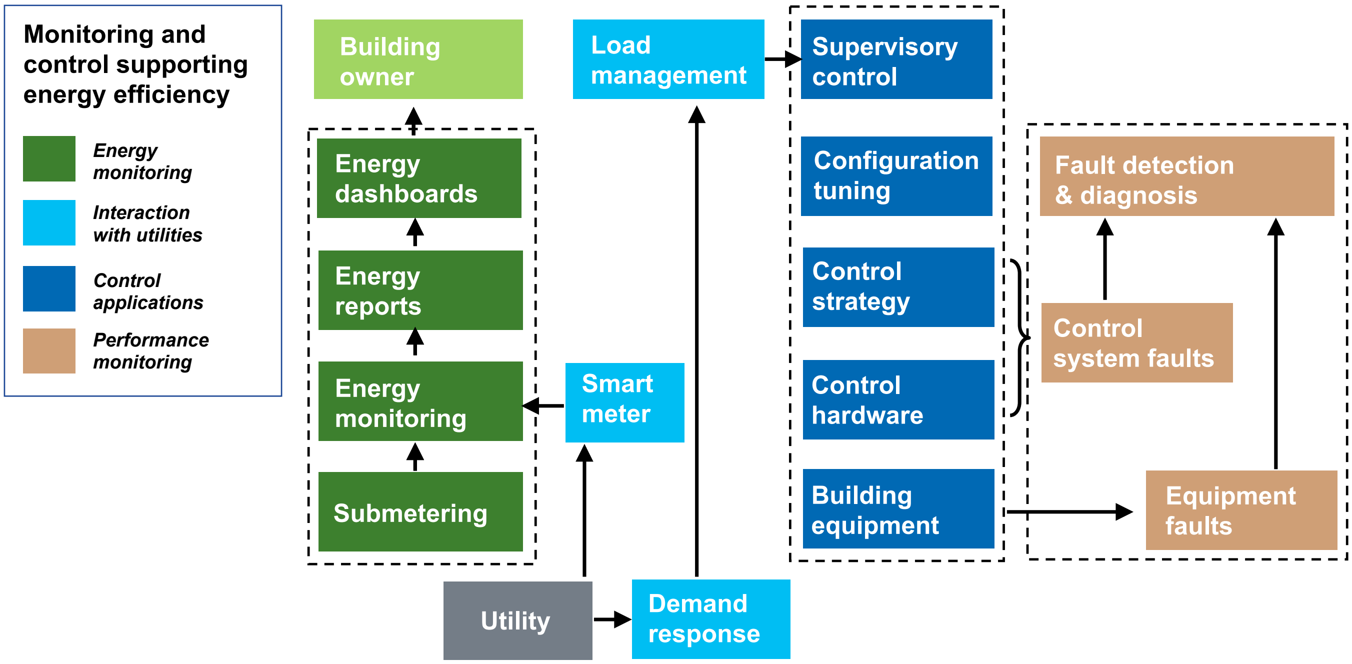
User experience. With social networking tools, occupants can provide instant feedback on their experienced comfort as well as receive explanations of system behaviors. In this way, occupants can be systematically engaged in energy management and building control. In the cloud environment, the social media data can be meshed with other real-time building data to create insights into the building's daily operation and implement improvements and cost-saving measures.



System view of HVAC

Contributor: Petr Stluka, Honeywell, Czech Republic

From: *The Impact of Control Technology,* 2nd ed., T. Samad and A.M. Annaswamy (eds.), 2014. Available at www.ieeecss.org.

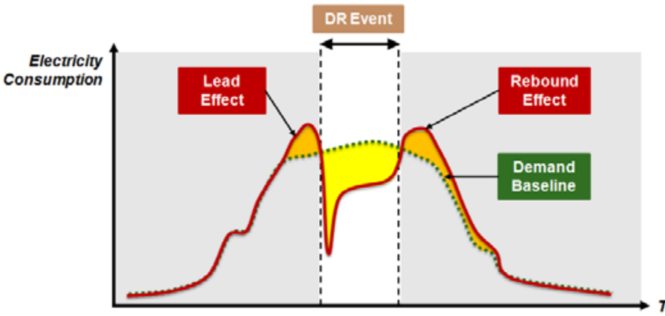


Challenges and Opportunities in Energy-Efficient Buildings

Multivariable HVAC supervisory control. The primary goal of HVAC control is to maintain occupants' thermal comfort and system energy efficiency. This requires adjustments of multiple setpoints—primarily temperatures and flow rates**.** Today these setpoints are either kept constant or manipulated by simple reset rules. An obvious opportunity exists for new robust multivariable supervisory control strategies that will leverage principles of model predictive control (MPC) to dynamically adapt key HVAC setpoints based on weather conditions, occupancy, and actual thermal comfort in zones. The challenge of developing reliable HVAC models for MPC might be addressed by moving the optimization engine to a cloud and coupling it with efficient analytics for identification of suitable models from the HVAC data.

Whole-building optimization. Economic optimization of building energy systems can be formulated to integrate all subsystems, including HVAC, lighting, onsite generation, and storage. The implementation of this approach is complicated by disturbances such as weather conditions and occupant behaviors and potentially also by dynamic pricing of electricity. However, the fundamental issue lies with buildingwide optimization models, which will always be hampered by significant inaccuracy, uncertainty, and lack of measurements. Distributed optimization approaches could be more viable; these would first divide the building into meaningful subsystems and then optimize each subsystem locally but not independently of others.

Building-to-grid integration. Recently, demand response (DR) has been recognized as a promising approach for the electricity market and an essential element of smart grid implementations. By sending changing power-price signals to building automation systems, adjustments of temperature setpoints, cycling of HVAC systems, or other actions can be initiated, and consequently energy use and expenditure can be reduced. A fundamental challenge is to enable the building to participate in demand response without violating thermal comfort. Advanced control strategies are needed that will manage building loads and use the building's thermal mass to implement various preheating strategies and adapt zone temperature trajectories. In addition to dynamic load management, in many cases, the scope of optimization could also encompass local generation and storage devices.



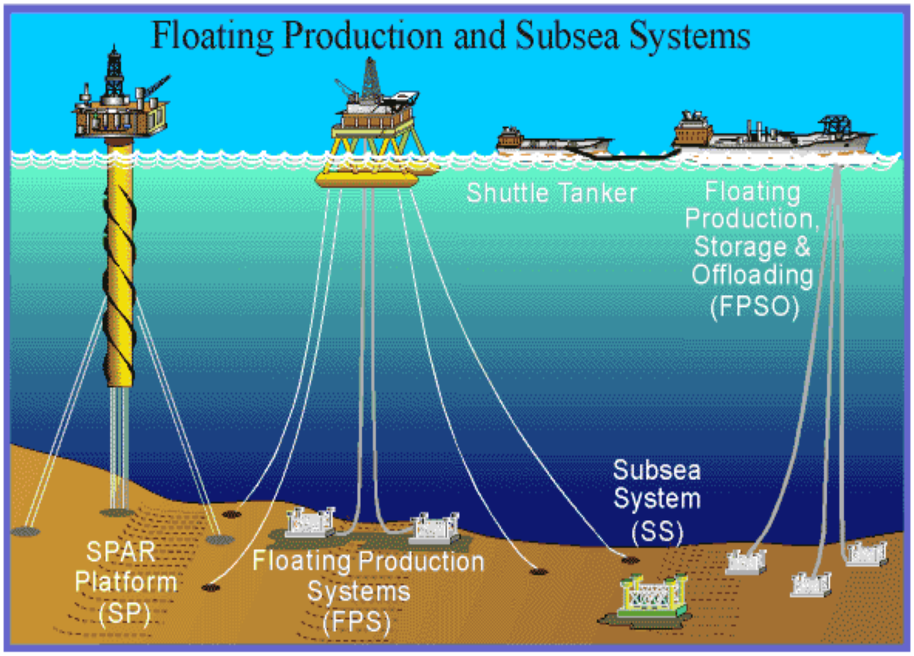
Building-to-grid integration: A DR event, which can result in significant reduction in electricity consumption (the yellow shaded region), can also result in increased consumption (orange) before and after the event. For optimizing demand response, models must be developed that incorporate the lead and rebound effects.

For more information: United Nations Environment Programme, Buildings and Climate Change: A Summary for Decision Makers, 2009, www.unep.org.

Control for Floating Structures in Offshore Engineering

The recent formation and growth of the global deepwater offshore industry has been driven by increased demand for oil and gas stemming from years of economic growth, reduced production in existing hydrocarbon fields, and depletion of shallow water reserves. These factors have encouraged operators to invest billions annually to chase this offshore frontier and the development of floating production and subsea systems as solutions for deepwater hydrocarbon extraction.

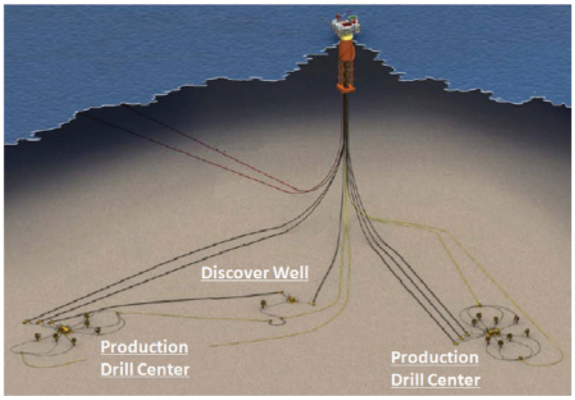
Currently, 15% of total offshore oil production is carried out in deep waters, and this proportion is expected to rise to 20% in the next few years. The harsher marine environment and need for subsea production systems in remote deepwater developments opens a set of challenges and opportunities for the control theorist and engineer.



Source: Minerals Management Service, U.S. Department of the Interior

A Critical Need for Technology

The April 2010 Deepwater Horizon accident in the Gulf of Mexico serves as a reminder of the risks and challenges in offshore operations. In the push toward exploration and production in deeper waters and harsher environments, control theorists and engineers working with colleagues in different disciplines will be challenged to forge a path forward with innovative technological approaches to safely supply the energy the world needs.



A view of the commercial subsea system (wells, manifold, and umbilical) on the seabed (Source: MMS Ocean Science, Nov. 2005)

Subsea Production Systems

Subsea systems must be installed accurately in a specified spatial position and compass heading within tight rotational, vertical, and lateral limits. The tolerances for a typical subsea installation are within 2.5 m of design location and within 2.5 degrees of design heading for large templates and are more stringent for the installation of manifolds into the templates.

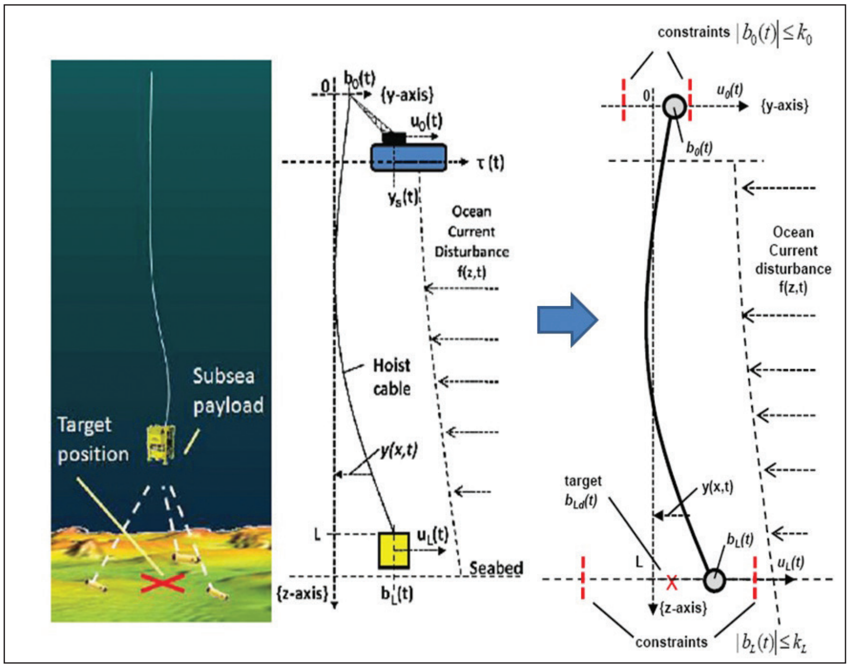
Traditional subsea installation methods include the use of guidelines or the use of ship dynamic positioning and crane manipulation to obtain the desired position and heading for the payload. Such methods become difficult in deeper waters due to the longer cable between the surface vessel and subsea hardware when near the seabed.

An intuitive solution to alleviate the precision placement problem is the addition of thrusters for localized positioning when the payload is near the target site. Control of the dynamic positioning of the subsea payload is challenging because of unpredictable disturbances such as fluctuating currents and transmission of motions from the surface vessel through the lift cable.

Contributors:* *Shuzhi Sam Ge, National University of Singapore, Singapore, and University of Electronic Science and Technology of China, China; Yoo Sang Choo and Bernard Voon Ee How, National University of Singapore, Singapore; Wei He, University of Electronic Science and Technology of China, China

Dynamics of the Lift Cable

With the trend toward installations in deeper waters, the longer cable increases the natural period of the cable and payload system, which in turn may lead to increased pendulum-like oscillations. Time-varying distributed currents may lead to large horizontal offsets between the surface ship and the target installation site. Investigation of the dynamics of the flexible lift cable to aid in the control design and operation planning is desirable and challenging.

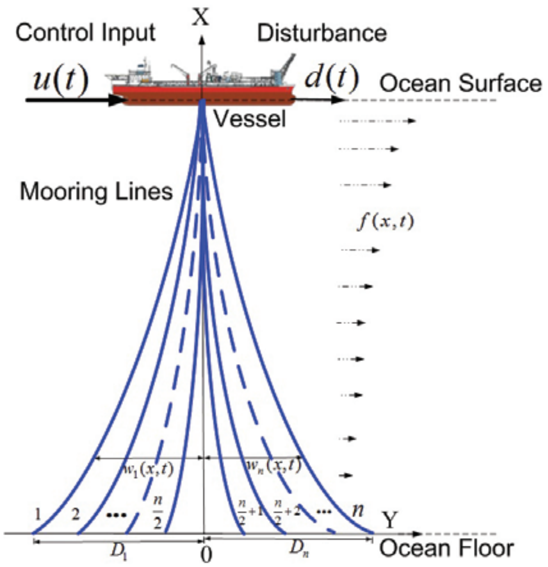


Positioning of the subsea hardware using thrusters (left), illustration of subsea positioning (center), and a schematic of the installation operation (right)

Thruster-Assisted Position Mooring Systems

Floating platforms, such as anchored floating production storage and offloading (FPSO) vessels with positioning systems, have been used widely. The two main types of positioning systems are dynamic positioning systems for free-floating vessels and thruster-assisted position mooring systems for anchored vessels. The thruster-assisted position mooring system is an economical solution for station keeping in deep water due to the long operational period in harsh environmental conditions. The thruster assistance is required in harsh environmental conditions to avoid the failure of mooring lines.

Mooring lines that span a great distance can produce large vibrations under relatively small disturbances, which will degrade the performance of the system and result in a larger offset from the target position of the vessel. Unknown time-varying ocean disturbances of the mooring lines lead to the appearance of oscillations, which make controlling the mooring system relatively difficult.



An FPSO vessel with thruster-assisted position mooring system

Control for Grid Responsiveness

Control room of a transmission system operator

Reliable electricity supply is largely taken for granted in the developed world. Very few electricity users think about the extensive infrastructure that is required to support ubiquitous availability of electrical energy. Even fewer are aware of the sophisticated analysis and control that underpins secure operation of these large-scale, highly distributed, nonlinear, hybrid dynamical systems.

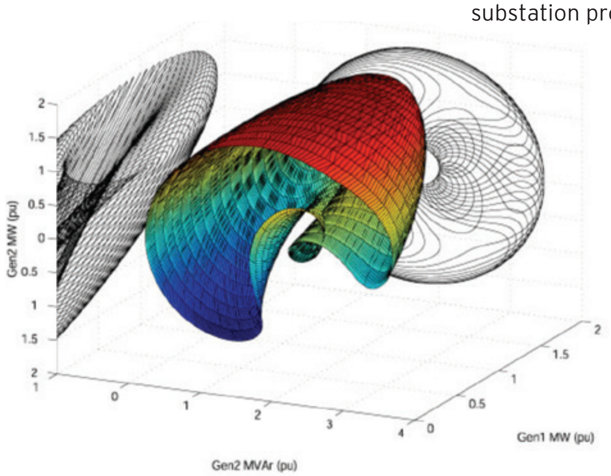
Power systems must evolve to incorporate new forms of generation and load. The variability inherent in large-scale renewable generation challenges existing regulation strategies. Plug-in electric vehicles, if adopted in large numbers, will introduce charging loads that must be carefully coordinated to avoid disruptive peaks in demand. Power transfers are continually increasing without a commensurate expansion of the transmission network, forcing system operation closer to limits. To meet these operational challenges, the grid must become more responsive.

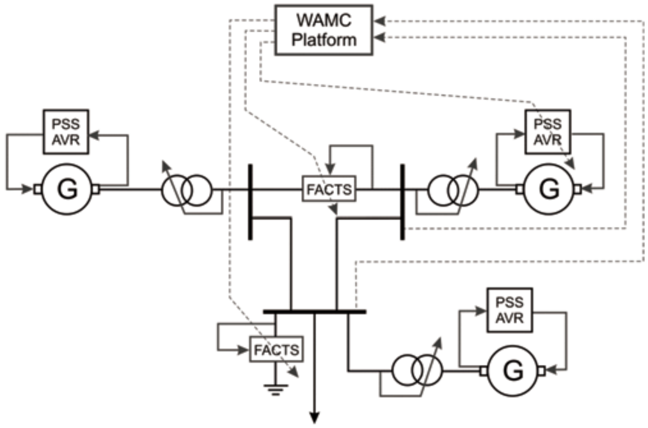
Enhanced grid responsiveness will rely on a range of available and emerging technologies. Phasor measurement units (PMUs) provide fast, accurate, time-stamped measurements that facilitate wide-area monitoring and control. Flexible AC transmission system (FACTS) devices use power electronics to control active and reactive power flows. Ubiquitous communication facilitates the participation of large numbers of loads in grid regulation. In all cases, control science and engineering will play a fundamental role in achieving stable, optimal operation.

Wide Area Monitoring and Control (WAMC)

Phasor measurement units provide geographically dispersed sensors that can supplement local measurements used by controllable devices, such as generators and FACTS installations. The wider view of system behavior offered by PMUs provides valuable information in determining optimal responses to systemwide events. Possibilities range from enhanced damping of inter-area oscillations to power flow modulation following large disturbances. To realize these benefits, however, controller designs must take into account signal latency and reliability.

PMU networks produce copious amounts of data. Sophisticated algorithms are required to extract information that is (1) valuable for alerting operators to system vulnerabilities, and (2) suited to closed-loop control applications. Security of communication networks is paramount, as PMUs are often tightly integrated into substation protection schemes.



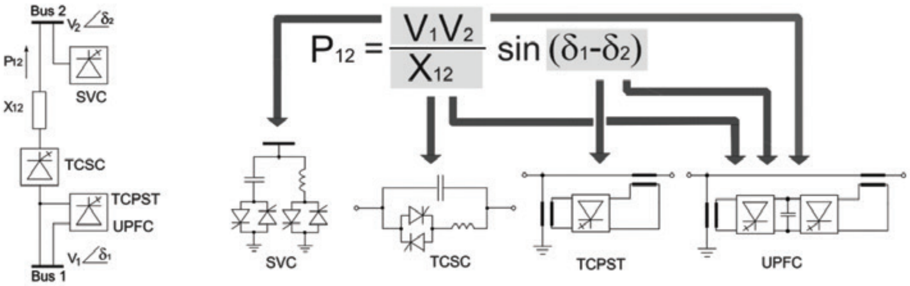


The colored figure shows the boundary of the power flow solution space for the example power system on the right. This manifold describes all combinations of the active and reactive power of one generator and the active power of a second generator. (The third generator is the "slack" generator, which balances the total supply with demand.) The black-and-white figures are projections of the colored figure onto axis pairs. The figure highlights the complexity that arises from the nonlinear nature of power systems, a complexity that cannot be avoided in real-world analysis and control applications. (Legend: AVR: automatic voltage regulation; PSS: power system stabilizer)

Contributor: Ian Hiskens, University of Michigan, USA

Flexible AC Transmission Systems (FACTS)

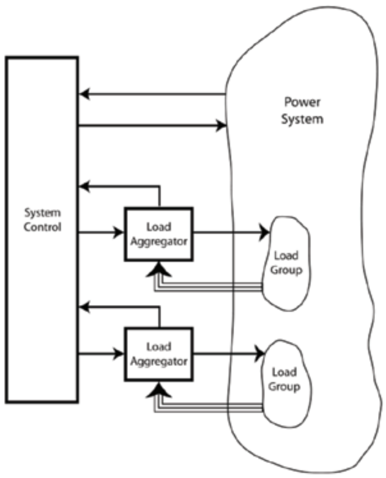
FACTS devices use the switching capability of power electronics to control voltages and currents in an AC grid. The most common FACTS devices are used to regulate bus voltages, for example, at the collector bus of a wind farm. FACTS devices are, however, also capable of controlling power flow over transmission lines. Without control, power will flow through an AC network in accordance with Kirchhoff's laws. This may overload some lines while leaving others underutilized. FACTS devices can redirect power to achieve more effective loading patterns.



Various FACTS devices

Examples of FACTS devices include static var compensators (SVCs—regulate voltage magnitudes), thyristor-controlled series capacitors (TCSCs—effectively regulate line impedances), thyristor-controlled phase shifting transformers (TCPSTs—regulate phase angle differences), and unified power flow controllers (UPFCs—regulate all of the above).

Optimal siting and sizing of FACTS schemes and their cost/benefit analyses involve nonconvex, nonlinear, mixed-integer optimization problems. Coordinated control of multiple FACTS devices must take into account the complexities inherent in regulation of a large geographically distributed nonlinear system.



A hierarchical control structure for integrating nondisruptive load control into power system operation

Load Control

Power system operation has traditionally relied on generation to ensure that supply and demand balance; however, because of the variability inherent in renewable energy production, that control philosophy will no longer be sufficient as renewable generation grows. It will become crucial for loads to participate in the regulation process.

To do so will require coordinated control of large numbers of autonomous devices. Centralized control seems impractical, with hierarchical and distributed control structures more likely to succeed. Using such control strategies, aggregate demand can track a specified signal, yet regulation of individual loads is achieved without any disruption to consumers. Such nondisruptive load control can be used, for example, to track the output of a wind farm, thereby minimizing its net variability. Numerous outstanding control questions remain to be addressed though.

Uncertainty in Power System Dynamics

Parameters associated with key power system models, in particular, loads and renewable generation, can never be known precisely. To ensure robust dynamic performance, controller designs must take into account plausible parameter ranges and system conditions. This is challenging due to the nonlinear, nonsmooth, large-scale nature of power systems.

Much work remains in the development of analytical results and numerical techniques that take uncertainty into account in the analysis and design of large-scale power systems.

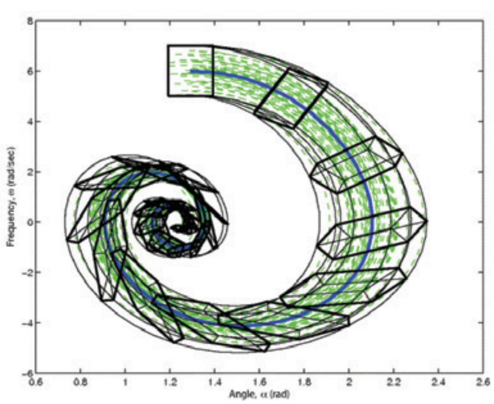


Illustration of phase angle (horizontal axis) and frequency (vertical axis) evolution in a power system showing nonlinear effects of parameter uncertainty. The complete uncertainty set generates a time-varying parallelotope that is mapped along with the nominal trajectory (the blue curve).

Control of Flapping-Wing Micro Air Vehicles

Drawings and manuscripts from antiquity reveal that humans have long dreamed of building machines that can fly using flapping wings. Although humans have been incredibly successful at building machines that fly, few of these aircraft fly using the modes most often observed in nature. Most natural flyers achieve flight through the use of flapping wings. One reason for the slow development of flapping-wing flight has been that the high aerodynamic efficiency of flapping wings that exists at the small scales found in nature disappears at the larger scales required for manned aircraft. Thus, practical applications for flapping-wing aircraft did not exist until the recent advancement of small-scale, unmanned air vehicle technology. The challenge of creating small, powered flapping-wing vehicles that perform some practical functions now appears to be within reach.

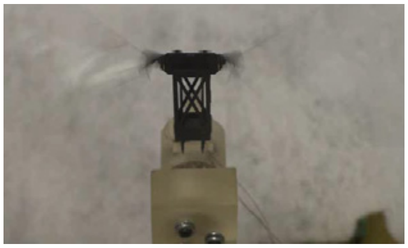
Numerous flapping-wing aircraft designers have been inspired by hummingbirds. (Photo by Bill Buchanan, U.S. Fish & Wildlife Service)

Beyond Conventional Flight Control

Most birds and many insects use periodic wing motion to propel themselves and maneuver. Most conventional flying machines are propelled by rotating machinery, achieve lift through rotating or fixed wings, and are controlled through the production of steady aerodynamic forces produced by rotors or movable wings. Many of the first powered flapping-wing micro air vehicles (MAVs) effectively replaced rotational propulsion modes with flapping wings and maintained control using conventional aerodynamic control surfaces or, in some cases, rotors. Recently, researchers have begun to develop aircraft that are controlled by manipulating the motion of the flapping wings themselves.

Control of a free-flying flapping-wing vehicle using only the flapping wings was achieved by the Aerovironment Hummingbird, a 19-gram aircraft that was powered by a DC motor and controlled by varying the angle of attack of each wing. Tiny piezoelectric flapping-wing aircraft in the 100-mg class have been produced by a research group at Harvard University; however, these aircraft have not yet achieved flight without being connected to an external power source. The interesting feature of control approaches that use only flapping wings is that the control forces and moments they produce are periodic rather than steady, as in a conventional aircraft. The periodic nature of the aerodynamics and the time scale separation between the vehicle flight dynamics and the wing oscillations allow the design of vehicles that can be controlled using a very small number of physical actuators.

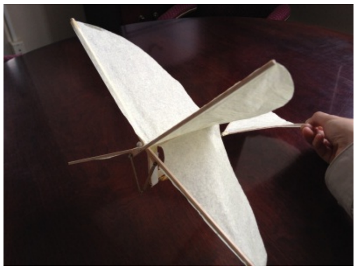
Conventional tail surfaces are used to control this flapping-wing MAV.



Two piezoelectric actuators enable independent control of the motions of the flapping wings on this test vehicle. (Photo by Maj. Michael Anderson, USAF)

Primitive Ornithopters

Ornithopters powered by rubber bands have been constructed since the 1800s. Although graceful and beautiful in flight, their practical value is limited. Many modern tail-controlled flapping-wing vehicles have borrowed elements from such designs. The four-wing ornithopter (right) propels itself by taking advantage of the interactions between the rigid leading edge and flexible trailing edges of the wings. The blowing action that occurs when the wings close and the suction produced when the wings open create an average horizontal component of force that propels the aircraft forward.



Balsa wood, tissue paper, wire, and rubber bands were used to create this ornithopter.

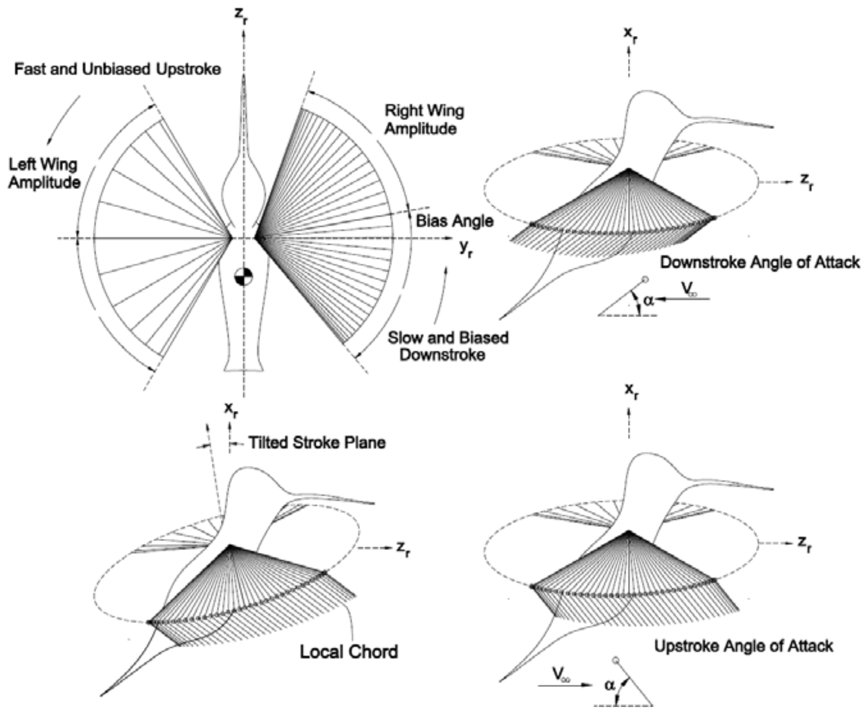
Contributors: David B. Doman and Michael W. Oppenheimer, U.S. Air Force Research Laboratory, USA

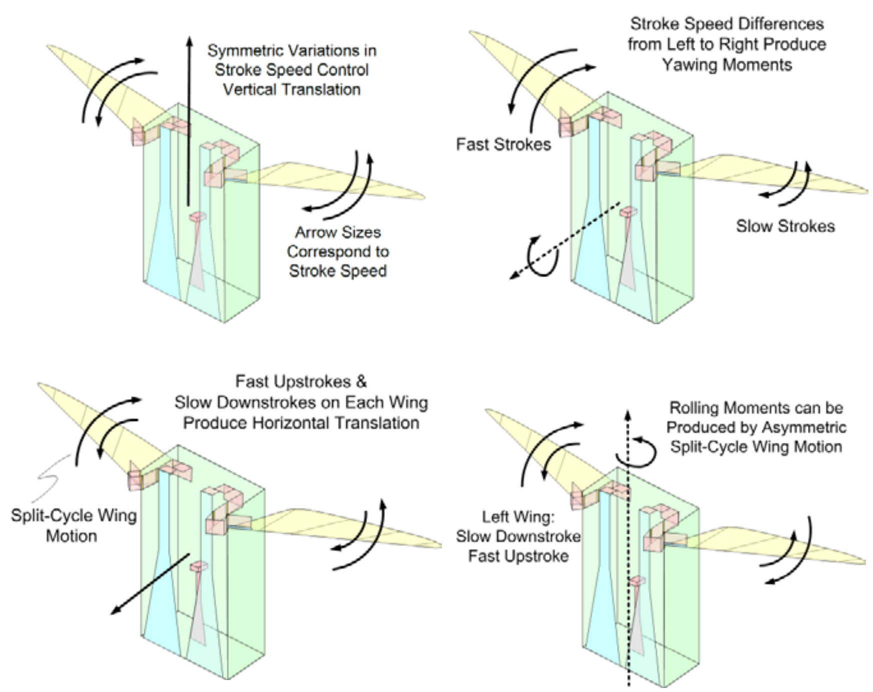
Exploiting Periodic Wing Motion for Control

Opportunities to explore control strategies for flapping-wing aircraft abound. Because the overall motion of the vehicle evolves on a much slower time scale than the motion of the individual wings, the vehicle motion is primarily influenced by time-averaged forces and moments. By manipulating a few variables that govern the periodic motion of two wings, the average forces and moments that are applied to the vehicle can be directly controlled. An exciting implication of this phenomenon is that the number of vehicle degrees of freedom that can be controlled can exceed the number of actuators that physically exist on the vehicle. Wing motion can be manipulated mechanically using numerous physical actuators or by using a single actuator whose motion is controlled by software. By shifting complexity from mechanical elements to software, the behavior of a small number of actuators can be governed by numerous virtual control variables that affect the time-averaged forces and moments applied to the vehicle. Some examples of wing motion parameters that may be used as virtual or physical control variables include wing stroke amplitude, wingbeat symmetric and asymmetric frequencies, wing stroke bias, angle of attack, and stroke-plane tilt angle. The strobed illustration to the right shows the effect of varying these parameters.

One method of controlling a flapping-wing aircraft is called split-cycle constant-period frequency modulation. This method works by using symmetric and asymmetric frequency as virtual control effectors, leaving the other previously mentioned variables fixed. The method allows the roll and yaw rotations and the horizontal and vertical translations to be directly controlled using two brushless DC motors or piezoelectric actuators that drive each wing independently. Differences in wingbeat period between the left and right wings produce a yawing moment. Collective changes in wingbeat period produce vertical accelerations. Differences between the upstroke and downstroke speeds that occur over each wingbeat period produce finite cycle-averaged drag forces that can be used for horizontal translation or the production of rolling moments.

Example Control Strategies

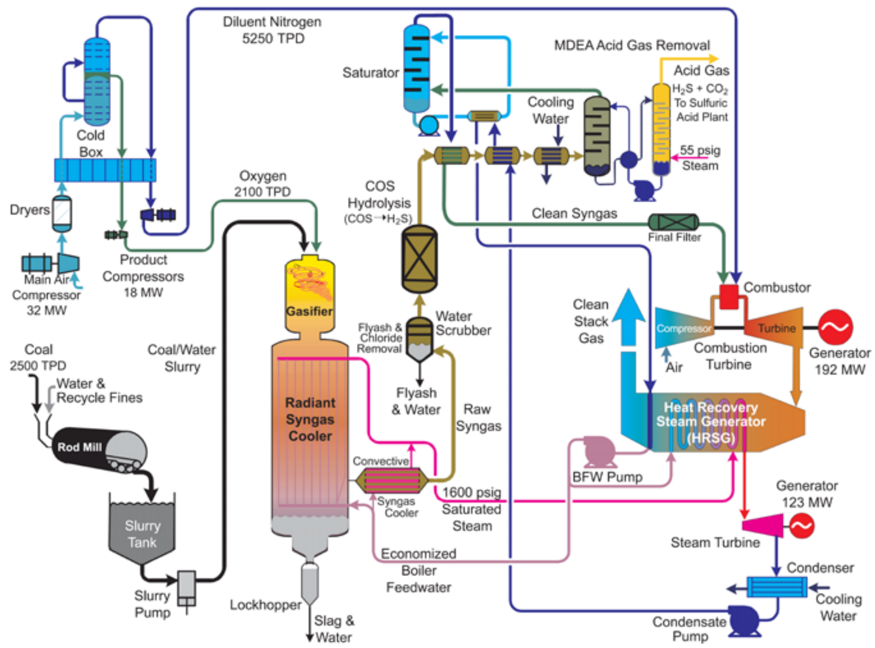




For more information: M.W. Oppenheimer, D.B. Doman, and D.O. Sigthorsson, Dynamics and control of a hovering biometric vehicle using biased wingbeat forcing functions, Journal of Guidance, Control, and Dynamics, 2011, pp. 204-217; D.B. Doman, M.W. Oppenheimer, and D.O. Sigthorsson, Dynamics and control of flapping wing micro air vehicles, Handbook of Unmanned Aerial Vehicles, Kimon P. Valavanis and George J. Vachtsvanos, eds., Springer, 2014 (to appear).

Control of Integrated Gasification Combined Cycle Power Plants with CO_2 Capture

Carbon dioxide (CO_2) emissions from today's coal-fired power-generation technology are a growing concern because of their implication in global climate change. Integrated gasification combined cycle (IGCC) power plants are an attractive prospect for clean coal power generation, with a few operational plants existing worldwide. The integration of CO_2 capture with IGCC is now being pursued, with the potential for significantly increased efficiency and lower cost of electricity than for CO_2-capture-integrated conventional pulverized coal plants. However, IGCC plants with CO_2 capture will require operation in a highly constrained and fluctuating environment. This complex environment requires the use of highly nonlinear dynamic models and poses several challenges in advanced control and sensors.



Schematic representation of an IGCC power plant

Pre-combustion versus Post-combustion CO_2 Capture

In conventional coal power plants, the fuel is pulverized and burned in a boiler. The "post-combustion" CO_2 by-product is emitted through the flue-gas stack. In IGCC plants, however, the coal is gasified and not fully combusted. The gasification process produces a synthesis gas containing "pre-combustion" CO_2 at much higher temperatures, pressures, and concentration levels, thereby facilitating the separation of the CO_2 from the synthesis gas. CO_2 capture requires the integration of a new, complex chemical engineering process. The more advanced capture systems include chemical solvents, whereas other capture processes at earlier stages of development employ novel methods such as solid sorbents or membranes.

CO_2 capture reduces the overall plant generation efficiency by 20–25% or more, but the efficiency loss is significantly lower for pre-combustion than for post-combustion technologies. In addition, the capital cost for an IGCC CO_2 capture unit is substantially lower than for the pulverized coal equivalent. Because of these factors, IGCC is often considered the preferred approach for clean coal power generation.

Current Status of IGCC Plants with CO_2 Capture in the U.S.

Several major IGCC power plants with full-scale pre-combustion CO_2 capture are moving forward in the U.S., including Mississippi Power's Kemper County lignite-fired 582-MW IGCC with 65% CO_2 capture, Summit Power's coal-fired 400-MW IGCC project with 90% CO_2 capture, and SCS Energy's petcoke-fired 421-MW IGCC with hydrogen production and 90% CO_2 capture. In these applications, the captured CO_2 will be used for enhanced oil recovery (EOR) from production wells. These large-scale projects will demonstrate the integration, operational, and control aspects of IGCC technology when coupled with CO_2 capture.

Mississippi Power's Kemper IGCC project with CO_2 capture (Source: www.biggerpieforum.org)

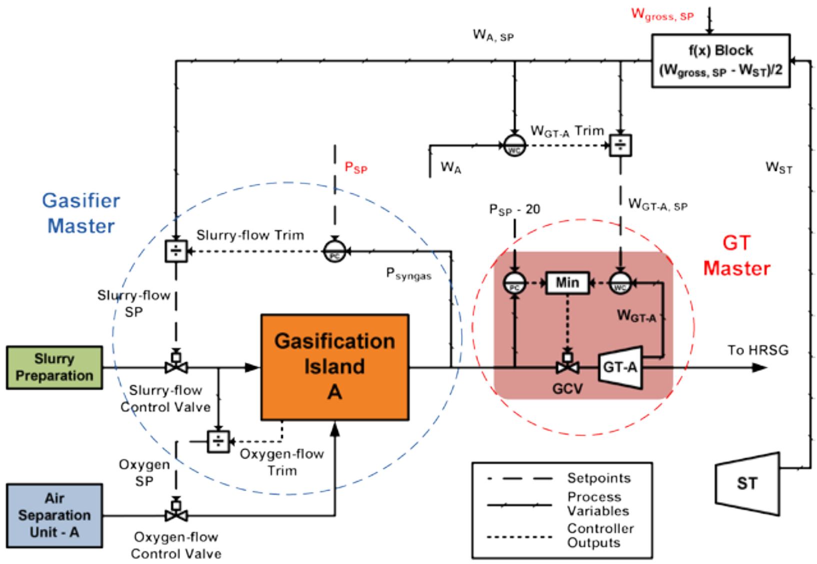
Contributors: Fernando V. Lima, Debangsu Bhattacharyya, and Richard Turton, West Virginia University, USA, and Priyadarshi Mahapatra and Stephen E. Zitney, National Energy Technology Laboratory, USA

Advanced Control and Power Plant Cycling

Significant penetration of intermittent and variable renewable energy sources (e.g., wind and solar) into the grid is likely to require IGCC plants to vary their generation output in concert with renewable generation and load variations. Novel modeling and control strategies for power plant cycling and load following, while maximizing economic criteria and minimizing emissions, are needed. Cycling operations can be performed by manipulating the throughput of the gasifier and combined cycle islands in tandem, as shown in the control architecture figure at right.

Other advanced control-related challenges include:

- Fast and accurate reduced models for use in model-based control of highly nonlinear and stiff processes, such as gasification
- Solving large-scale equation systems in the multiple-software industrial automation environment
- Developing equipment damage models to assess the impact of IGCC cycling operations and control



Improved control architecture (GT leader/gasifier follower) for IGCC load following. Nomenclature: W – work, P – pressure, SP – controller setpoint, GT – gas turbine, ST – steam turbine, GCV – gas control valve (inlet guide valve), HRSG – heat recovery steam generator.

Sensors and Estimation

Because of ever-tightening environmental emission limits, the accurate estimation of pollutants (e.g., CO, CO_2, H_2S, COS, NH_3) emitted by power plants is becoming crucial. Monitoring the compositions of key process streams is also important for plant efficiency and safety. However, available composition sensors, especially for trace species, are costly, maintenance-intensive, insufficiently accurate, and do not provide real-time estimates. Development of real-time or near-real-time sensors for state estimation is required to improve advanced process monitoring and control of such species. Optimal sensor placement strategies are also needed for monitoring, disturbance rejection, and fault diagnosis.

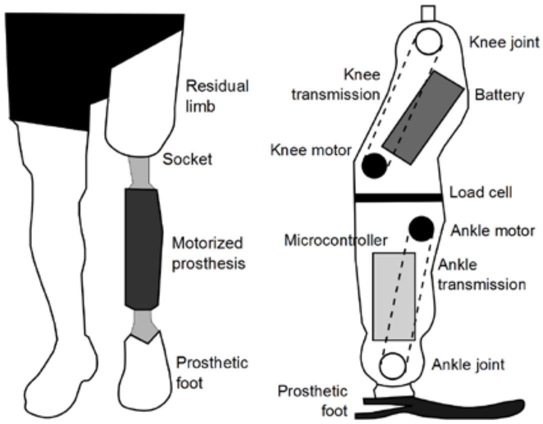
Platform for Testing IGCC Control and Sensor Strategies

The AVESTAR® Center at the National Energy Technology Laboratory (NETL) and West Virginia University (WVU) provides a real-time IGCC dynamic simulator for developing and testing advanced control and sensor placement strategies. The figure on the left shows a distillation column in the virtual plant environment for which a control system was designed. The high-fidelity IGCC model can simulate plant performance over a range of operating scenarios, including variable-load operation, startup, shutdown, and variable CO_2 capture rates. The IGCC simulator is being used to develop novel model predictive control (MPC) strategies to improve ramp rates and load-following operation while satisfying CO_2 emission constraints. Distributed and hierarchical MPC of large-scale networked systems with embedded sensors and controllers is another area of active research.

For more information: S.E. Zitney et al., AVESTAR Center: Dynamic simulation-based collaboration toward achieving operational excellence for IGCC plants with carbon capture, Proc. of the 29th Annual International Pittsburgh Coal Conference, Pittsburgh, PA, October 15-18, 2012; http://www.netl.doe.gov; http://www.che.cemr.wvu.edu.

Control of Powered Prosthetic Legs

Amputee locomotion is slower, less stable, and requires more metabolic energy than able-bodied locomotion. Lower-limb amputees fall more frequently than able-bodied individuals and often struggle to navigate inclines such as ramps, hills, and especially stairs. These challenges can be attributed largely to the use of *mechanically passive* prosthetic legs, which do not contribute positively to the energetics of gait, as do the muscles of the biological leg. Powered (or robotic) prosthetic legs could significantly improve mobility and quality of life for lower-limb amputees (including nearly one million Americans), but control challenges limit the performance and clinical feasibility of today's devices.



Drawing of above-knee amputee wearing a powered/motorized prosthetic leg (left), and enlarged schematic diagram of the experimental Vanderbilt *prosthetic leg (right)*

State of the Art in Lower-Limb Prosthetic Control Systems

With the addition of sensors and motors, prosthetic legs must continuously make control decisions throughout the gait cycle, thus increasing the complexity of these devices. This complexity is currently handled by discretizing the gait cycle into multiple periods, each having its own separate control model. Each control model may enforce desired stiffness and viscosity characteristics or track predefined patterns of angles, velocities, or torques at the joints. To switch between control models at appropriate times, the prosthetic leg uses sensor measurements to estimate the phase—or location in the gait cycle.

Limitations of the State of the Art

This approach to prosthetic leg control poses two key problems: (1) reliability of the phase estimate for switching control models, and (2) difficulty of tuning control parameters for several control models to each patient and task. An error in the phase estimate can cause the prosthesis to enact the wrong control model at the wrong time (e.g., swing period during stance), potentially causing the patient to fall. Even if the phase estimate is correct, each control model must be carefully tuned by a team of clinicians or researchers to work correctly for a particular patient performing a particular task. Some prosthetic control systems have five discrete periods of gait with more than a dozen control parameters per joint per period. Multiple tasks (e.g., walking, standing, stair climbing) add up to hundreds of parameters for a multi-joint prosthetic leg, presenting a critical challenge to the clinical viability of these high-tech devices.

How Can Control Theory Contribute?

- Nonlinear filtering techniques for accurately estimating the phase of gait
- Optimization methods to *automatically* determine patient-specific control parameters for multiple periods of gait and different tasks
- Nonlinear control methods to unify the entire gait cycle under one control law
- Simultaneous stabilization theory to operate across multiple periods, tasks, and patients
- Formal verification methods to certify safe operating conditions

Application to Orthotics/Exoskeletons

Control theory can also be applied to powered orthoses or exoskeletons that replicate lost leg function after spinal cord injury or assist locomotion after stroke. This application will present new challenges in designing control systems around human limbs, modeling limb mass/inertia, and compensating for muscle spasticity (i.e., abnormal resistance to muscle stretch due to a neurological impairment).

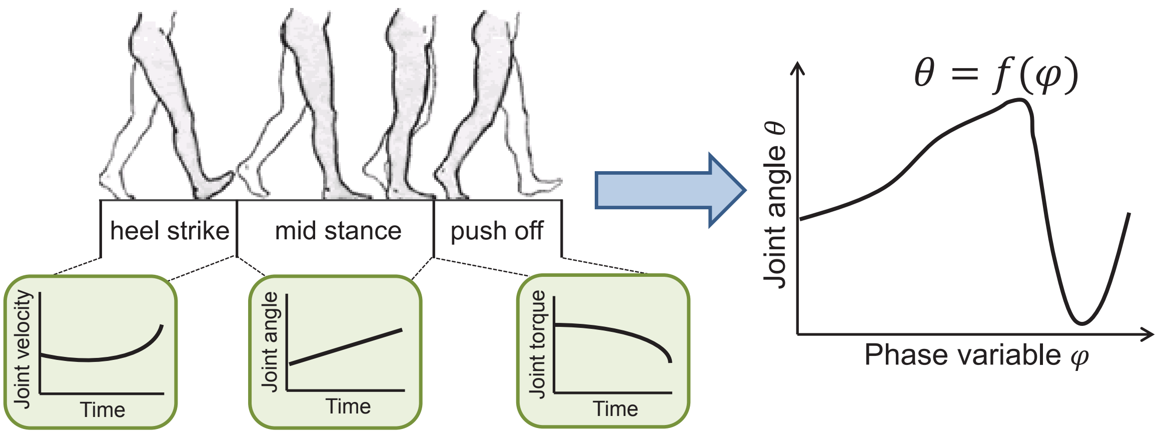
Contributors: Robert D. Gregg, University of Texas at Dallas, USA, and Levi J. Hargrove and Jonathon W. Sensinger, Rehabilitation Institute of Chicago and Northwestern University, USA

From Walking Robots to Prosthetic Legs: Phase-Based Virtual Constraint Control

Some of these challenges could be addressed by parametrizing a nonlinear control model with a mechanical representation of the gait cycle phase, which could be continuously measured by a prosthesis to match the human body's progression through the cycle. Feedback controllers for autonomous walking robots have been developed that "virtually" enforce kinematic constraints, which define desired joint patterns as functions of a mechanical phase variable (e.g., hip position). These phase-based patterns, known as *virtual constraints*, have recently enabled bipedal robots to walk, run, and climb stairs, presenting an emerging opportunity to address a key roadblock in prosthetic technology.

Recent Results

Researchers are currently investigating the use of biologically inspired virtual constraints to make prosthetic legs more robust and easily tuned than with controllers used to date. In particular, researchers at the Rehabilitation Institute of Chicago programmed the powered *Vanderbilt leg* to control its knee and ankle patterns based on the heel-to-toe movement of the center of pressure—the point on the foot sole where the cumulative reaction force is imparted against the ground. A recent study successfully tested this prosthetic control system on three above-knee amputee subjects.



Conceptual diagram of shift from sequential control to virtual constraint control, where joint patterns are characterized by continuous functions of a mechanical phase variable

Remaining Challenges With Virtual Constraints

- Measuring a phase variable from the prosthesis during the swing period
- Modeling virtual constraints for different tasks such as stair climbing
- Feedback linearization to precisely enforce virtual constraints for high-performance tasks such as running
- Neural interfaces for subconsciously commanding the leg to perform different tasks
- Clinical trials with virtual constraints in a take-home prosthetic leg

An above-knee amputee testing a virtual constraint controller on the Vanderbilt leg at the Rehabilitation Institute of Chicago

For more information: R.D. Gregg and J.W. Sensinger, Towards biomimetic virtual constraint control of a powered prosthetic leg, IEEE Transactions on Control Systems Technology, 2013; R.D. Gregg et al., Experimental effective shape control of a powered transfemoral prosthesis, IEEE International Conference on Rehabilitation Robotics, Seattle, WA, 2013.

Control of Tokamak Plasmas

Controlled thermonuclear fusion has the potential to address the global need for sustainable energy. Energy can be generated by the fusion of deuterium and tritium (isotopes of hydrogen extracted from water and the Earth's crust) in a harmless way (no direct radioactive waste is generated and the radioactivity of the structure decays rapidly). Fusion devices using magnetic confinement of the plasma, such as tokamaks, can thus be envisaged as a major carbon-free energy resource for the future. Significant research challenges still need to be addressed, however, before tokamaks can be reliably operated and commercially viable. The controls community will play a crucial role in resolving these challenges.

Tokamaks

Tokamaks use a magnetic field to confine a plasma in the shape of a torus. The charged particles follow a helicoidal trajectory according to the field created by controlled magnets, which thus set the position and shape of the plasma. Radio-frequency antennas allow selective action on electrons or ions and modify internal plasma properties such as current and temperature. The plasma is fueled by pellets shot at high speed toward the plasma center, and neutral particles are injected to increase the plasma momentum and energy.

The ITER Tokamak, an international project involving seven members (the European Union, Russia, the U.S., Japan, China, Korea, and India), plans to start operation during the next decade. It is expected to produce 500 MW from 50 MW of input power.

Tokamak control is becoming increasingly important to the success of magnetic fusion research and will be crucial for ITER. Feedback control of the main plasma macroscopic parameters, such as plasma position and shape, total current, and density, is now reasonably well mastered in the various worldwide tokamaks. However, control of internal plasma dynamics and radial profiles (1-D distributions) is still in its infancy. This control is likely to be crucial for robust stability and to maintain high-efficiency tokamak operation.

Challenges in Plasma Physics for ITER

New control methods will be required for meeting five important objectives for tokamak operation.

C1 – Magnetohydrodynamics (MHD) stabilization: Nonaxisymmetric electric currents cause perturbed magnetic fields inside (e.g., magnetic islands) or outside (e.g., resistive wall modes) of the plasma, as well as central plasma relaxations (e.g., sawteeth). MHD instabilities evolve at a fast time scale (~ 10^{-6} – 10^{-3} sec) and need to be addressed in both the poloidal and toroidal directions.

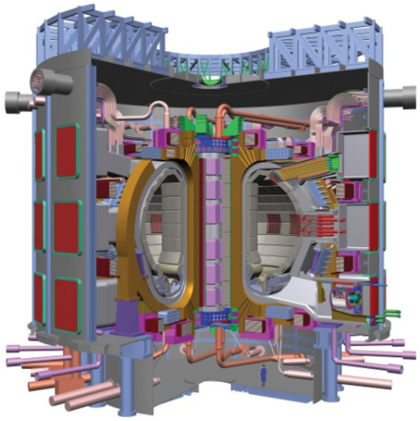
C2 – Heat confinement: The plasma core must be at very high temperatures (up to 150 x 10^6 K, 10 times the central temperature of the sun) while having an edge temperature that can be sustained by the plasma-facing components. Maintaining large temperature gradients is thus essential to achieving an efficient "burn" control while preserving the plasma shell.

C3 – Steady-state operation: "Steady-state" dynamics evolve more slowly than the current density diffusion time (~1 – 100 sec) and relate to the ability to continue the tokamak operation indefinitely; that is, the plasma pulse is terminated by the operator's choice. The so-called "safety-factor" q (calculated as the ratio between the toroidal and the poloidal magnetic flux gradients) and the pressure profiles provide indicators on the potential avoidance of MHD instabilities.

C4 – Control of plasma purity: An impurity flux is driven by different transport phenomena (e.g., ash transport, gas puffing at the plasma boundary, and impurity removal) as well as plasma-wall interactions. This problem is related to both design aspects (e.g., optimal diverter and plasma-facing components) and real-time feedback.

C5 – Plasma self-heating with α-particles: The α-particles (He^{2+}) produced by the fusion reaction are charged and are trapped by the magnetic field, transferring their energy to the plasma. They thus provide an extra heat source and induce a local nonlinear feedback. Anisotropic transport analysis and burn control must be combined to control this phenomenon.

Although each of the above challenges is mostly considered as an independent control problem, the automation system will ultimately have to deal with the strong couplings that exist between the various plasma dynamics and the multiple roles of each actuator. For example, the q-profile is a key parameter for both the global stability of plasma discharges and an enhanced confinement of the plasma energy (C1 and C3). Other examples include the couplings between the temperature and the safety-factor dynamics (C2 and C3), and the multiple effects of using an RF antenna at electron cyclotron frequency (C1, C3, and C5).



The ITER Tokamak (www.iter.org)

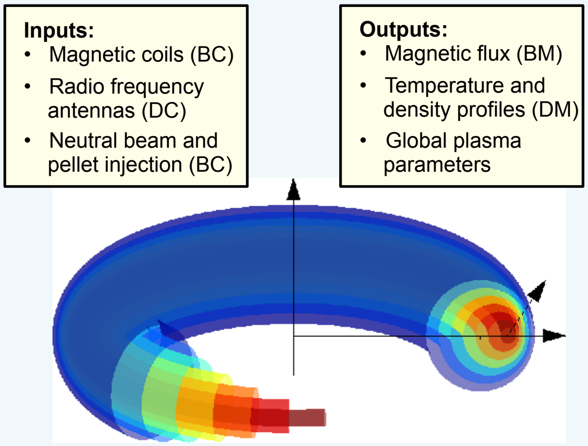
Contributor: Emmanuel Witrant, University Joseph Fourier, France

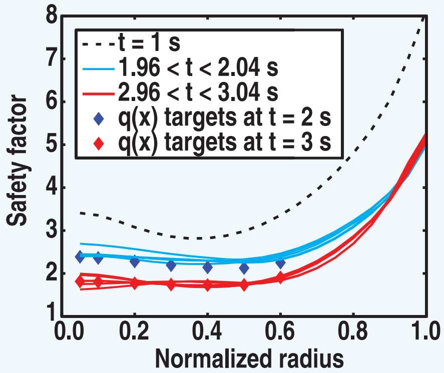
Tokamak Automation

The plasma in a tokamak is actuated with magnetic coils of various sizes and positions, with antennas emitting at different radio frequencies, and with the injection of pellets and neutral particles. In terms of sensing capabilities, distributed measurements of the temperature and density are available, along with the boundary values of the magnetic flux. Real-time equilibrium reconstruction codes provide virtual sensors for the magnetic flux inside the plasma. The resulting integration and sensor fusion of complex diagnostics, combined with the assignment of the actuators to specific regulation tasks, pose control-engineering challenges.

Example: Safety-Factor Profile Control

For "steady-state" operation, the state-space variables can be averaged on surfaces of identical magnetic flux (identified with different colors on the left figure below), and the radial profile is regulated in the 1-D space. Both boundary and distributed controls (BC and DC) and measurements (BM and DM) are available. The first experimental results (right figure below) of integrated control of the safety-factor and pressure profiles were published in May 2013 (Moreau et al., *Nuclear Fusion* 53, 2013).





Experimental results of controller capability to track changes in the target safety-factor profile q(x) (at t = 2 and 3 sec), from an arbitrary value (at t = 1 sec)

Control Challenges for Distributed-Parameter Systems

The control issues associated with tokamaks involve the spatiotemporal dynamics of transport phenomena (magnetic flux, heat, densities, etc.) in the anisotropic plasma medium. The physical models typically involve inhomogeneous partial differential equations (PDEs, mostly of parabolic or hyperbolic type) with transport coefficients that differ by several orders of magnitude, depending on their location, and involve nonlinear couplings between the physical variables. New results are thus sought on the following topics.

Identification and estimation, possibly with unknown inputs, of time- and space-varying transport parameters. The wide range of tokamak instrumentation provides an exceptionally rich database for evaluating new estimation strategies in the PDE framework.

Stabilization with computation constraints for high-order linear systems with multiple time-varying delays. Such models can be used to describe convective transport and MHD instabilities based on modal analysis.

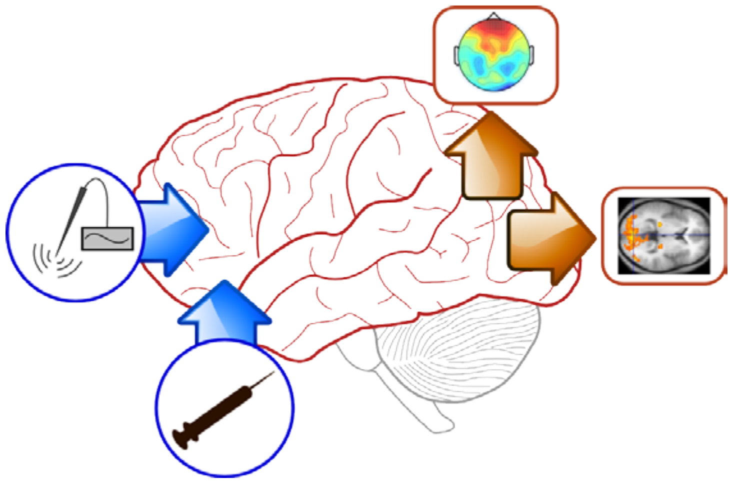
Robust PDE control for the regulation of 1-D transport equations (e.g., safety factor, temperature, and density), which results in a profile control in the radial direction.

Optimal reference design to provide scenarios that integrate the multiple coupling constraints for the feedback strategies. These scenarios can be computed offline and use advanced physical models that include the complexity of plasma interconnected dynamics.

For more information: Join the IEEE Control Systems Society Technical Committee on Power Generation; visit http://www.ieeecss.org/technical-activities/power-generation.

Control-Theoretic Approaches in Neuroscience and Brain Medicine

Emerging technologies in neuroscience and brain medicine are providing ever more sophisticated ways of both measuring and manipulating neuronal circuits *in vivo*. As evidenced by therapies such as deep brain stimulation, such technologies hold immense medical promise. The full realization of this promise will require the recognition of these technologies—and conventional treatments involving pharmacologic agents—as inputs and outputs of a dynamical system. Appropriately and tractably modeling brain dynamics will open new avenues for the exploration of control-theoretic approaches to elucidating brain function and optimizing new therapeutic strategies.



Emerging technologies for brain stimulation, coupled with pharmacologic manipulation, are providing an ever-increasing ability to "actuate" brain dynamics. Paired with modalities for measuring these dynamics, new opportunities in neuronal control theory will emerge.

Current Paradigms in Pharmacologic and Stimulation-Based Neural Control

Closed-Loop Control of Anesthesia: General anesthesia is the most prevalent "actuator" of human brain dynamics in current clinical practice. Despite efforts to develop closed-loop anesthetic delivery solutions, several challenges remain related to ensuring robustness in the presence of significant patient variability and unreliable feedback signal quality. By developing a new class of nonlinear biophysical models that more accurately reflect drug action, new and improved control designs will be possible.

Deep Brain Stimulation (DBS) for Treatment of Parkinson's Disease: In DBS, specialized electrodes are implanted in certain brain structures and electric current is delivered to disrupt pathological neural activity. At present, this stimulation consists of periodic pulse trains whose frequency is set empirically. Biophysical modeling and principled application of control theory will enable the design and optimization of closed-loop controllers for delivering this therapy in the next generation of DBS technology.

Applications in Basic Neuroscience: Optogenetics

Optogenetics is a powerful and popular technique in experimental neuroscience that uses genetic manipulation to make specific classes of neurons sensitive to light. These neurons can then be activated, that is, caused to spike, through illumination (e.g., via fiber optics) *in vivo*. The power of this technique lies in the ability to probe how certain populations of neurons are involved in behavior and function. Currently, this technique is used in open loop, and illumination takes the form of constant pulses. Large populations of neurons are turned on or off en masse. Control theory is certain to play an important role as neuroscientists seek to induce more sophisticated patterns of activation and to manipulate neural activity in real time.

How Can Control Theory Contribute?

Many control theoretic notions—dynamics, stability, feedback, identification—are of direct importance to neuroscience and brain medicine. Particular areas of interest include:

- Robust methods for designing feedback controllers in the presence of patient uncertainty
- Dimensionality reduction for developing tractable biophysical models of neuronal activity in response to exogenous stimulation
- System identification of neuronal dynamics at multiple scales
- Constrained design of persistent excitation for estimating neuronal circuits
- Analysis and characterization of dynamics in neuronal networks toward uncovering basic brain function
- Optimization of therapeutic strategies using multiple types of drugs or stimulation modalities

Contributor: ShiNung Ching, Washington University in St. Louis, USA

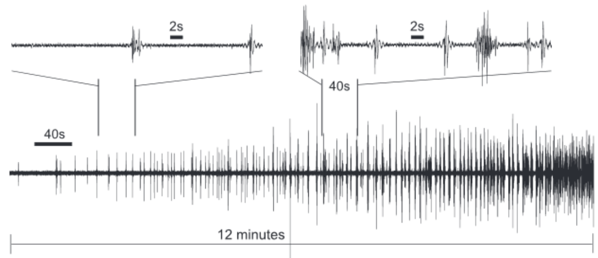
Toward New Therapies in Brain Medicine: The Dynamics of Neurological Disease

An emerging trend in clinical neurophysiology is to associate certain brain dynamics (e.g., oscillations, synchrony) with neuropathology, disease, or clinical state. Modeling these dynamics will open the door to new interventions motivated by control theory.

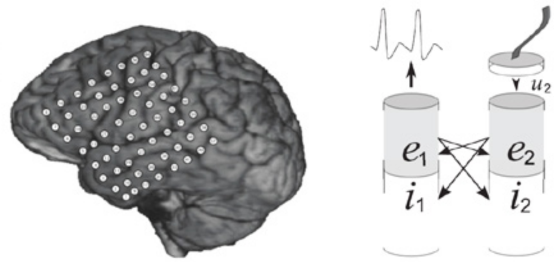
Control of Medically Induced Coma via Closed-Loop Anesthetic Delivery: A medically induced coma (a very deep level of general anesthesia) is characterized by bistable brain activity (see right). Recent nonlinear models have shed light on the underlying dynamics, opening the possibility of new nonlinear closed-loop designs.

Control of Seizure Activity Using Surface Stimulation: In an epileptic seizure, a focal region of pathological activity kindles an event that spreads through the brain in a network cascade. Grids of electrodes could be used in a distributed control scheme to induce stability (see right).

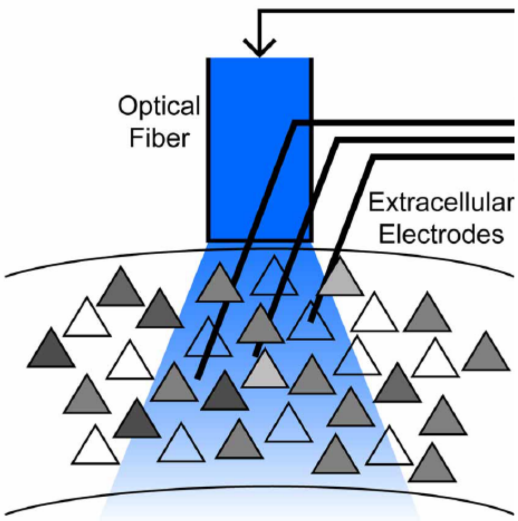
Control of Neuronal Spiking Patterns: Rather than simply activating/deactivating entire brain regions, an emerging challenge is that of true neuronal "control," in which specific patterns of brain activity are induced, potentially even at the scale of individual neurons (see below).



Twelve minutes of EEG activity showing emergence from a medically induced coma. Accounting for the nonlinear dynamics (note the bistability) is paramount in developing closed-loop methods to regulate such brain states.



Grids of stimulating electrodes could be used to control pathological activity. The spatial undersampling introduces challenges in distributed and decentralized control.



Optogenetics schematic. A single light source may affect hundreds of neurons. What is the appropriate notion of controllability in this setting?

The Challenge of Underactuation

Whether by pharmacology or stimulation, neuronal control is fundamentally underactuated: only a few inputs are used to affect a highly complicated system consisting of a vast number of interacting units (neurons, brain regions). The classical notions of controllability and observability are likely to be impractical and, arguably, not meaningful in this setting. Instead, new theoretical questions emerge:

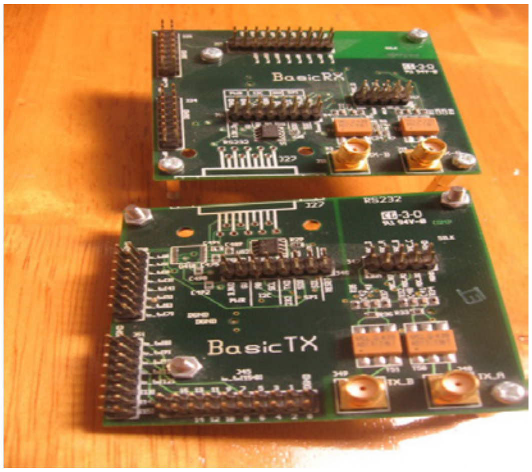
- Rather than specifying complete state trajectories, what patterns (e.g., the order in which neurons are activated) can be achieved?
- How can desired activity be induced while satisfying strict constraints (e.g., stimulation power, drug quantity)?
- Can the intrinsic sparsity in neural activity be exploited for purposes of control?

For more information: S. Ching et al., Real-time closed-loop control in a rodent model of medically induced coma, Anesthesiology, 2013; S. Ching et al., Distributed control in a mesoscale cortical network model: Implications for seizure suppression, Physical Review E, 2013; S. Ching and J.T. Ritt, Control strategies for underactuated neural ensembles driven by optogenetic stimulation, Frontiers in Neural Circuits, 2013.

Controlling Modern Radars

To the layman, radars evoke the image of large rotating parabolic antennas bolted to a stationary structure and managed by human operators. Although such radars still exist, modern radars can be rather small, and their antennas may be electronically steered. Modern radars are typically mobile, either flying on aircraft, moving on the ground, or located aboard ships. Their outputs are exploited by algorithms as complex as those used to produce the outputs themselves. Below the surface, the changes are equally impressive, mainly brought about by advances in antenna technology and signal processing. Despite earlier isolated efforts, radar control has only attracted the interest of researchers in the past 20 years, under the headings of sensor management and sensor scheduling. The wide adoption of active electronically steered array (AESA) antennas starting in the 1970s and the conception of software-defined radio (SDR) in the 1990s bring a combination of agility and flexibility to radar systems operation and design that is only now being fully exploited.

The face of the COBRA DANE AESA in Alaska; the antenna contains 34,000 modules, each measuring 5 inches in diameter.



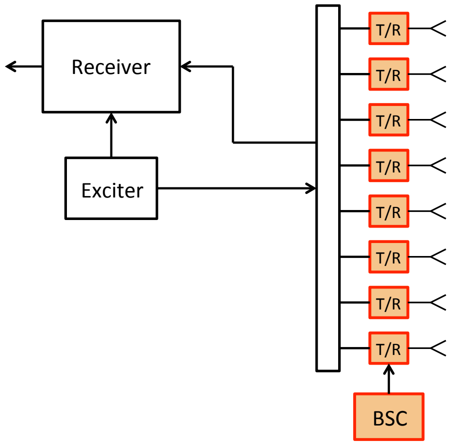
Basic transmitter and receiver daughterboards for the Universal Software Radio Peripheral (USRP)

Enabling Technology I: Software-Defined Radio

Software-defined radio is an emerging technology in which several radio frequency (RF) functions, such as mixing, filtering, and modulation/demodulation, are performed in software instead of hardware. An example of SDR is the GNU Radio project, a free and open source development toolkit providing signal processing blocks to implement SDRs.

Enabling Technology II: AESA Antennas

Active electronically steered array antennas are obtained by combining numerous miniature transmit/receive (T/R) modules placed behind each radiating element, as shown at right. The beam-steering controller (BSC) controls the phases and amplitudes of the radio waves transmitted and received by each radiating element. AESA antennas are far more agile than their mechanically steered counterparts, allowing the formation of almost arbitrary space-time configurations for the antenna beam. The ability to rapidly switch beams permits multiple radar functions to be performed, interleaved in time.



Source: George Stimson, "Introduction to Airborne Radar," SciTech Publishing, Inc., 1998; reprinted with permission.

Contributor: João B.D. Cabrera, BAE Systems, USA

Modern Radars and Echolocation in Bats

Although echo feedback is only recently being used for waveform adaptation in radars, nature has it is own version of the process in the form of the echolocation systems of various species of bats.

Echolocation signals in bats are very brief sounds, varying in duration from 0.3 to 300 ms and in frequency from 12 to 200 kHz. In most species, the sounds consist of either frequency-modulated (FM) components alone or a combination of a constant-frequency component with FM components. Most species studied to date show changes in their sonar signal parameters (for example, duration, bandwidth, and repetition rate) with varying foraging conditions, such as their proximity to vegetation, water, and buildings. Sonar signal design is widely believed to reflect the bat's control over acoustic information gathered from the echoes.

Interestingly, bats may or may not be capable of multipulse, coherent processing as performed by pulse-Doppler radars. However, bats (and dolphins) are believed to produce synthetic aperture sonar images without coherent processing.

Research Challenges at the Crossroads of Signal Processing, Estimation, and Resource Management

Radar waveforms (the transmitted signals) are selected on the basis of the desired application. Fundamental uncertainty principles apply; as an example, one cannot have arbitrarily precise estimates of range and radial velocity at the same time. The faint echo reflection is processed in the receiver using advanced signal processing techniques—phase array processing—developed over the course of several decades. In most cases, however, the waveform remains unchanged for large periods of time (i.e., the transmitter/receiver system operates in open loop).

The advent of SDR systems enabled programmability of the waveform generator and opened the possibility of closing the feedback loop in modern radars. Some of the theoretical underpinnings for echo feedback were established in the early 1990s, but much work still needs to be done on the theoretical front:

- Fundamental performance bounds of the Cramér-Rao type need to be established for radar systems in the presence of echo feedback;
- Coding, coherent processing, and waveform adaptation need to be brought together within a unified framework;
- Mechanisms for real-time, pulse-to-pulse interleaving of adapted waveforms need to be developed. Multiple-input, multiple-output (MIMO) radars and radar networks make the problem even more formidable as the multiple transmitter/receiver pairs need to operate in coordination.

As the ultimate challenge, consider a network of unmanned aerial vehicles (UAVs), each carrying a radar capable of waveform adaptation. To complete a surveillance mission successfully, the UAV ensemble needs to coordinate its radar operation. At the same time, the UAVs exchange information about the environment, providing yet another input to the control loop. The radar controllers must optimally combine their native echo reflections with the inputs from the other UAVs to produce their transmission schedules.

For more information: S. Haykin, Cognitive radar—a way of the future, IEEE Signal Processing Magazine, January 2006; F. Gini and M. Rangaswamy (eds.), Knowledge Based Radar Detection, Tracking and Classification, John Wiley & Sons, Inc., 2008.

Design Science for Cyberphysical Systems

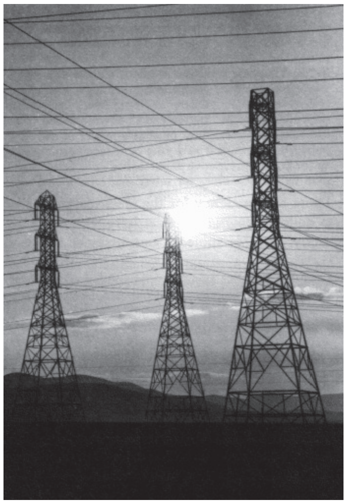
The integration of physical systems through computing and networking has become the most pervasive application of information technology, a trend now known as cyberphysical systems (CPSs). In many science and technology domains—transportation, healthcare, energy, and manufacturing automation—this intersection has yielded disruptive technologies, created new industries, and rearranged the status quo in entire economic sectors. The deep integration of computational, physical, and human aspects of engineered systems has spurred the development of cross-disciplinary approaches that are yielding new methods in science and engineering design. The tight integration of physical and information processes in CPSs necessitates the development of a new systems science, one that ensures efficiency, resiliency, robustness, safety, scalability, and security.

Why Is Systems and Control Relevant?

Control science and engineering have achieved remarkable success in the analysis and design of physical systems. New theories and tools are now needed to design networked heterogeneous systems, along with their automation and controls, that take into account interactions among the physical, communication, and computation systems.

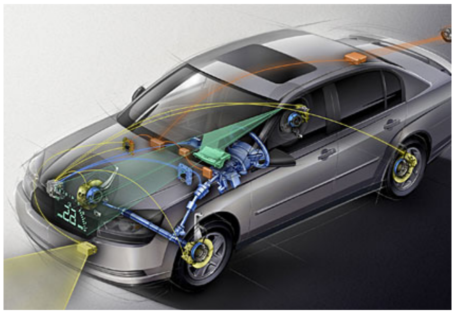
CPSs represent an exciting opportunity for the controls community to extend its legacy of rigorous theoretical foundations and engineering methods and to thereby play a crucial role in ensuring the performance, reliability, safety, and feasibility of complex engineered systems.

Designing Complex CPSs



Among the hardest problems facing researchers and engineers are those associated with producing robust integrations of physical systems and computational processes for cyberphysical systems. Cyber components are steadily increasing and are essential for delivering advanced capabilities in airplanes, cars, spacecraft, power grids, medical devices, physical infrastructures, and other applications. CPSs are ubiquitous: well over 90% of all microprocessors are now used for cyberphysical systems—not in stand-alone computers or enterprise systems. But theoretical foundations and automated tools for designing tightly integrated computational and physical systems—encompassing modeling, control, verification, optimization, and the like—are lacking.

Contributors: Xenofon Koutsoukos, Vanderbilt University, USA, and Panos Antsaklis, University of Notre Dame, USA

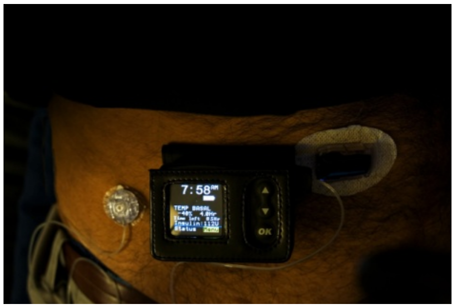


Automotive Systems

Cars and trucks are as much cyber as physical entities! A modern automobile can have more than 100 processors and more than 10 million lines of code. Multiple vehicle functions are under software control, such as engine, transmission, aftertreatment, suspension, and driver controls. Semiautonomous (and fully autonomous) cars are being developed and will pose further challenges for automation.

SCADA

Supervisory control and data acquisition (SCADA) systems perform vital functions in national critical infrastructures, such as electric power transmission and distribution, oil and gas pipelines, water and wastewater pipelines, and transportation systems. SCADA functions include monitoring and control of large-scale complex physical processes that may be distributed among multiple sites over long distances.



Biomedical Devices

Human lives are directly at stake in the operation of biomedical devices, especially implanted ones. Increasingly, such devices are incorporating embedded sensing, control, and actuation. Examples include insulin pumps, pacemakers, and implantable cardiac defibrillators.

CPS Design: Cross-Cutting Research Needs

Cyberphysicality spans the gamut of engineering domains. Although CPSs also pose application-specific requirements, control scientists and engineers can have an impact across industry and application sectors by addressing several broad research challenges:

- *Science of CPS integration.* CPS architectures need to bring together complex control algorithms, communication protocols, and computational platforms in integrated networked systems that provide real-time, high-confidence assurance of performance and robustness.
- *Dynamic configuration and scalability.* In many applications, CPSs will have plug-and-play components, including physical and cyber elements. CPS control must function seamlessly under reconfiguration and as the systems scale.
- *Co-design of system/platform/control.* Historical practices of developing the physical system, then the information technology platform, and then the control approach and algorithm are no longer tenable—an integrated, synthetic design paradigm is needed.
- *Human-automation interaction.* Human operators and users are an inherent part of many CPSs and are an additional source of complexity in control system design. In several respects, semiautonomous systems are harder to design than fully autonomous ones!
- *Verification and validation.* Most CPSs operate in safety- or mission-critical environments; costs of failures are typically intolerable. Automation and control design processes must provide assurance of safe, reliable operation under normal and abnormal conditions.
- *Cybersecurity.* From medical devices to the power grid, CPSs are cyber-connected systems. Security against and resilience under cyber attacks are an increasingly critical need.

Discrete-Event Control Theory for Flexible Manufacturing

Flexibility in Manufacturing

Flexibility in the manufacturing context refers to the ability of a manufacturing system to react to and accommodate changes. Changes may occur due to various causes: machine breakdown, system reconfiguration, and market demands. Correspondingly, flexibility is needed at different levels. For example, a machine may perform a variety of operations and a manufacturing cell may produce a range of products. Flexible manufacturing has been widely recognized as essential for reducing production costs, enhancing operational reliability, and adapting to variable needs in today's markets.

A flexible manufacturing system (Image courtesy of Cincinnati Milacron, Ohio)

Flexible Manufacturing Systems

In the past few decades, flexible manufacturing systems (FMSs) have attracted significant attention in industry as well as research communities. An FMS typically consists of a number of automatic machine tools linked by material handling/transportation elements (robotic manipulators, automated guided vehicles) and storage (buffers, inventories). Besides hardware, FMSs have sophisticated computer-control software that executes monitoring, scheduling, routing, and supervision tasks.

Flexible Manufacturing Systems: Challenges

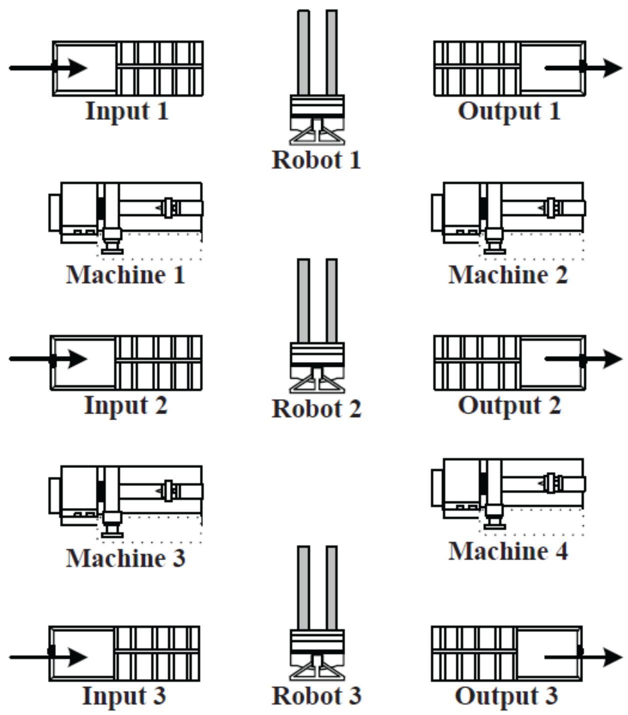
Despite the considerable advances in flexible manufacturing, today's state-of-the-art FMSs have several shortcomings:

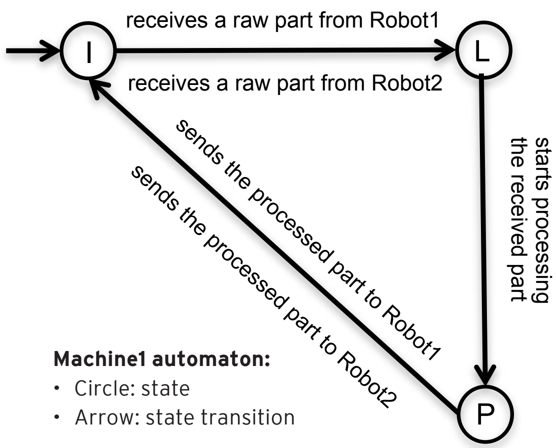
- **Trial-and-error design and verification:** In current system and control design, behavioral requirements of FMSs are specified by experts/managers, and software is coded by software engineers. This frequently leads to coding-testing trial-and-error loops, making the design process inefficient, control software prone to error, and FMSs prohibitively difficult to expand in functionality and scale.
- **Limited scalability:** Many FMSs have a central controller that requires extensive communications for even moderate-scale systems. Such a control and communication architecture makes it difficult to increase system size and production capacity to meet rapid changes in market demand.
- **Poor fault tolerance:** Component failure in an FMS can cause cascading malfunctions; replacement of a component may require interruption of system operation. Efficient reconfiguration strategies are lacking that can allow unaffected parts of FMSs to continue to operate while recovering from a failure.

Why Discrete-Event Systems?

The dynamics of FMSs exhibit the typical characteristics of discrete-event systems (DESs). DESs are discrete in time and (usually) in state space; they are event-driven, with events occurring asynchronously; and they may be nondeterministic (i.e., capable of transitional choices by internal chance or other mechanisms not necessarily modeled by the system analyst). The theoretical machinery developed for DESs can help resolve outstanding issues with FMSs.

Contributor: Kai Cai, Osaka City University, Japan





Left: An FMS consisting of machine tools, routing robots, and input and output conveyors. This FMS can produce three different products: the first uses machines 1 and 2, the second uses machines 3 and 4, and the third uses all four machines. (Image courtesy of W. Chao et al., Int. J. of Control, 86(1), 9-21, 2013)

Above: Automaton model of machine 1. The model has three states—"I" (idle), "L" (loaded), and "P" (processing)—with state transitions triggered by distinct events.

By modeling FMSs as automata, as shown here, formal supervisory control methods can be applied to address challenging control design problems in flexible manufacturing.

Why Supervisory Control Theory?

Supervisory control theory for DESs was proposed by P.J. Ramadge and W.M. Wonham in the 1980s. The theory employs finite automata for modeling and regular languages as analytic tools and has developed a compelling design methodology suitable for addressing the challenges faced by FMSs.

- **Formal models and correct-by-construction synthesis:** With uncontrolled FMS behavior and control requirements both modeled as finite automata, the theory automatically synthesizes supervisors that ensure safe, deadlock-free, and maximally permissive behavior. The synthesized supervisors are also finite automata, which can be translated into standard FMS control software such as for programmable logic controllers.
- **Modular control architecture and supervisor localization:** For large-scale DESs, the theory explores horizontal modularity (decentralized control), vertical modularity (hierarchical control), or a combination of both. Built on these modular control architectures, supervisor localization has recently been developed that decomposes modular supervisors into local control logics for individual distributed components. These techniques are useful for improving FMS modularity and scalability.
- **Automated diagnosis and reconfiguration:** Supervisory control theory provides diagnosis methods for efficient detection and isolation of failure events. Automatic reconfiguration approaches have also emerged that enable a DES to continue operation in reduced-performance modes. These methods will enhance the reconfigurability of manufacturing systems in response to failures.

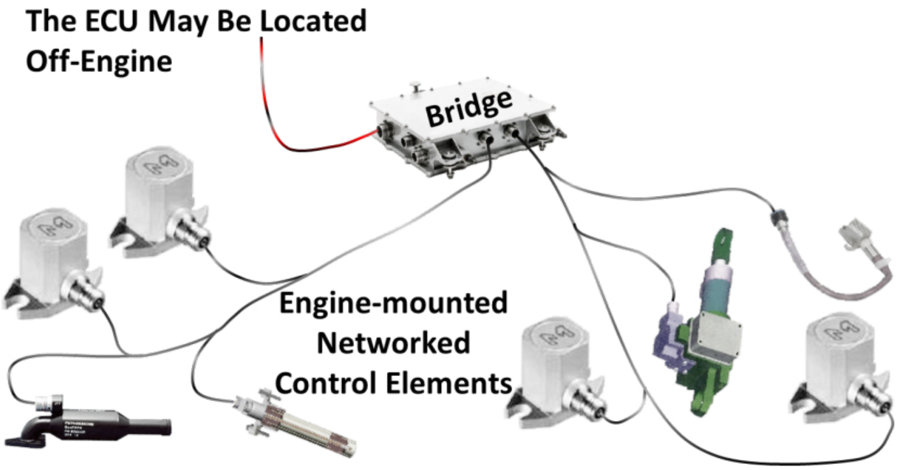
Impact on Industry

We anticipate that research in the application of discrete-event supervisory control theory to flexible manufacturing systems will help companies at the forefront of this technology achieve numerous significant benefits that are critical for competitiveness in global markets:

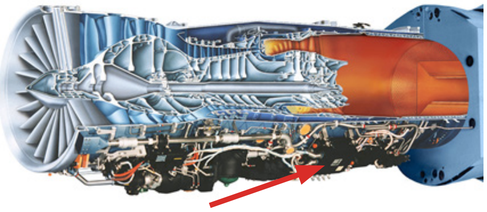
- Improved flexibility of manufacturing with assurance of performance and safety
- Reduced delay and waste in manufacturing processes as a result of optimized controller designs
- Enhanced operational reliability against component and subsystem failure
- Improved scalability of production with modular designs
- Automatic and rapid reconfiguration of manufacturing lines in response to market changes

Distributed Control for Turbine Propulsion

Development of a modular, distributed control architecture for turbine propulsion systems is necessary for achieving improved performance and capability in aviation. These high-performance engines and their control systems are expected to perform at peak efficiency for decades while subjected to some of the most severe environmental conditions, implementation constraints, and the constant threat of electronics obsolescence.



Notional high-temperature, networked control elements remaining on the engine in a future distributed control architecture



Currently, the ECU for military aircraft engines is core mounted, requiring fuel cooling. As the core becomes smaller and hotter, both the size and cooling requirements for the ECU become problematic.

The present ECU for commercial aircraft engines is mounted on the cold fan casing. As the fan diameter increases, the push is to reduce the gap between the casing and the nacelle for weight and drag reduction. The size of the ECU becomes an issue.

From Centralized to Distributed Control

Today's modern full authority digital engine control (FADEC) is based on technology developed in the early 1980s. This architecture features a digital electronic engine control unit (ECU) that also accommodates analog input/output circuitry to operate the control system's sensors and actuators. The present ECU is typically engine-mounted in close proximity to these control elements to minimize the overall impact of weight on the engine, of which the harness accounts for a large percentage. Consequently, the ECU is exposed to the most severe environmental extremes that must be accommodated in its design and then certified to the most stringent safety and reliability requirements.

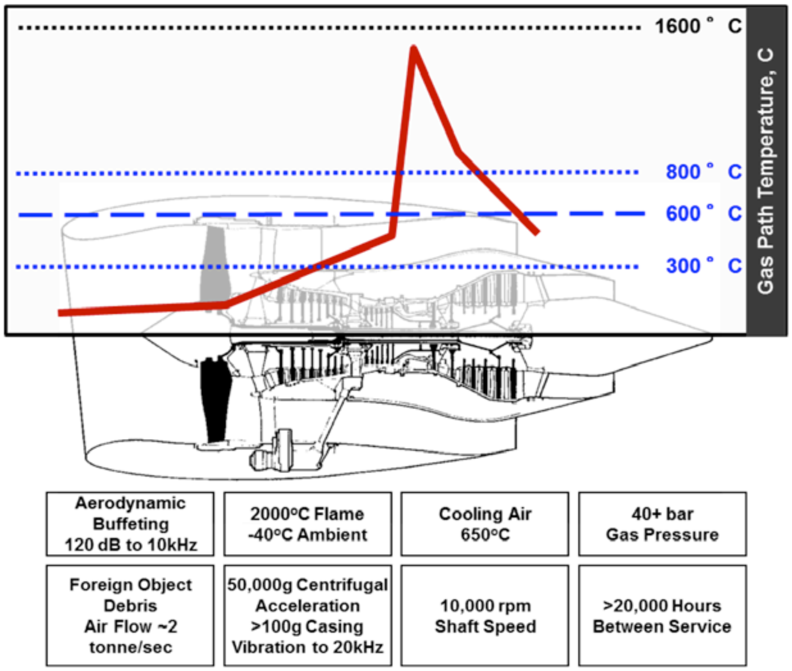
Unfortunately, the limitations of this architecture are becoming increasingly apparent. The improvements in turbine technology have increased engine operating temperatures and reduced the opportunities to locate control hardware. Likewise, the certification process inhibits the introduction of new capabilities while obsolescence drives the need for unplanned modifications.

Migrating to a distributed control architecture will allow the replacement of complex analog interfaces with digitization at the source, embedded processing, and common network interfaces at the system effectors and the ECU. The modularization of control functionality affords new opportunities to improve turbine capability by increasing the capability of existing control functions or, more simply, expanding the number of elements.

Contributors: Dennis Culley and Sanjay Garg, NASA Glenn Research Center, USA

Control System Constraints

- Harness weight can be several hundred pounds on large engines. "Wiring" the harness through the engine structure with multiple connectors adds weight and complexity, as well as raising reliability and fault-tolerance issues.
- The size of the ECU and the cooling requirements pose a challenge as turbine engine technology moves to smaller and hotter cores and larger fans for efficient operation.
- ECUs use commercial electronics, which often become obsolete in a few years, whereas ECUs have a lifetime of 10+ years. This requires complex and costly ECU parts management.
- Unplanned redesign and upgrades are prohibitively costly due to certification issues.
- Emerging active component control concepts for more efficient engine operation cannot be accommodated in the current control architecture.



Temperature extremes and vibration in the aero-engine environment are severe. The minimum temperature is due to ambient conditions at altitude, whereas the maximum temperature occurs in the combustor. Keeping the engine running under most fault conditions is essential.

Challenges Facing Distributed Control for Turbine Propulsion

A fundamental challenge facing implementation of distributed control on turbine engines is the lack of suitable electronics operating at high temperatures. Aircraft engine temperatures typically exceed the ratings of commercially available electronics used in automotive applications. This is especially true at soakback conditions when airflow through the engine is eliminated. Existing commercial electronics capable of operating above 200°C, typically silicon-on-insulator, are prohibitively expensive and have an indeterminate life when operating for extended periods at those temperatures. In response, engine manufacturers and their suppliers have formed a collaboration through the Distributed Engine Control Working Group (DECWG™) to address precompetitive technologies such as high-temperature electronics.

For the long term, significant effort is also being expended to increase the high-temperature capability of electronics for power as well as small-signal analog and digital applications. The progress in silicon carbide has been most notable, especially in the area of power electronics where commercial devices are available.

Additional challenges for distributed control architectures, several of which are posed by the high-temperature electronics issue, include the following:

- High-bit-rate networking in a high-temperature environment
- Robust and deterministic communication protocol
- Engine safety and stability
- Control system integration
- System modeling and analysis
- Verification and validation
- System certification and partial certification
- Supply chain viability
- Engine-airframe integration

For more information: J.D. Cressler and H.A. Mantooth (eds.), Extreme Environment Electronics, CRC Press, 2012; J. DeCastro et al., Analysis of decentralization and fault-tolerance concepts for distributed engine control, AIAA 2009-4884, 45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, 2009; D. Culley, Transition in gas turbine control system architecture: Modular, distributed, and embedded, ASME Turbo Expo 2010: Power for Land, Sea, and Air, 2010.

Dynamics and Control for the Artificial Pancreas

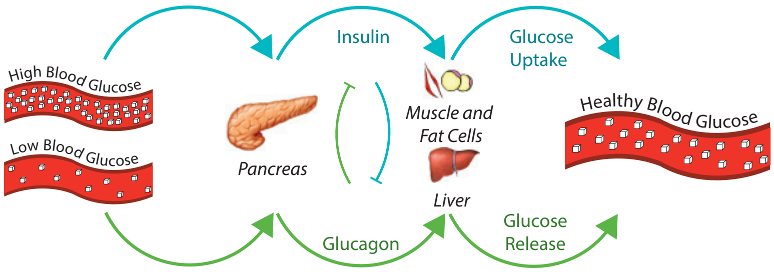
Healthy Regulation of Blood Glucose

The human body uses a combination of opposing manipulated variables (dual control) to achieve regulation of blood glucose using hormones produced by the pancreas, similar to a driver using the brake and gas pedals in an automobile. Insulin functions as the "brake pedal," lowering the blood glucose by stimulating the uptake of glucose by muscle, fat, and kidney cells. The counterregulatory hormone glucagon acts primarily to break down glycogen in the liver, yielding glucose and an elevation in blood glucose levels (acting as the "gas pedal").

Relevance of Systems and Control

The systems and control community is playing a crucial role in developing strategies for the reliable automation of blood glucose monitoring and regulation in several ways:

- Advanced control design (for example, model predictive control)
- Design of "verifiable" algorithms for regulatory approval processes
- Safety and fault analysis for medical delivery systems
- Algorithms to monitor the patient and the health of the system
- Advanced glucose calibration algorithms

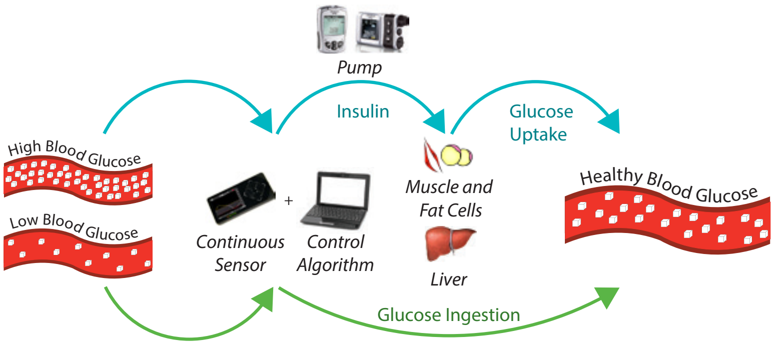


Type 1 Diabetes

- Type 1 diabetes is an autoimmune disease leading to insufficient or no production of insulin by the pancreas. When untreated, the disease results in very high blood glucose levels.
- Current insulin therapies require frequent user intervention (insulin administration and blood glucose measurements). These typically open-loop methods are often ineffective in maintaining blood glucose in the normal range, resulting in large fluctuations in glucose.
- While attempting to control blood glucose intensively, hypoglycemia events (low blood sugar) frequently occur due to insulin stacking or overdosing.
- Conversely, hyperglycemia (elevated blood glucose) may lead to long-term vascular complications.
- The common insulin administration route is the subcutaneous one via either multiple daily injection (MDI) or a continuous subcutaneous insulin infusion (CSII) pump.
- Type 1 diabetes affects as many as 3 million individuals in the U.S. with associated annual medical costs of $15 billion.

The Artificial Pancreas Vision

Although type 1 diabetes is currently incurable, the development of a reliable artificial pancreas would considerably improve the lifestyle of subjects with this disease. In addition to control algorithms, an automated, fully closed-loop artificial pancreas will require sensors and actuators (see below). Recent developments in continuous blood glucose monitoring (sampling rates of approximately 1–5/min) and innovations in insulin pumps (including telemetry) are promising in this context—the enabling technologies for control engineering to make an impact are well along in development (see schematic below with devices integrated into the system).



Contributor: Francis J. Doyle III, University of California at Santa Barbara, USA

Barriers and Challenges

Several technical, policy, and cultural barriers must be addressed before a viable artificial pancreas can be developed:

- Performance metrics for closed-loop control are still a subject of discussion.
- Glucose sensor reliability and accuracy remain an issue.
- Delay in insulin action makes systems sluggish.
- Intrasubject variability is a significant challenge (for example, hour-to-hour changes in insulin sensitivity).
- Closed-loop trials must meet strict regulatory standards.

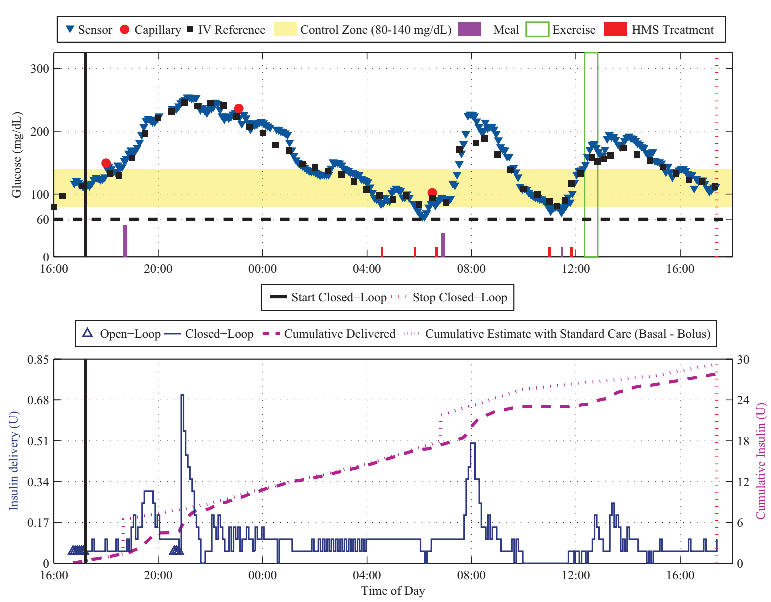
Several specific engineering challenges must also be resolved:

- Redundant arrays of glucose sensors based on different measurement methods
- Alternative insulin delivery routes, such as intraperitoneal or inhalable formulations
- Multiple-chamber pumps that will allow delivery of insulin and/or glucagon and amylin
- Personalization and dynamic predictive modeling of individual patients
- Communication and interfacing standards for the artificial pancreas
- Well-defined, clinically oriented benchmark scenarios to evaluate control design
- Faster-acting insulin formulations

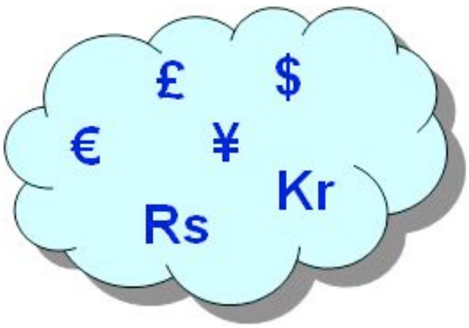
The Artificial Pancreas System and Preliminary Closed-Loop Trials

Intensive research is under way in all facets of the artificial pancreas. Multiple control strategies are being explored in human clinical studies around the world with good initial results. As demonstrated, safe glucose regulation, minimizing hyperglycemia, and overcoming meal disturbances are being evaluated in human clinical studies. A recent milestone has been the development of the Artificial Pancreas System (APS©) platform at the University of California at Santa Barbara in collaboration with the Sansum Diabetes Research Institute. It is also being used to link glucose sensors and insulin pumps using wireless protocols. The APS© platform provides a flexible mechanism for integrating hardware (such as glucose sensors and insulin pumps, in addition to computational devices), software algorithms, and human-machine interfaces (see top figure at right).

The APS© is being used in closed-loop trials around the world to test the efficacy of a variety of algorithms for blood glucose control as well as other advanced control applications. Several clinical trials using zone model predictive control with unannounced meals and exercise have been performed at the Sansum Diabetes Research Institute with good results (see bottom figure at right).



Estimating Heavy-Tailed Distributions in Finance



Financial engineering is about risk assessment and risk management. At the individual and institutional levels, the key objective is to maximize the expected return on one's investment portfolio while staying within the limits of acceptable risk.

Completely eliminating risk is impossible; for example, a bank settles for bounding risk at a 99% (or some other preselected) level, ensuring that the reserves on hand are sufficient to meet all contingencies with a probability of 0.99. These probability computations are based on historical observations—and therein lies the problem.

Crises are caused not by run-of-the-mill events, but by "extreme events"—sometimes by a confluence of extreme events referred to as a "perfect storm." So an important part of financial engineering is estimating "tail probabilities," that is, probabilities of very rare events. It is safe to say that most existing methods for fitting probability distributions to observed data are poorly suited for estimating tail probabilities. This is the challenge that provides an opportunity for control scientists and engineers.

Value at Risk (VaR)

The financial industry uses the concept of VaR (value at risk) as a metric to quantify the risk of a "position" or investment portfolio. The 1% VaR is the 99th percentile of the probability distribution function (of an individual or institutional portfolio), or equivalently, the 1st percentile of the complementary distribution function.

For example, if a portfolio of stocks has a one-day 1% VaR of $1 million, the probability is 0.01 that the portfolio will fall in value by more than $1 million over a one-day period if there is no trading. Obviously, estimating VaR accurately is crucial for financial institutions.

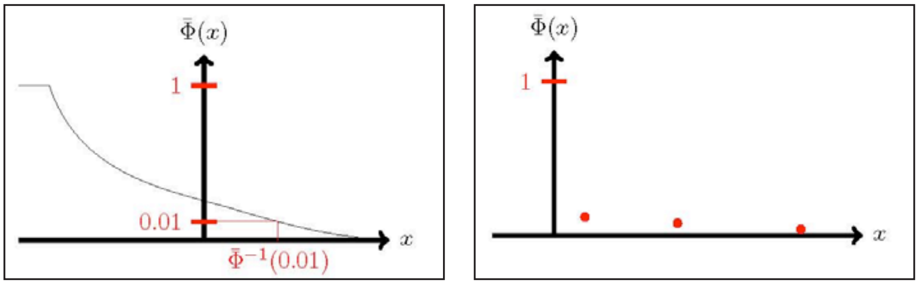


Illustration of the value at risk concept. *The plot on the left shows the complementary probability distribution of the potential loss and the VaR at the 0.01 (1%) level—the loss will be ≥ $\bar{\Phi}^{-1}$ (0.01) with probability of 0.01. The plot on the right shows the three most extreme points out of 250 (one year's data), illustrating the difficulty of estimating the 1% VaR (diagram not to scale).*

Estimating VaR

The complementary distribution function $\bar{\Phi}(\cdot)$ is often estimated using historical records. This approach poses two difficulties:

- An inadequate number of observations, and
- Improper modeling assumptions.

When historical records are used to estimate $\bar{\Phi}(\cdot)$, often a standard distribution function (such as Gaussian, Laplacian, or Pareto) is fitted to observations. Although these may give a good approximation, the key is to obtain a good fit for the "tail" of the distribution because that is where it is most crucial to estimate the risk correctly.

With daily closing records stretching over one year, only 250 data points exist for each random variable; so only 12.5 data points are available to estimate $\bar{\Phi}^{-1}$ (0.05), and a mere 2.5 data points are available to estimate $\bar{\Phi}^{-1}$ (0.01). Going back over periods longer than one year is risky as the process statistics are nonstationary and will have changed. Enlarging the number of samples by pooling data from multiple sources, such as prices of multiple stocks or multiple commodities, is also dangerous if these measurements are highly correlated; the apparent multiplicity of samples is then illusory.

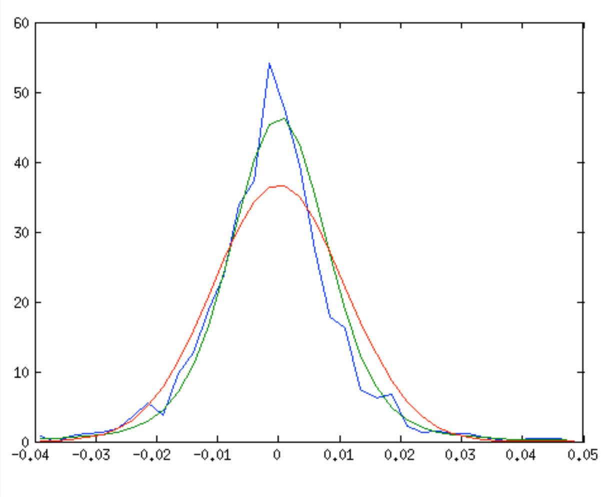
Contributor: Mathukumalli Vidayasagar, University of Texas at Dallas, USA

Heavy-Tailed Random Variables

Much of financial engineering is based on so-called "complete markets" and on the use of the Black-Scholes formula. A complete market is one in which every price movement can be "replicated" by a hedging strategy—an unrealistic assumption. A closely related assumption is that asset prices follow a log-normal distribution, or in other words, the daily fluctuations in prices viewed as percentage changes follow a Gaussian distribution. On the contrary, studies of actual asset prices show that they do not follow a log-normal distribution.

Long-term averages of asset returns can be shown to follow a "stable" distribution. Each stable distribution has an associated "exponent" $0 < \alpha \leq 2$. The Gaussian is the only stable distribution with finite variance and has $\alpha = 2$. All other stable distributions have $\alpha < 2$ and have infinite variance. If $\alpha < 1$, then even the mean can be infinite, but this situation rarely arises with actual financial data. Such random variables are said to be "heavy-tailed." Moreover, as shown in the figure to the right, real asset movements are better approximated by stable distributions with α well below the critical value of 2. Note that the smaller the value of α, the more slowly the tails decay and the greater the scope for error when Gaussian approximations are used.

Averages of heavy-tailed random variables still follow the law of large numbers (the average converges in probability to the true mean as the number of samples increases) but do not follow the central limit theorem (fluctuations about the true mean are not necessarily Gaussian). In fact, large excursions about the mean are far more "bursty" with heavy-tailed random variables. In short, "rare" events are not as rare as log-normal models would predict. This may be one reason why large swings (ten or more standard deviations when log-normal approximations are used) are far more frequent than a log-normal model would predict.



Daily Returns of the Dow Jones Industrial Average: *The plot shows the daily DJIA fractional returns from January 2000 to March 2007—1,833 samples in all. The green curve, with α = 1.6819, is the best stable fit and fits the data far better than a Gaussian (red curve, α = 2). Note that the best stable fit is also skewed, with negative returns more prevalent than positive returns. Skewed stable distributions are represented by a nonzero value for a second parameter, β (equal to −0.0651 here).*

Other Applications for Estimation of Heavy-Tailed Distributions

In addition to financial engineering, heavy-tailed distributions arise in numerous other applications. Examples include extremes in weather (e.g., rainfall) and Internet traffic. Because of their asymptotic behavior, heavy-tailed stable distributions are also referred to as "power laws."

Role for Control Engineers

Several engineers trained in stochastic control and filtering have made the transition to financial engineering, both in academia and industry. The problems of estimating parameters and signal models are common to both areas. Moreover, dealing with heavy-tailed random variables, the need for which is only now being appreciated, requires sound training in probability theory of the kind imparted to control engineers.

For more information: S.T. Rachev (Ed.), Handbook of Heavy-Tailed Distributions in Finance, Elsevier/North Holland, 2003; W.E. Leland et al., On the self-similar nature of ethernet traffic, IEEE/ACM Trans. Networking, 2(1), 1-15, 1994.

Geoengineering the Earth's Climate: The World's Largest Control Problem

Impacts and risks associated with climate change will continue to exist, even with aggressive reductions in CO_2 emissions. *Is it possible to intentionally intervene in the climate system to reduce risk?*

The climate can be "engineered" in two ways: by removing CO_2 from the atmosphere (slow and expensive) or by blocking some sunlight—this could be fast and cheap.

Solar geoengineering or *solar radiation management* refers to any large-scale intentional technique that reflects some incoming sunlight back to space, cooling the planet. This constitutes a control problem.

Feedback to Manage Uncertainty

A common question in geoengineering is: *"How can we engineer a system that we don't understand?"* The answer, of course, is with feedback.

Feedback design requires a dynamic model describing global mean temperature response to radiative forcing (defined as the difference between the radiant energy received by the earth at the upper atmosphere and the energy radiated back into space).

Simple dynamic models have been developed that are feasible for use in feedback control. These models show a good match, for control purposes, with complex climate models such as the HADCM3L coupled atmosphere-ocean general circulation model (AOGCM).

How Might the Planet Be Cooled?

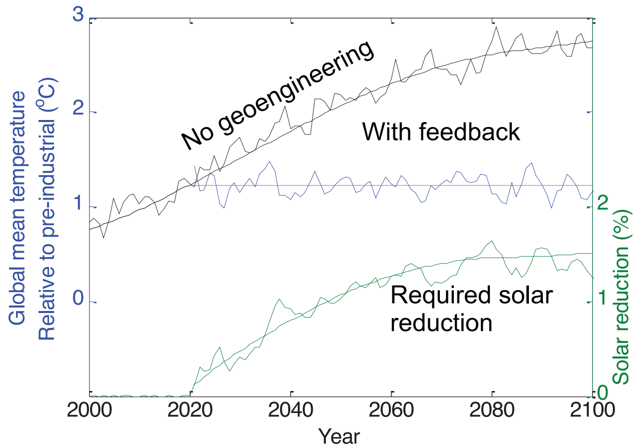
No solar geoengineering experiments have been conducted or are currently planned; however, we can learn from nature about possible "actuation mechanisms" that would act to cool the planet. These include creating a layer of stratospheric sulfate aerosols (as have resulted from volcanic eruptions; see figure below) or brightening of marine stratocumulus clouds (as in "ship tracks"—more reflective clouds caused by ship exhaust).

2009 Sarychev volcanic eruption as seen from the International Space Station (Source: NASA)

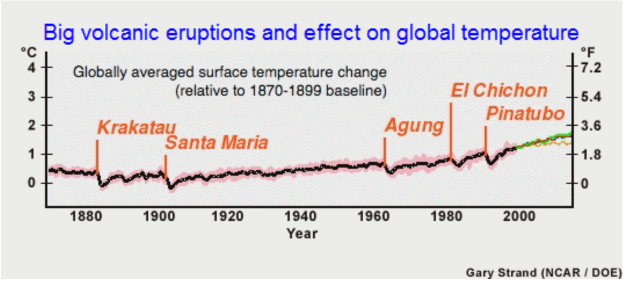
Risks of Using Solar Geoengineering

- May lead to delay in reducing CO_2 emissions, thus requiring geoengineering for centuries
- Possible "side effects" such as ozone loss, whiter skies, and unknown unknowns
- Will not compensate climate changes precisely, leading to regional differences
- Uncertainty: the need to design for a system that is not fully understood
- Must get it right the first time; experimentation is not possible

Control theory can help address these



Simulations with the HADCM3L AOGCM demonstrate adjusting solar reduction using feedback of the global mean temperature to maintain a target level. A simple PI controller was used, with feedback gains designed using a reduced-order dynamic model. The resulting model predictions are shown as smooth lines and match the full climate model (the lines that include variability).

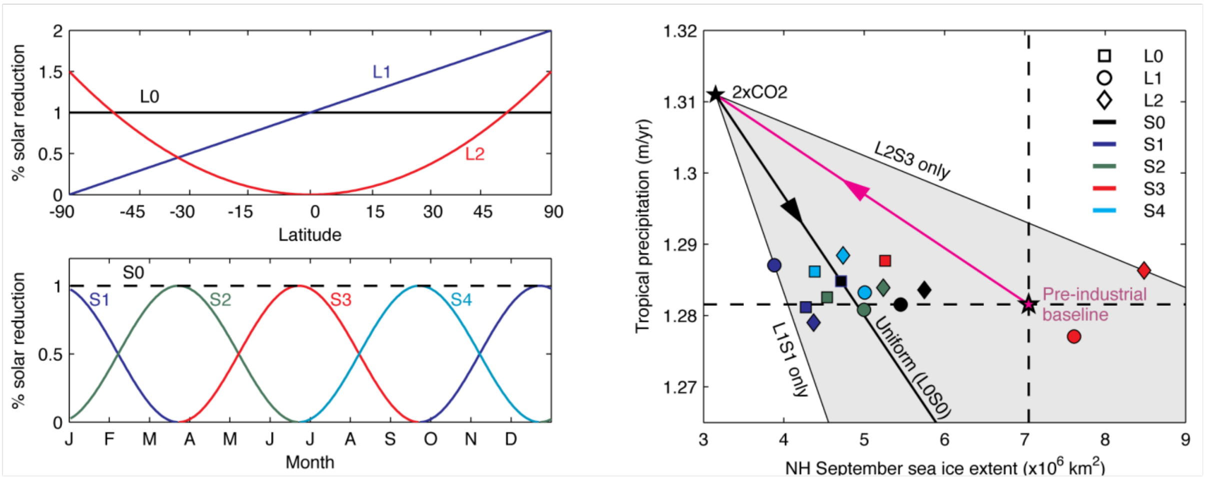


Major volcanic eruptions in recent history and their effect on global temperature

Contributors: Douglas MacMartin, California Institute of Technology, USA, and Ben Kravitz, Pacific Northwest National Laboratory, USA

Optimization to Minimize Impacts

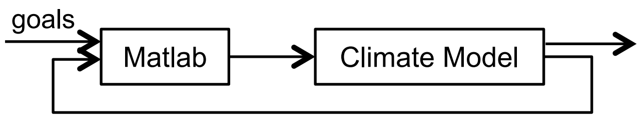
Compensation of climate change with solar geoengineering is not perfect, so some regions might benefit from more geoengineering than other regions, leading to the question: *"Who gets to set the thermostat?"* However, we can vary the distribution of solar reduction in space and time and we can choose distributions to maximize benefits, such as restoration of northern hemisphere (NH) sea ice. See the figure below for a simulation example.



The example above considers 15 "basis functions" (shown at left), with solar reduction uniform or peaking in each season and spatial reduction constant or varying with latitude. The right panel shows how two specific climate impacts of a doubling of CO_2 concentration relative to the pre-industrial baseline can be ameliorated by solar geoengineering: a uniform solar reduction cannot perfectly compensate for increased CO_2, but varying the distribution of solar reduction can influence multiple variables—any climate in the shaded region can be chosen by a linear combination of the basis functions. (Of course, we care about more than these two variables; we can't independently change the climate everywhere with only 15 degrees of freedom.)

More Opportunities for Control!

Control theory and techniques can be widely useful in climate science. Wrapping an external feedback loop can be used to:



Explore human/climate interactions:

- Geoengineering—how to manage multiple goals despite deep uncertainty, high background variability (poor signal-to-noise ratios), and the need to "get it right" the first time without prior experimentation
- Determine optimal emission pathways to achieve climate goals with learning about uncertainty

Explore natural climate behavior:

- Example: permafrost thaw releases methane, causing warming, causing more methane release, and so on
- Understand climate variability, e.g., El Niño or meridional overturning dynamics: use system-identification tools to evaluate subsystem process dynamics from data and in different models

Explore climate model behavior:

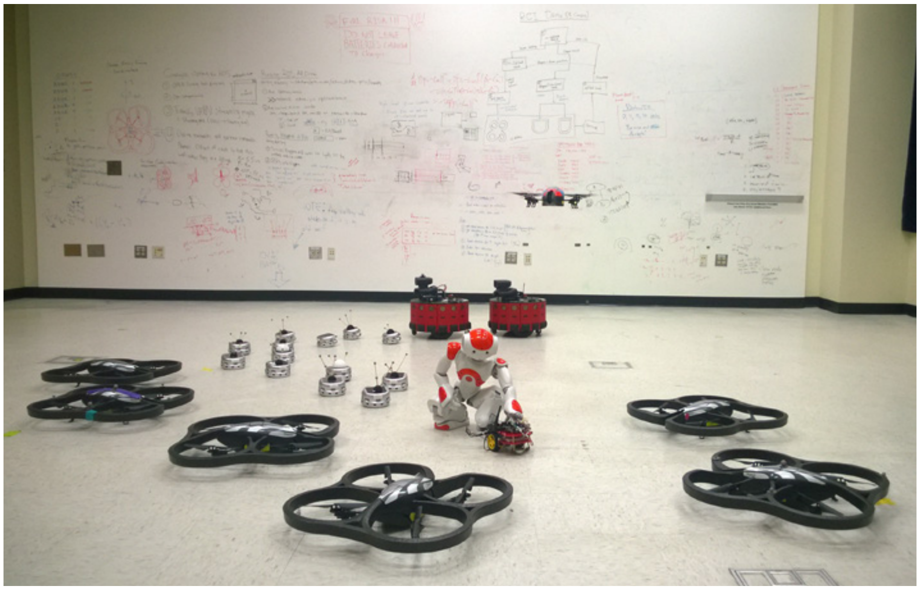
- Can we use feedback to automate and accelerate model tuning?
- Can we use feedback to artificially alter time constants to accelerate model convergence?
- Compare different greenhouse gases: e.g., how much CH_4 is equivalent to how much CO_2?

For more information: D.G. MacMartin et al., Dynamics of the coupled human-climate system resulting from closed-loop control of solar geoengineering, Climate Dynamics, 2013; D.G. MacMartin et al., Management of tradeoffs in geoengineering through optimal choice of non-uniform radiative forcing, Nature Climate Change, vol. 3, pp. 365-368, 2013. B. Kravitz et al., Explicit feedback and the management of uncertainty in meeting climate objectives with solar geoengineering, Environmental Research Letters, vol. 9, no. 4, 2014.

Human Interactions With Complex Networks

Human-Swarm Interactions

Imagine you are surrounded by a million robot mosquitoes and you have a single joystick you can use for interacting with the swarm. How should this interaction be structured?



Appropriate abstractions of the swarm/complex network of agents are needed that are controllable (Can the operator use the abstractions to achieve the desired performance?) and observable (Are the abstractions transparent, i.e., can they be inferred by and acted upon by the operator?).

Inverting the Many-to-One Relationship

As many facets of society move toward greater levels of automation, the current many-to-one relationship, wherein multiple operators are required to control a single dynamical agent (e.g., an autonomous vehicle), is not sustainable. Instead, a single operator needs to be able to influence and control large collections of agents over an interconnected network. At its core, the problem of devising effective control strategies for making human operators able to control complex networks aims at inverting the current many-to-one relationship.

Multi-Agent Robotics

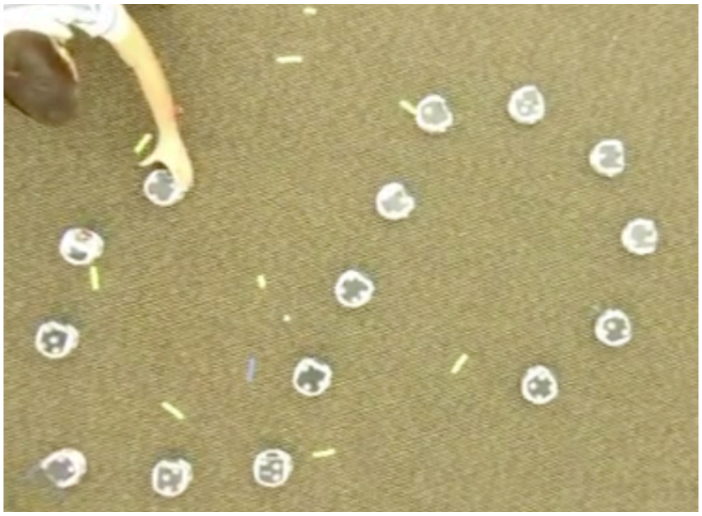
One particular application domain where human operators must control large collections of agents is multi-agent robotics. Key objectives that the interactions must support include

- Formation control (How can the robots be driven to particular shapes?);
- Coverage control (How can they be made to cover an area?);
- Swarming and flocking (How can coordinated behaviors be enforced?).

All of these objectives can be cast in terms of desired global geometries and shapes, and the human operator should be able to specify the shape and influence the multi-robot network to achieve the desired shape.

Leader-Follower Interactions

The image below shows an example of the leader-follower paradigm for controlling teams of robots wherein the user takes control of individual leader-agents in the network. By manipulating the leaders, the rest of the network can be controlled indirectly through interagent couplings effected by wireless communication. Alternatives to the leader-follower paradigm include boundary control (boundary agents are manipulated), fluid-based interactions (the operator "stirs" the team of robots), and behavioral interactions (behaviors rather than positions are manipulated by the operator).



Contributor: Magnus Egerstedt, Georgia Institute of Technology, USA

Operator Decision Support

One significant challenge facing the successful deployment of unmanned aerial vehicles (UAVs) in unstructured environments is the level of human involvement needed to carry out the mission. Control strategies are needed that will allow pilots to control and coordinate multiple UAVs.

Affordances and User Interfaces

An affordance is a relation between an object and a user wherein the object allows the user to perform a particular action. The affordances identified when controlling a swarm include stretching the swarm, molding it into a particular shape, splitting and merging sub-swarms, and mixing of different swarms. The image at right shows one possible "swarm interface" that allows the operator to interact with the swarm by manipulating the shape of a deformable medium (clay).

Influence vs. Control?

In a network of agents that are updating their states according to some interaction protocol, one can ask how easy or hard it is to influence such a network. The answer to this question depends on several factors, such as the form of the interaction dynamics, the structure and dynamics of the underlying information-exchange network, and the mechanism whereby the external influence is injected into the network. But key among these factors is what is meant by *influence* itself. This term can mean anything from very precise, global coordination of all agents to instantaneous (or short-term) notions of ensemble-level trends. Several interesting and important questions need to be answered, such as:

- What are appropriate system-theoretic notions for characterizing how easy or hard it is to influence complex networks?
- Can social influence be characterized in system-theoretic terms, thus bridging two different ways in which influence is understood?
- Some networks may be easy to control algorithmically but hard for human operators to control. How can this distinction be quantified in a precise manner?

For more information: J.P. de la Croix and M. Egerstedt, Controllability characterizations of leader-based swarm interactions, AAAI Symposium on Human Control of Bio-Inspired Swarms, Arlington, VA, 2012; Y.Y. Liu, J.J. Slotine, and A.L. Barabasi, Controllability of complex networks, Nature, 2011, pp. 167-173.

Lithium-Ion Battery Management

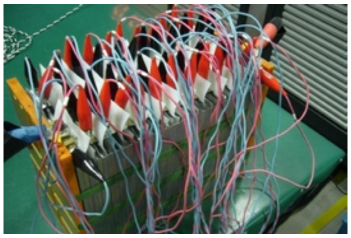
Control and monitoring remain outstanding challenges in lithium-ion batteries for high-power applications despite the progress of the technology.

Since the commercialization of lithium-ion rechargeable batteries by Sony Corporation in 1991, the basic principles of lithium intercalation and rocking-chair operation have remained unchanged, but new material types and enhanced chemical compounds for electrodes and electrolytes are being continuously introduced.

Capacity improvements have been incremental, and commercially available 200-Wh/kg cells are still far from the theoretical values of about 600 Wh/kg for current chemistries and up to 11 kWh/kg for the holy grail of lithium-air cells (an energy capacity that is close to gasoline's 13 kWh/kg). Specific power, on the other hand, has significantly improved and is getting close to a milestone of 100C for charge and discharge current (i.e., current limit in amperes is a hundredfold of cell capacity in ampere-hours).

High energy and power density make lithium-ion a technology of choice for automotive and aerospace applications; however, safety issues and susceptibility to rapid aging are drawbacks, which are currently addressed by conservative control strategies that sacrifice battery capacity and power.

Lithium-Ion Battery Safety

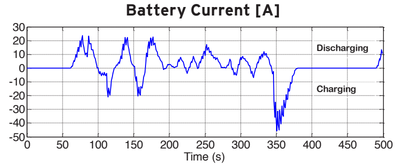


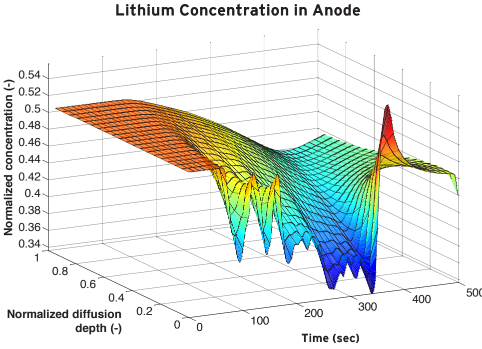
Experimental 24-V/100-Ah lithium-ion battery pack with 40 cells for automotive applications

Lithium-ion technology is prone to thermal runaways with low temperature threshold. In combination with a flammable organic electrolyte solvent and metal-oxide content in electrodes, these runaways can lead to violent self-sustaining fires that can spread rapidly among cells in a battery pack (a phenomenon that has prompted the recall of tens of millions of Sony batteries, concerns about the Chevrolet Volt fire risk, the grounding of Boeing 787s, etc.). Safety problems arise from both design/manufacturing issues (contamination of metallic particles leading to slow dendrite formation and short-circuiting of separators) and also from insufficient insight into battery operation and battery internal state. Protection systems should not rely solely on measurements, but should be able to estimate and predict internal state for early detection of hazardous events.

Lithium-Ion Battery Aging

Lithium-ion batteries have a tendency toward rapid aging (capacity and power decrease, internal resistivity increases) if used close to operational constraints (overcharging, full discharge, excessive current and temperature, etc.). Since battery cost is a significant part of the total electric or hybrid vehicle cost, the powertrain controller should optimize for mileage and battery amortization jointly. This requires modeling battery aging and battery state-of-health estimation; however, aging is a combination of complex mechanisms (formation of ion transport inhibiting deposits, lithium plating, passivation, crystal grid collapse, etc.).





Under high-current operation, high variation of lithium concentration on the anode surface (normalized diffusion depth of 0) creates "hidden waves" in the internal lithium concentration profile (normalized diffusion depth >0). Knowledge of the internal profile is essential for estimating parameters such as state of charge, but voltage measurements only reflect the surface concentration.

Contributors: Pavel Trnka and Vladimir Havlena, Honeywell, Czech Republic

From: *The Impact of Control Technology*, 2nd ed., T. Samad and A.M. Annaswamy (eds.), 2014. Available at www.ieeecss.org.

Challenges: Control-Oriented Modeling for High-Power Applications

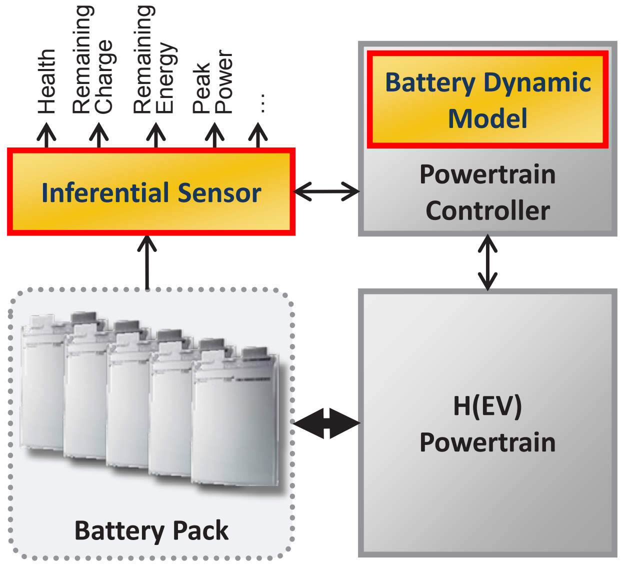
- High current density causes spatial imbalance of active species resulting from fast electrochemical phenomena and slow diffusion kinetics. This distributed-parameters behavior cannot be ignored.
- Accurate electrochemical models are based on partial differential algebraic equations driven by the Butler-Volmer kinetic equation and diffusion laws; these models are too complex for closed-loop control.
- Modeling of the aging mechanism is not tractable even on the level of electrochemical models (e.g., phenomena on the solid-electrolyte interface) and has to be approximated by empirical models.
- Successful single-particle models have limited applicability for high-power applications.
- The main challenge is bridging the gap between high-fidelity electrochemical models and control-oriented models targeted to embedded applications.

Challenges: Experiment Design

- Battery experiments consume substantial time and resources, especially lifetime cycling and testing under a wide range of conditions.
- Parameter identification of electrochemical models is difficult.
- Electrochemical impedance spectroscopy only captures steady-state behavior.
- Major challenges are the design of reasonable experiments for aging models and parameter identification algorithms.

Challenges: Battery Aware Optimization

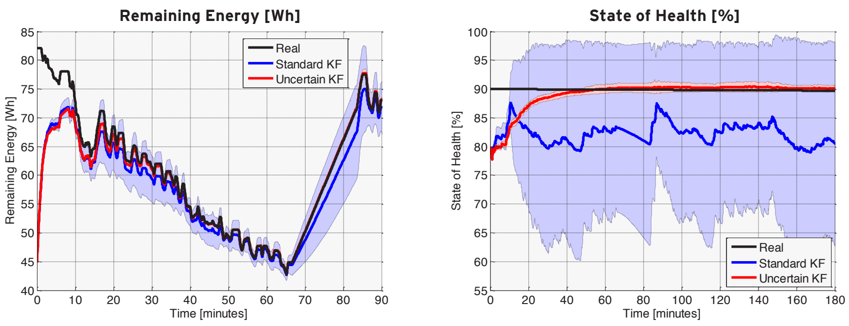
- Powertrain optimization with a criterion combining mileage and battery amortization.



An inferential sensor for a Li-ion battery pack is needed for several variables in hybrid or electric vehicle powertrains.

Challenges: Inferential Sensor Design

- Design of an inferential sensor must be robust to manufacturing variability (see graphs below).
- Variables of interest, apart from state of charge, include remaining energy, current peak power, heat production rate, and remaining effective charge capacity.
- Early detection of thermal runaways is necessary.
- Observability issues arise from limited measurements and distributed-parameter-based models.



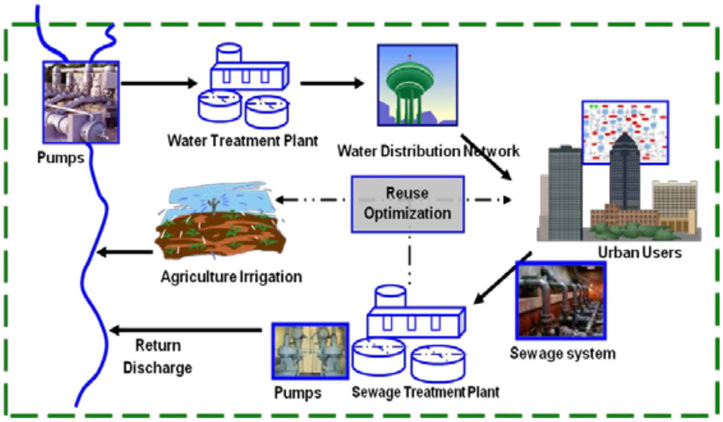
Inferential sensing of battery remaining energy (left) and state of health (right) under manufacturing variability. The black trajectory is a simulated model with "real" parameters. The blue and red trajectories show estimation with a standard Kalman filter (KF) and a Kalman filter for uncertain models over multiple runs with model parametrization taken randomly from manufacturing uncertainty. Solid lines are mean values, and shaded bands illustrate the standard deviation range in each case. Knowledge of, and design for, manufacturing uncertainty is essential for accurate estimation.

Management of Complex Water Networks

Water security is one of the most tangible and fastest-growing social, political, and economic challenges faced today. It is also a fast-unfolding environmental crisis. In every sector, the demand for water is expected to increase and analysis suggests that the world will face a 40% global shortfall between forecast demand and available supply by 2030.

This outlook bears potential for crisis and conflict since water lies at the heart of everything that is essential for human life: food, sanitation, energy, production of goods, transport, and the biosphere. Water ensures not only mere survival of humans, but also social well-being and economic growth. In addition, water is a renewable, yet not inexhaustible, resource—it cannot withstand constant over-extraction and being depleted faster than being renewed. What is more, water cannot be substituted.

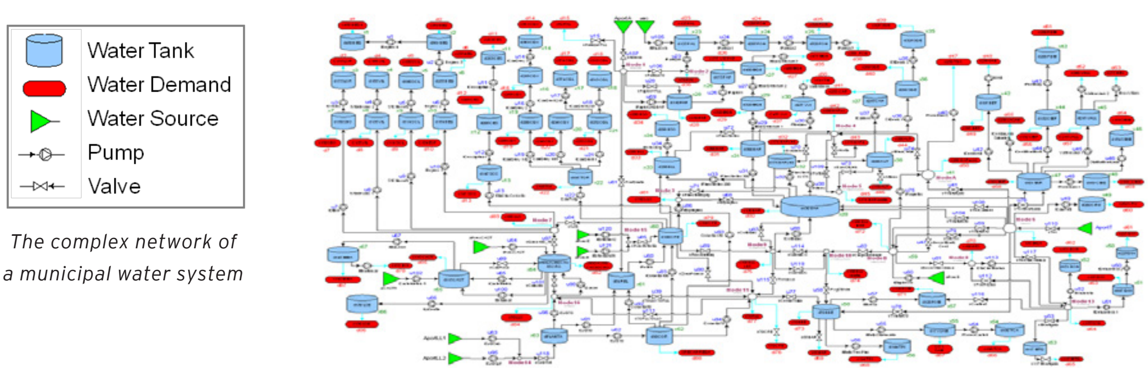
—World Economic Forum 2012



Municipal Water Management System Scope

Control and optimization solutions for municipal water management systems will need to encompass a set of interacting layers that define a variety of problems.

- The water cycle starts with water sources and mechanical/chemical/biological water treatment; here, fairly standard advanced process control approaches can be used to produce a given amount of potable water with predefined quality and minimized consumption of energy and chemicals.
- The next step is to transport the water to the tanks and water towers; optimal energy-efficient control of pumping stations and valves is closely related to network topology, altitude profile, storage capacities, pumping efficiency curves, and electricity cost profile. For model maintenance and real-time responsiveness to network topology changes, tight integration with a geographic information system (GIS) is beneficial. Energy efficiency can be significantly improved if reliable prediction of water demand/consumption is available. The transport layer also contributes to more effective quality control by mixing water from different sources; quality constraints related to water aging are also treated here.
- The water from the tanks is distributed to end users; optimal control of this distribution layer covers pressure zone monitoring and control by booster pumps and reduction valves to guarantee minimum pressure to all consumers and reduce leakage by maximum pressure reduction.
- The sewage water collection and treatment process follows a similar hierarchy of layers and also usually involves storm water management.



The complex network of a municipal water system

Contributors: ***Vladimir Havlena and Pavel Trnka, Honeywell, Czech Republic; Bill Sheridan, Honeywell, USA***

Municipal Water Management System: Hierarchy and Time Scale

Economic optimization and planning of network topology and instrumentation updates is done over several years.

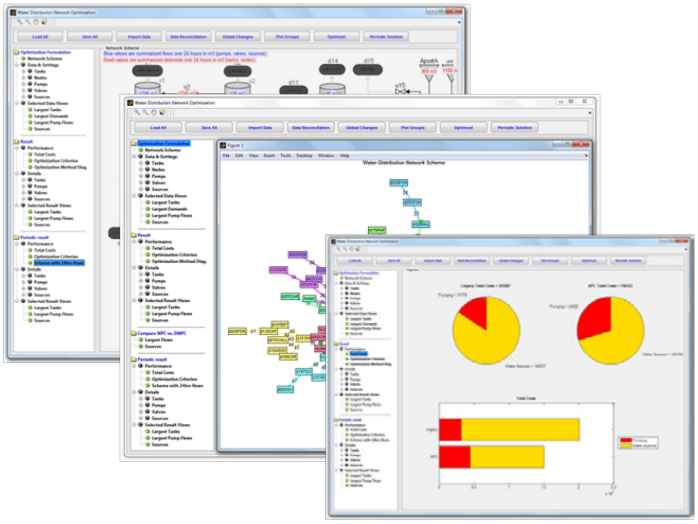
The water treatment and transport layers operate with a control/optimization horizon of up to 1 week with a sampling rate from 15 minutes to 1 hour to capture periodic demand, varying production costs, and electricity tariffs. Typically, heuristic control laws with intermittent operator interventions are used today. The next-generation solution of this large-scale spatially distributed control and optimization problem can be achieved by distributed model predictive control. Prediction of electric energy cost and water demand is a key component of such a solution.

The distribution layer is typically operated by programmable logic controllers (PLCs) with sampling periods of 0.1–1 sec. Several preprogrammed pressure levels (day/night) or fixed pressure setpoint profiles are used in the distribution network feeding points. The next-generation solution should provide active pressure zone control based on network pressure profile monitoring by sensor networks.

Challenges: Modeling

Mathematical models are the primary information source for advanced model-based control and real-time optimization techniques. A succession of reduced-complexity models for individual layers of the optimization, control, and monitoring hierarchy, with automated adaptation to network changes (network extensions, manual adjustment of routing valves), is needed. For facilitating model maintenance, tight integration with detailed reference hydraulic models available in a GIS is desirable.

Traditional issues such as modeling of nonlinear mechanical control valves (altitude valve, check valve) and maintenance of pump performance curves (mostly tabulated) will have to be resolved on a large scale. Expertise in new domains such as modeling of water aging/chlorination will also be needed.



Challenges: Planning and Scheduling

Operation planning is based on uncertain predictions of water availability and demand profiles. Estimated availability from the wells with lowest treatment cost may depend on rainfall prediction. Emerging technologies, such as rainwater harvesting and wastewater reuse, may reduce the level of uncertainty.

Models can also be reused as a decision support tool for long-term strategic planning and what-if scenario analysis.

Challenges: Network Control Theory

Modeling and control of the network with flows following Kirchhoff's laws (incompressible fluids) can build on structural information. Examples include model reduction for control/optimization/monitoring; aggregation in time and space (skeletonization); and structural controllability/observability analysis based on network topology. Advances in network control theory will also be beneficial for optimizing locations of additional actuators and sensors in the network to meet more stringent performance criteria.

Challenges: State Estimation

Network monitoring operates with limited information subject to unmeasured disturbances. Raw data reconciliation and leakage assessment are essential to minimize loss of treated water. Monitoring approaches should incorporate bottom-up (pressure zone modeling) and top-down (balancing individual metered areas) approaches and leverage the capabilities of smart sensor networks. The Bayesian approach for nonlinear state estimation enables use of model-based techniques to detect abrupt changes (bursts) as well as background losses. Overdosing of chemicals can be reduced if a reliable estimate of water quality/chlorine concentration is available. Maintenance of buried pipes based on proven corrosion monitoring algorithms can save significant investment costs as well as reduce water losses.

Challenges: Advanced Control and Real-Time Optimization

The advanced control and real-time optimization problem has a well-defined structure described by varying network topology and multi-objective optimization criteria. For the network operator, independent operation of individual subsystems with strong interactions must be enabled, with a progressive area-by-area approach to the commissioning of an overall solution. For large municipal applications, computational complexity is still prohibitive for closed-loop real-time applications; however, feasible computation time can be reached by decentralized solutions based, for example, on dual decomposition methods. The control strategy must be robust to communication failures and demand prediction uncertainty.

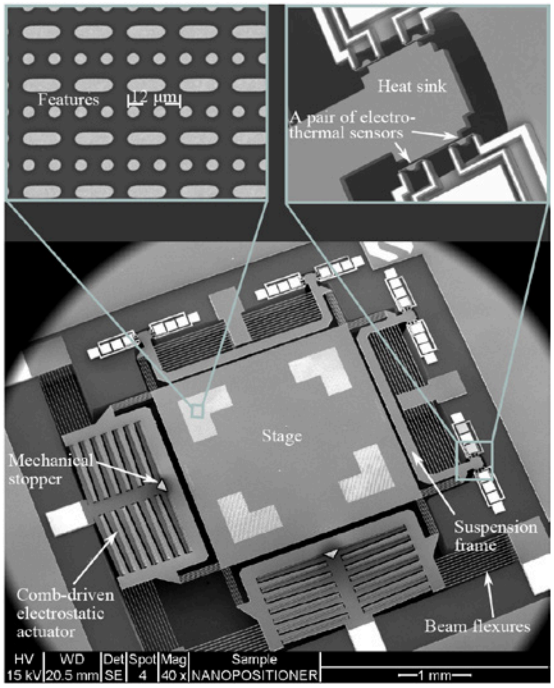
MEMS-Based Nanopositioning for On-Chip Atomic Force Microscopy

An atomic force microscope (AFM) operates by running a sharp tip, positioned at the end of a microcantilever, over a sample in a raster pattern such that a 3-D image is obtained based on the z-axis deflection and in-plane position of the tip. Typically, a laser beam is focused on the end of the cantilever, with the deflection of the reflected beam indicating the height of the cantilever and therefore the topography of the sample. The AFM is one of the most versatile methods for imaging structures at nanometer scale. The ability to operate in a non-vacuum environment gives the AFM a significant advantage over competing microscopy methods such as the transmission electron microscope (TEM) and the scanning electron microscope (SEM). Apart from imaging, the AFM is used to manipulate matter at nanometer scale and is viewed as the dominant tool in nanorobotics. The AFM's ability to image and manipulate matter at nanometer scale is entirely dependent on the use of several feedback loops. This gives rise to numerous opportunities and a significant need to apply advanced feedback control methods in AFM.

MEMS Nanopositioner

An important component of an atomic force microscope is a nanopositioner that moves the sample, relative to the probe, in a raster pattern. A typical AFM nanopositioner is a large, heavy flexure-guided mechanism machined from a solid block of steel or aluminum, with incorporated actuators and displacement sensors. The most widely used actuation technology for nanopositioning is the piezoelectric stack actuator, which can generate a large amount of force with a small stroke. These actuators suffer from nonlinearities such as hysteresis and creep, which are difficult to address using feedforward methods. Furthermore, from an electrical viewpoint, they are large capacitors that require complex and expensive low-noise, linear amplifiers for their operation. There is significant interest in moving away from using piezoelectric actuators in nanopositioning systems.

One promising approach is to develop micro-electromechanical (MEMS) nanopositioners that can function as scanning stages of future atomic force microscopes. These miniaturized systems potentially hold several advantages over conventional macro-sized nanopositioners. Qualities such as increased operating bandwidths, lower unit manufacturing costs, simpler bulk fabrication, and a much smaller packaged size mean that MEMS-based nanopositioners represent an attractive solution for many applications, particularly for atomic force microscopy.



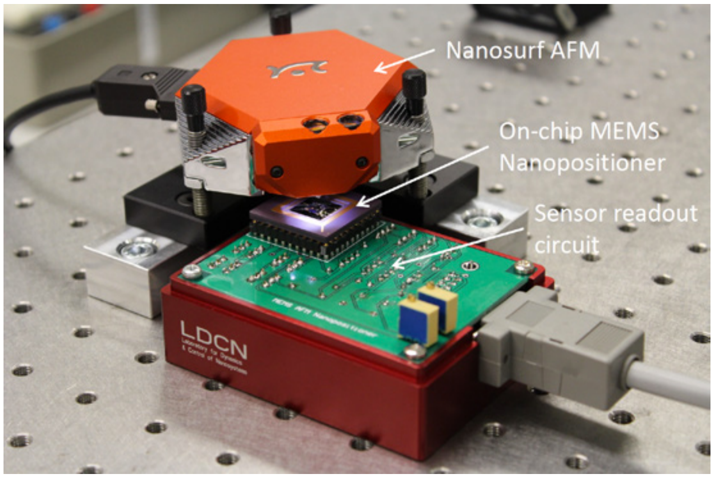
The device shown at left is a nanopositioner fabricated using a silicon-on-insulator MEMS process. The design features integrated electrothermal sensors that enable real-time measurements of the stage displacement along the x and y directions. The scanner has two mechanical degrees of freedom, with electrostatic comb-finger actuators being used to position a 3-mm × 3-mm stage along the planar x and y directions. The mechanical design of the nanopositioner is based on a parallel-kinematic configuration, and a series of beam flexures around the perimeter of the stage are used to position the stage along the x and y axes and also to decouple the motions of the two axes. Each of the nanopostioner's electrostatic actuators features interdigitated comb fingers with dimensions chosen to maximize the force generated by the actuator for a given actuation voltage.

Contributor: S.O. Reza Moheimani, University of Newcastle, Australia

Control of MEMS-Based AFM

The Need for Feedforward and Feedback Control

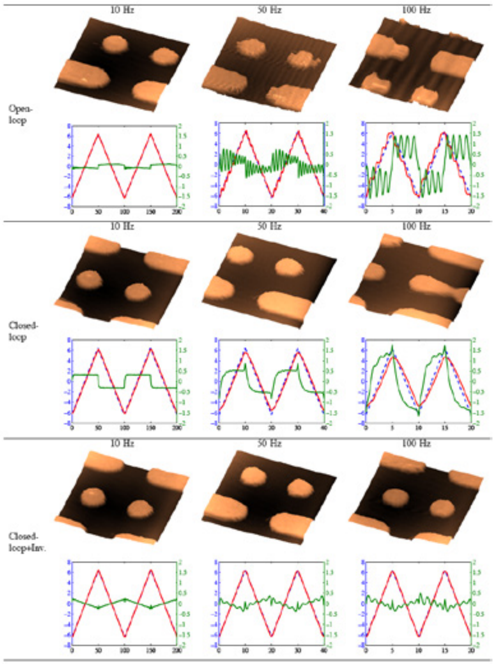
MEMS nanopositioners are typically highly resonant systems, and their high-speed operation is prone to scan-induced vibration. Therefore, to achieve the required positioning accuracies, which may be on the order of fractions of a nanometer, feedback control is essential. Sensor noise is a key issue in MEMS nanopositioners. Electrothermal displacement sensors are highly suitable for MEMS nanopositioning since they can be realized with a very small footprint; however, they suffer from flicker noise and low-frequency drift. The control system must be able to deal with these issues. The presence of cross-coupling between the two lateral axes of the nanopositioner is another contributing factor to low image quality, which can be improved by proper design of the feedback control loop. Finally, the closed-loop bandwidth achievable with a feedback controller may not be enough for high-speed scans. Thus, a feedforward controller may have to be used in addition to feedback control.



Experimental setup of the AFM and MEMS nanopositioner in a scan-by-sample mode

Benefits of Control: Experimental Results

A control system was designed and implemented on the MEMS nanopositioner. The experimental setup consisting of the MEMS scanner mounted on a printed circuit board, together with the readout circuitry and a Nanosurf EasyScan2 AFM, is illustrated in the above image. The experiments were performed in the "scan-by-sample" mode where the scan table, which is deposited with calibration features (illustrated in the SEM micrograph on the previous page), was moved in relation to the static probe. An image area of 12.7 µm × 12.7 µm was scanned at 10 Hz, 50 Hz, and 100 Hz in open-loop, closed-loop, and closed-loop with inversion-based feedforward. The figure below plots the three-dimensional topography images, the fast x-axis displacements, and tracking errors of the nanopositioner. The important role of control in improving image quality at high scan speeds is evident.



AFM scan results obtained at 10 Hz, 50 Hz, and 100 Hz in open-loop, closed-loop, and closed-loop with inversion-based feedforward. 3-D topography of the sample is plotted. The fast axis displacements (µm vs. ms) are plotted in red and tracking errors (µm vs. ms) are plotted in green. Reference signals are plotted in blue.

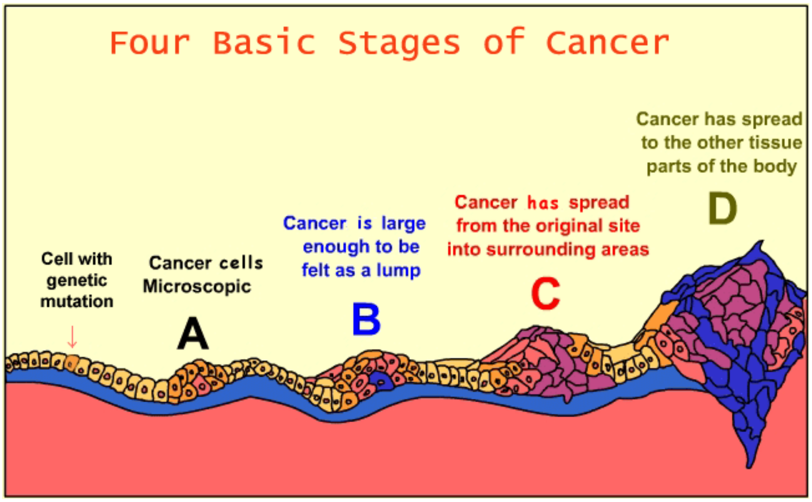
This work was supported by the Australian Research Council. The author thanks the research team in the Laboratory for Dynamics and Control of Nanosystems, particularly Anthony Fowler, Ali Mohammadi, and Yuen Yong, for their suggestions and contributions.

Modeling Cancer Dynamics and Tumor Heterogeneity

Cancer is the second leading cause of death in most countries, both advanced and developing. Worldwide, cancer accounts for roughly one out of every eight deaths, or about 7.6 million per year. The wide variety of cancer manifestations makes it one of the most challenging diseases for the medical community to treat. Recent advances in experimentation provide an opening for the mathematically minded to study several aspects of cancer, such as:

- Modeling the progression of cancer, specifically, explaining the observed pattern of progression whereby a tumor has a long quiescent period lasting many years, often followed by a period of very rapid growth.
- Capturing the response of tumors to therapy whereby they become both more homogeneous and more resistant to therapy.
- Predicting the response of tumors to a combination of two drugs on the basis of the response to each drug individually.

The first two of these challenges are discussed in more detail on the following page.



Source: National Cancer Center of Singapore, http://www.ncs.com.sg

A Brief Sketch of Cancer

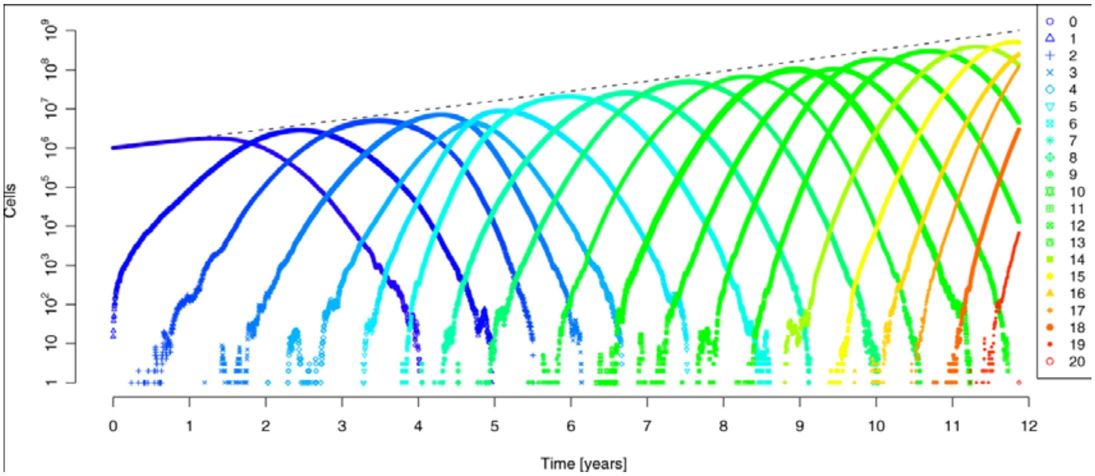
The human body contains about 50 trillion cells, each of which contains (in principle) an identical copy of one's DNA. Human DNA consists of roughly 22,000 genes. Cells undergo division at a rate of roughly once a day; they also undergo programmed death, known as apoptosis. With so much division going on, occasionally errors in DNA replication can occur, leading to mutations. Also, exposure to external agents such as radiation or other carcinogenic substances can cause mutations. When a cell's DNA gets mutated, it might gain a "fitness advantage" over normal cells whereby it divides more quickly or dies at a slower rate, or both. Starting from a small benign population of cells with mutated DNA, the number of such cells usually exhibits a very slow and sustained growth pattern, often lasting many years, before (in many cases) suddenly bursting into a period of rapid growth.

Models of Tumor Growth: Today's State of the Art

Models for tumor progression are based on population genetics. One of the more biologically realistic models is the so-called Wright-Fisher model wherein a population of fixed size consists of two or more subtypes and individual cells make a transition from one subtype to another in a Markovian manner.

Simulation results with this approach are illustrated in the figure on the following page. Each gene is assumed to undergo mutation at a fixed rate, which is the same for all genes. Also, a mutation in any one gene confers a fitness advantage to the cell, and the overall fitness advantage is a simple sum of the advantages arising from individual gene mutations. Finally, the mutations in each gene are assumed to be independent of those in all other genes. With this model, and realistic parameters for various rates, the growth in the number of cells that contain a certain number of mutations can be simulated. The figure on the following page shows that as time passes, the average number of cells with mutations increases linearly; however, this linear progression is not supported by experimental evidence on the growth of actual tumors.

Contributor: Mathukumalli Vidayasagar, University of Texas at Dallas, USA



Growth in the populations of cells with increasingly many mutations based on the Wright-Fisher model. It takes approximately 12 years for cells with 20 mutations to begin to appear and roughly 20 years to reach a critical size of one billion (not shown in figure). (Source: N. Beerenwinkel et al., Genetic progression and the waiting time to cancer, PloS Computational Biology, vol. 3, no. 1, e225, 2007)

Challenge: Nonlinear and Probabilistic Models for Tumor Growth

One challenge is to replace the linear growth of the Wright-Fisher model. In reality, the number of cancerous cells does *not* grow linearly; rather, it initially increases at a linear, or a slow exponential, rate, followed by a period of rapid growth. Thus, the challenge is to construct a probabilistic model that shows this behavior *without* assuming that the physiological parameters change over time, as that would be unrealistic.

One suggested approach is to divide genes into two categories: "passengers" whose mutations are by themselves only marginally harmful and "drivers" that trigger rapid tumor growth when coupled with a sufficiently large number of passenger mutations. Biological studies strongly support the notion that mutations can indeed be divided into drivers and passengers. Thus, it may be possible to postulate a branching process model whereby the mutation rate for drivers is far less than for passengers; the passenger mutations accumulate at a rate similar to that in the figure above, but when a driver mutation takes place, tumor growth is accelerated. Although this sounds like a plausible model, the challenge is to turn this hand-waving argument into something more rigorous *and* to validate the postulated model through biological experiments.

Challenge: Modeling the Heterogeneity and Therapy Response of Tumors

Until recently, the dominant view was that a tumor consists of a mixture of one's normal DNA and one dominant mutant DNA. By determining the dominant mutation, the physician can then choose an appropriate therapy, which usually consists of killing off all mutant cells. The usual picture is that practically all patients "respond" initially in that the tumor shrinks or even "disappears" (meaning that it is too small to be detected); however, in many cases, the tumor recurs. The challenge therefore is to develop dynamical models for tumor growth before and after therapy whereby tumor heterogeneity and recurrence are built into the model.

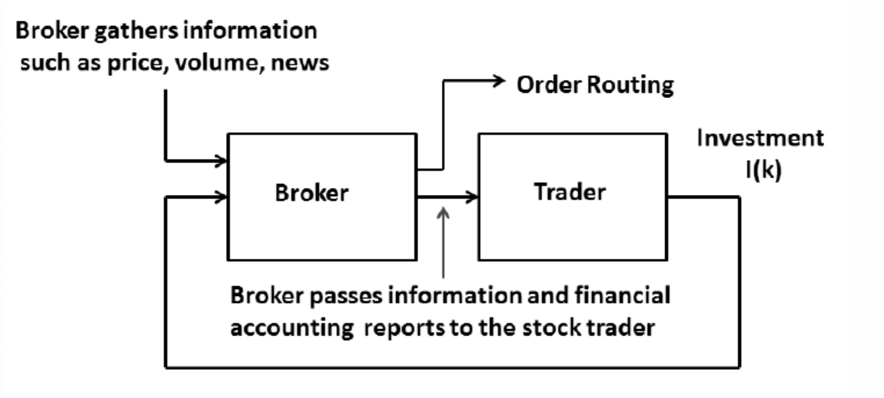
When experimental techniques only permitted measurements of tens of thousands of cells at a time, it was reasonable to assume that tumors were quite homogeneous in having just one dominant mutation. But experimental techniques have now improved to a point where it is possible to study hundreds of cells and occasionally even a single cell. These studies suggest that a tumor can contain many different mutant forms of DNA. If therapy targets the most dominant mutation, the second most dominant mutation could take over; and so on. Thus, a working paradigm for therapy, which is yet to be validated, is as follows: When a tumor first grows, it consists of a mixture of several mutant DNAs, each of them having a fitness advantage over cells with normal DNA. When some therapy is applied, the tumor initially shrinks because the dominant mutation(s) are killed off, but the tumor eventually recurs due to other mutants. Moreover, the recurring tumor is often resistant to the original therapy.

Therefore, if it is possible to construct a mathematical model of such behavior and then to "identify the parameters" in the model based on molecular analysis of the tumors, the physician could be ready with a Plan B and a Plan C, etc., right from the outset. These mathematical models would have to be validated through experiments, but any progress in this direction would result in a fundamental paradigm shift in cancer therapy.

For more information: S. Mukherji, The Emperor of All Maladies, Fourth Estate Press, London, UK, 2011; W.J. Ewens, Mathematical Population Genetics, Springer-Verlag, Berlin, 2004; L.D. Wood et al., The genomic landscape of human colorectal and breast cancers, Science, vol. 318, no. 108, pp. 1108–1113, 2007.

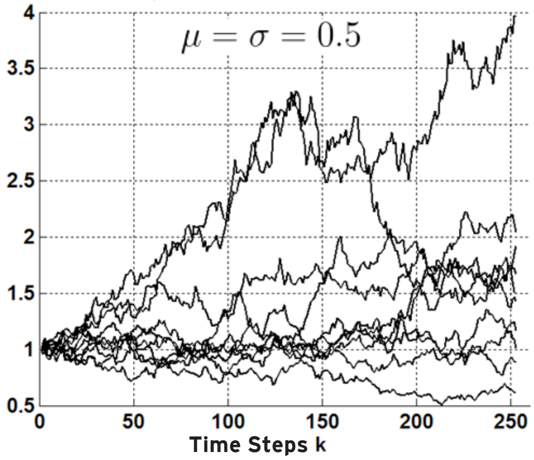
Opportunities for Control Theory in Stock Trading Research

In the world of stock trading, *technical analysis* generally refers to a class of strategies and indicators that are based on the use of price and volume data. Despite the heavy use of technical analysis by many professional market managers on Wall Street and individual investors alike, such strategies are one of the most controversial subjects in the area of finance. Although proponents of the Efficient Market Hypothesis have long claimed that technical analysis has little to no theoretical justification, there is an increasing body of literature to the contrary. Based on statistical analysis of back-tests using empirical data, numerous authors have argued that many technical analysis methods have significant validity. This dichotomy between theory and statistics is being addressed by ongoing research in the control community that models these popular methods from a feedback control perspective. In the process, this research opens the door to a rigorous new framework for theoretical analysis that is lacking in the existing finance literature.



On Stock Trading Without a Model

In its "purest" form, technical analysis involves little or no reliance on an underlying model for stock prices. Instead, price patterns or other patterns such as those of gains and losses or trading volume are used to determine how an investment should be modulated over time. Many refer to this approach as "model free" or "reactive"; that is, the controller determines the investment level by reacting to the observed pattern without regard to an explicit prediction of future prices. This method is consistent with the viewpoint of robusticians in the control field (i.e., distrust of the underlying price model dictates the development of strategies along these lines).



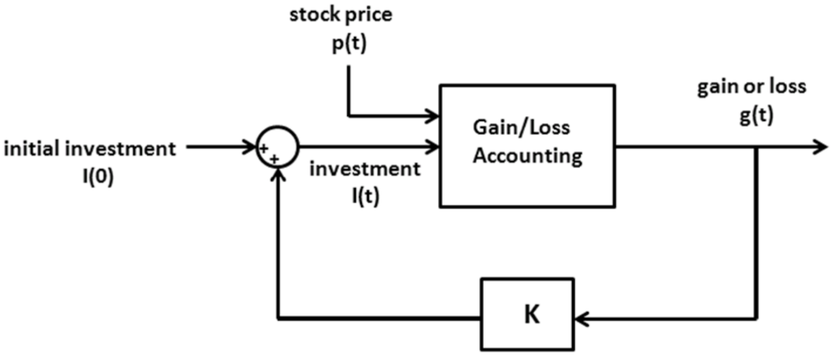
Brownian motion sample paths for price

Benchmarking Considerations for Control-Based Stock Trading Strategies

Analogous to performance evaluation standards used for algorithms in nonlinear optimization, stock trading strategies should "prove their worth" via back-tests using a sufficiently diverse group of benchmark price classes. Such classes include geometric Brownian motion (shown left), historical data for both stocks and sectors, and various other classes of stochastic processes. From a practical perspective, a theory, however elegant, is only as good as its back-tested performance. To this end, a focal point of ongoing research is the development of new benchmark stock price classes to serve as "proving grounds" for technically based control strategies in conjunction with several metrics for evaluating risk and return.

***Contributors:** James A. Primbs, California State University, Fullerton, USA, and B. Ross Barmish, University of Wisconsin-Madison, USA*

Considerations Particular to Feedback-Based Stock Trading



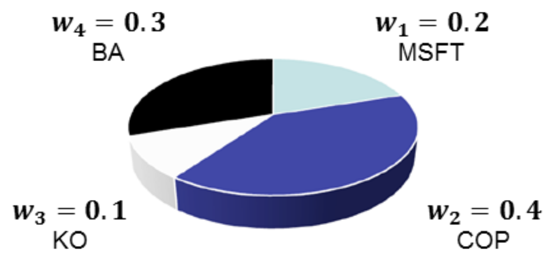
Reformulation of stock trading as a control system with feedback K and input p(t)

Modeling Technical Analysis via Feedback

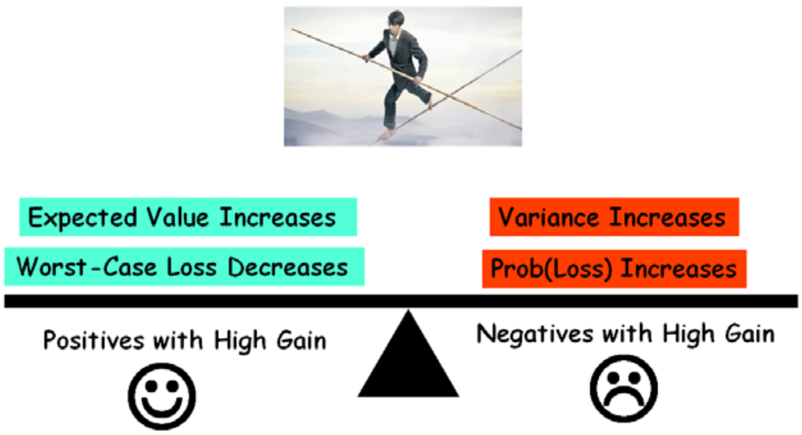
Many popular technical analysis strategies are modeled naturally as feedback loops. For example, the key characteristics of many pyramid-style trend-following strategies, as espoused by legendary traders such as Jesse Livermore and William O'Neil, can be captured via a model involving a simple proportional feedback mapping on cumulative trading gains and losses. The block diagram at left with a feedback control gain K corresponds to such a strategy and allows researchers to carry out a rigorous mathematical analysis of various performance-related quantities such as profits and losses. In this setting, the terminology "control specific risk" is used when studying the effects of K on performance.

Open Research Problem: Portfolio Balancing and Feedback-Based Technical Analysis

A *portfolio* is a collection of assets consisting of stocks, bonds, and cash. Modern portfolio theory deals largely with how to allocate wealth among these assets. The figure at right depicts this process for a portfolio consisting of Microsoft (MSFT), Conoco Phillips (COP), Coca-Cola (KO), and Boeing (BA). The *i*-th weight indicates the percentage of the wealth allocated to the *i*-th asset. Over time, portfolios are typically rebalanced based on market conditions; that is, as time evolves, various criteria are used to change the portfolio weights. An open research problem involves the development of new model-free technical analysis methods aimed at rebalancing a portfolio based on feedback control.



Example of portfolio weights



Tuning feedback gain K is a balancing act.

Challenges in Selection of Feedback Gains

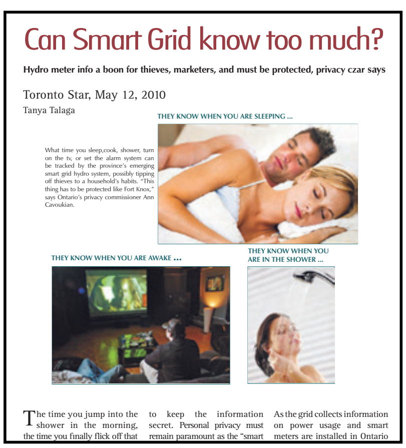
Strategies based on feedback exhibit highly non-Gaussian return characteristics. Thus, the use of traditional measures of performance and risk are rendered unsuitable. Choosing appropriate feedback gains in a stock trading setting is a balancing act involving different aspects of the profit and loss distribution. One particular challenge is to formulate and solve new classes of optimization problems with many more parameters than the simple illustration at top left. In a portfolio context, there may be many stocks, each with its own feedback gain. In addition, the formulation and analysis of new optimization problems should include consideration of control-specific risk as described above.

Preserving Privacy in Cyberphysical Systems

The operation of emerging large-scale monitoring and control systems, such as intelligent transportation systems or smart grids, relies on information continuously provided by and about their users. The result can be an undesirable loss of privacy for the participants, which could delay or compromise adoption of these new technologies, thereby putting their promised benefits at risk.

Disclosure limitation has long been recognized as an important issue in the analysis of statistical databases. More recently, the need has arisen for new theories and tools that can protect the individuals around whom sensor networks and other smart information sources are being built for purposes of collecting dynamic data.

The systems in question produce public aggregate signals from their users' data—for example, the average velocity on a road segment. The main challenge then is to ensure that individual information cannot be inferred from the released signals. Systems and control science provides fundamental insights into such problems, most crucially on how to rigorously establish tradeoffs between achievable system performance and privacy.



Can Smart Grid know too much?

Hydro meter info a boon for thieves, marketers, and must be protected, privacy czar says

Toronto Star, May 12, 2010

Tanya Talaga

THEY KNOW WHEN YOU ARE SLEEPING ...

What time you sleep,cook, shower, turn on the tv, or set the alarm system can be tracked by the province's emerging smart grid hydro system, possibly tipping off thieves to a household's habits. "This thing has to be protected like Fort Knox," says Ontario's privacy commissioner Ann Cavoukian.

THEY KNOW WHEN YOU ARE AWAKE ...

THEY KNOW WHEN YOU ARE IN THE SHOWER ...

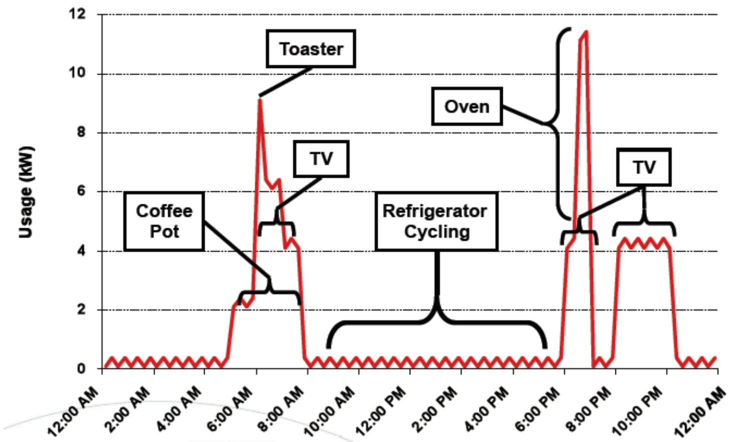
The time you jump into the shower in the morning, the time you finally flick off that to keep the information secret. Personal privacy must remain paramount as the "smart As the grid collects information on power usage and smart meters are installed in Ontario

Source: Toronto Star, May 2010

Definitions of Privacy

Defining a quantitative notion of privacy is a delicate question but is a necessary first step in providing rigorous guarantees. Privacy breaches generally arise from the possibility of linking the information released about a group of people with some additional publicly available information. Hence, simply anonymizing a dataset is usually far from enough to guarantee privacy.

Several formal definitions of privacy have been proposed for studying the tradeoff between the accuracy of the information released and the degree of privacy a given system provides. Information-theoretic definitions have a rigorous foundation but require statistical modeling of the available public information, a difficult task. A popular notion of privacy is k-anonymity, which requires that, from a released output, one cannot distinguish the information of an individual from that of k-1 others. A strong notion of privacy that has also become popular in recent years is differential privacy, a term characterizing certain algorithms releasing randomized outputs whose probability distribution is centered around the desired answer while being insensitive to the presence of any particular individual in the dataset.

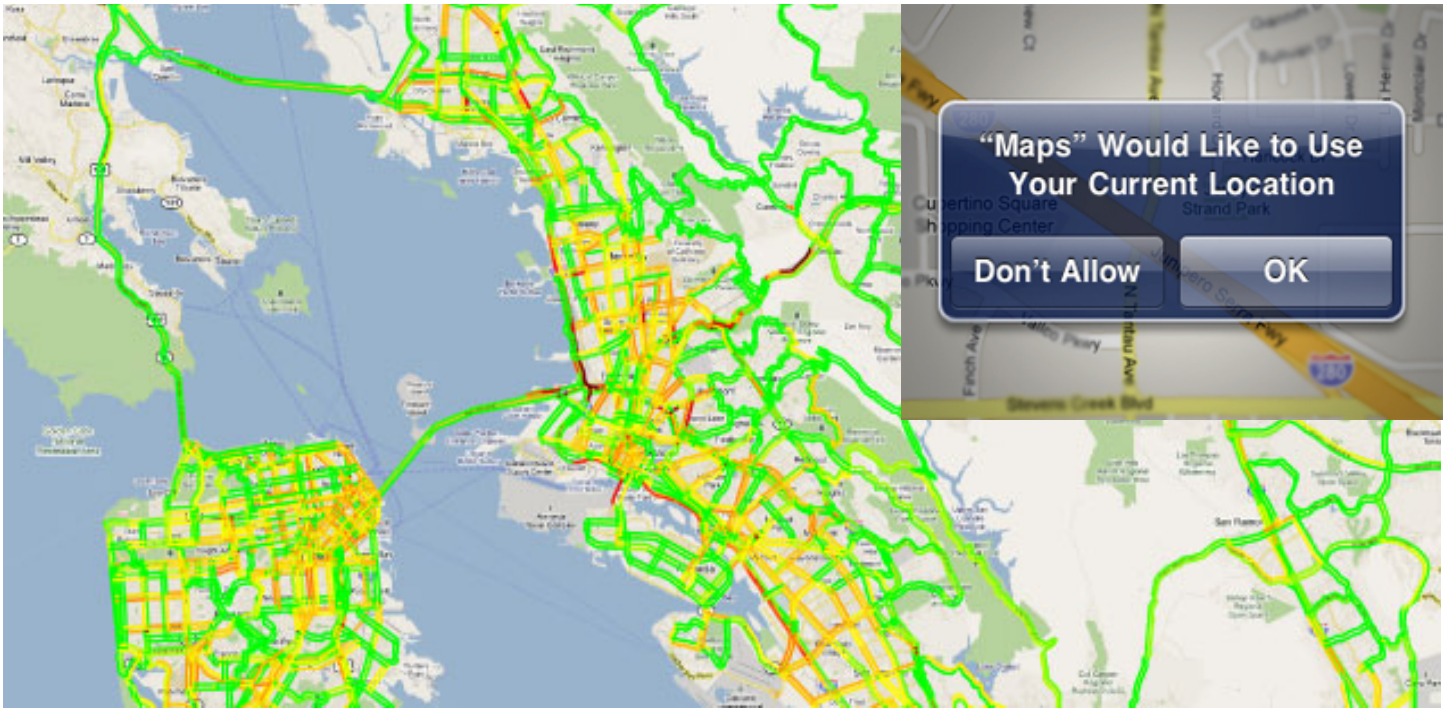


Source: M.J. Hertzler, Xcel Energy

The Smart Grid and Privacy

Privacy concerns related to cyberphysical systems (CPSs) have been targeted in particular at smart meters, halting their installation in some areas. In the 1980s, researchers noted that the appliances being turned on and off in a home can be identified by simply observing the sharp changes in total power consumption recorded by the meter, a technique called nonintrusive appliance load monitoring (see figure above). The danger is that by transmitting this data at frequent intervals, smart meters could be used to monitor the activity of a household. In fact, this data is already being broadcast every 30 seconds by certain automatic meter reading systems currently installed in millions of homes.

Contributors: Jerome Le Ny, Polytechnique Montreal, Canada, and George J. Pappas, University of Pennsylvania, USA



Source: Mobile Millennium, University of California, Berkeley

Privacy Issues in Intelligent Transportation Systems

Emerging traffic monitoring systems fuse data from a variety of sources, including feeds from GPS-enabled smartphones. Anonymity is not enough to guarantee privacy of the traces because most people can be identified just from knowing their two or three most frequently visited places (home, work, etc.). Schemes based on k-anonymity have been proposed to make the user tracking problem more difficult. Similar privacy issues arise with the smartcards used today in many public transportation systems.

Why Is Systems and Control Relevant?

Popular notions of privacy such as k-anonymity require modification before they can be effective in the context of dynamic, real-time systems. Control theory (e.g., optimal estimation) provides important tools for integrating privacy constraints into these systems without sacrificing too much performance. Intuitively, privacy is related to the system-theoretic notion of observability. Moreover, the sensitivity of an output to the data of specific individuals, which is a crucial object of study in the design of differentially private mechanisms, is also related to the standard notion of system gain.

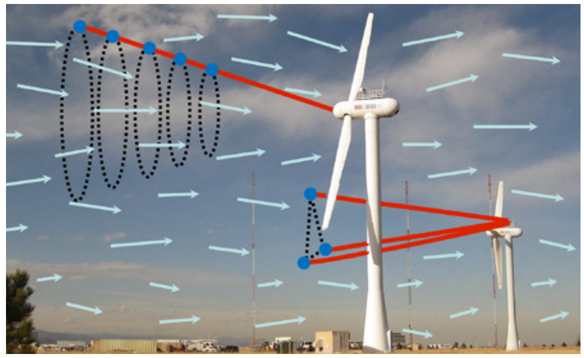
These examples show that control scientists are well equipped to make fundamental contributions to the design of rigorous privacy schemes that can be integrated to the core of many cyberphysical systems. These user protection mechanisms will increase trust and thereby encourage the adoption of these systems, thus indirectly, and perhaps paradoxically, also contributing to improving their efficiency and performance.

For more information: I. Rouf et al., Neighborhood watch: Security and privacy analysis of automatic meter reading systems, Proc. 2012 ACM Conference on Computer and Communication Security; B. Hoh et al., Enhancing privacy and accuracy in probe vehicle based traffic monitoring via virtual trip lines, IEEE Transactions on Mobile Computing, 2012; J. Le Ny and G. J. Pappas, Differentially private filtering, to appear in IEEE Transactions on Automatic Control.

Preview Control of Wind Turbines

Wind is not only the energy source for wind turbines but also the main disturbance to the control system. Thus, knowledge of the incoming wind is valuable for optimizing energy production and reducing structural loads. "Preview control" techniques can incorporate this knowledge.

Advances in lidar technology provide new opportunities for preview control of wind turbines; however, even with state-of-the-art lidar systems, the disturbance cannot be measured perfectly. This requires control research to address two coupled aspects. On one hand, the complex wind field can be reduced to wind characteristics such as speed, direction, or shears, and a control problem can be formulated to address changes in the disturbances.
On the other hand, the performance of the preview controller depends on how well the deduced disturbances correlate with the actual disturbances acting on the turbine. A thorough understanding of the nature of the wind, as well as signal processing and wind measurement principles, are mandatory for developing accurate estimation techniques that enable successful preview control algorithms.



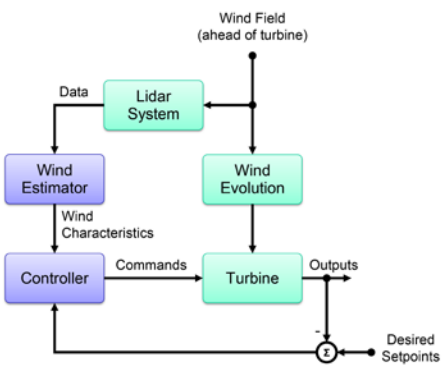
Example lidar configurations: A lidar mounted in the hub of the two-bladed turbine (in the foreground) measures wind speeds by scanning multiple circles in front of the turbine. A nacelle-mounted lidar on the three-bladed turbine (farther away) measures wind speeds by focusing three beams at different locations in front of the wind turbine.

Control Challenges

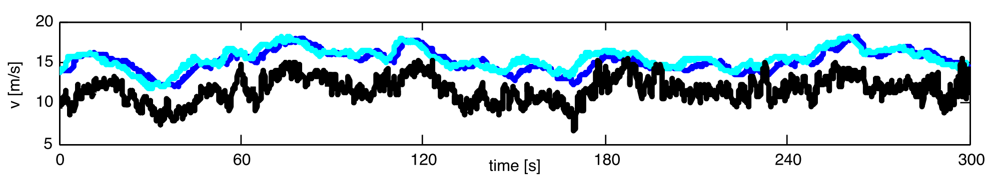
Control strategies need to be robust to the changing information quality of preview wind measurements and the long-term uncertainties present in wind modeling assumptions. The highly flexible turbine structure reacts to disturbances in many ways; thus, baseline controllers must be enhanced by added feedback loops to damp dominant tower and blade modes. Further application of preview control must address multiple complex control goals while coping with changing information quality.

Estimation Challenges

Extracting information from measured lidar data is challenging for several reasons. First, the incoming wind is an evolving three-dimensional vector field. Second, current lidar systems can only detect the speed of aerosols in the line-of-sight direction of the laser beam, which makes exact detection of the three-dimensional wind field impossible. Third, the turbine itself has an (induction-zone) effect on the wind approaching it. In the near term, estimating simplified wind characteristics such as the rotor effective wind speed can be useful for control design. In the long term, estimating more complete incoming wind field information may further improve control performance.



General block diagram of preview control of wind turbines



Rotor effective wind speed estimated from turbine data (dark blue) and preview wind measurements from lidar data (light blue). Single-point wind speed measurements from a nacelle anemometer (black) show reduced speeds due to the induction-zone effect.

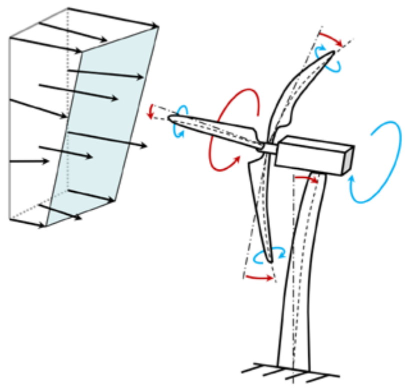
Contributors: David Schlipf, University of Stuttgart, Germany, and Lucy Pao, University of Colorado Boulder, USA

Controller: Current State of the Art and Initial Results

Traditional feedback controllers for wind turbines consist of several single-input, single-output (SISO) loops. They are disturbed by diverse wind characteristics, and wind preview measurements can be used to improve performance. In high wind speeds, preview control can assist the blade pitch controller in regulating the rotor speed as well as mitigating structural loads. This can lead to lower operational and manufacturing costs. Basic feedforward controllers have already been tested successfully on real wind turbines, and more advanced "optimal" controllers such as model predictive controllers (MPCs) have been shown to achieve higher load reductions in simulations.

To reduce the structural loads introduced by the inhomogeneous wind over the rotor disk, the blades are often pitched individually based on blade bending measurements. Several feedforward and MPC-based controllers show further load reductions if the blade-effective wind speeds or the wind shears are known in advance.

Although optimizing energy production has also been addressed, assisting the generator torque controller in low wind speeds is challenging due to its nonlinearity, and only minimal improvements have been achieved with preview control. More promising are approaches to prevent shutdowns due to overspeeding and to use wind direction preview to improve yaw control.

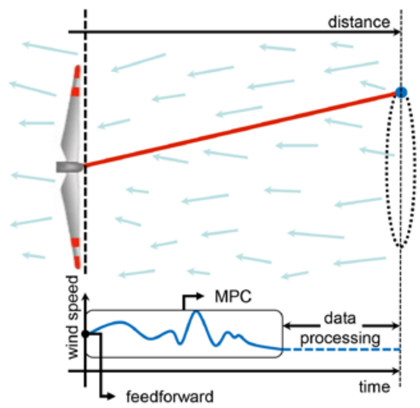


The wind vector field is reduced to rotor-effective wind speed and horizontal and vertical shears. Main turbine modes (red): rotor motion and tower and blade bending. Control inputs (blue): blade pitch angles and generator torque.

Estimator: Current State of the Art and Initial Results

Wind characteristics (e.g., wind speed) are usually predicted from the raw data measured in front of the turbine using reduced-order models of the wind and lidar system. Initial preview controllers have generally assumed that the turbulence in the flow is unchanged as the wind approaches the turbine at the mean speed, the estimate of which is continuously adjusted.

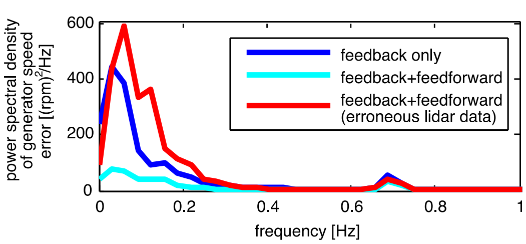
Comparing the wind prediction to the delayed disturbance estimate from turbine data and models reveals that only low-frequency measurement data are correlated to the wind sensed by the turbine. This has been confirmed by theoretical correlation models, which can also be used to optimize the lidar scan configuration and to design optimal prefilters. Filtering is necessary to avoid harmful control action due to measurements at uncorrelated frequencies.



Lidar measurements at a fixed preview distance are used to predict the incoming wind. Some feedforward methods use only the expected wind speed at the rotor plane, whereas more complex methods such as MPC use the entire preview.

Next Steps/Continuing Challenges

Initial field test results of lidar-based collective-pitch feedforward control show significant improvement in speed regulation when wind speeds are estimated well. However, in the case of low correlation between the preview wind measurements and the wind speeds that affect the turbine, the feedforward controller has a negative impact. Therefore, adaptive elements in estimator and control design are crucial to improve preview control of wind turbines reliably and will pave the way to applying advanced multivariable controllers in the nonlinear transition between partial and full power production, where most of the structural loads occur. Furthermore, developing appropriate dynamic wind models and advanced estimators is necessary for successful individual blade pitch preview controllers on real turbines.



Field test results: Generator speed regulation can be improved significantly with feedforward control if there are "good" preview wind speed measurements.

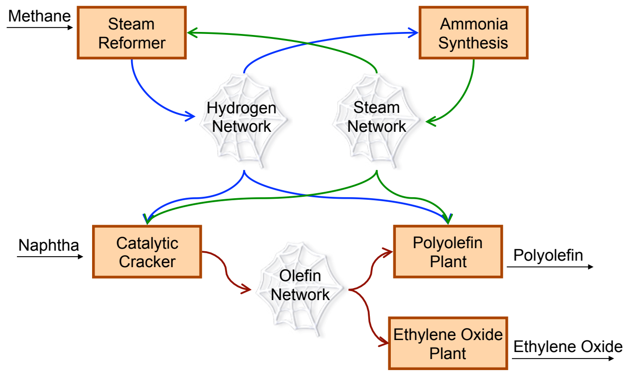
For more information, visit www.nrel.gov/wind, www.uni-stuttgart.de/windenergie, www.vindenergi.dtu.dk, www.cener.com/en/wind-energy, www.ecn.nl/expertises/wind-energy, www.windenergy.citg.tudelft.nl, ecee.colorado.edu/~pao/research/wind.html.

Process Manufacturing Networks

Growing competition on the global scale, the transition from supply-driven to demand-driven markets, and the tightening of process safety and environmental regulations are all placing increasing pressure on process manufacturing and operations. Leveraging the full economic potential of a process plant while maintaining a high level of sustainability requires the supply-chain-conscious optimization of plant operations in real time. Optimal plant operation must accommodate both the interactions with other plants in the associated supply chain and the dynamics of raw material, energy supply, and product demand. In particular, availability and prices of raw materials and energy may change quickly in global markets. Furthermore, production is faced with an increasing diversity of product types and grades.

A Manufacturing Site as a Network

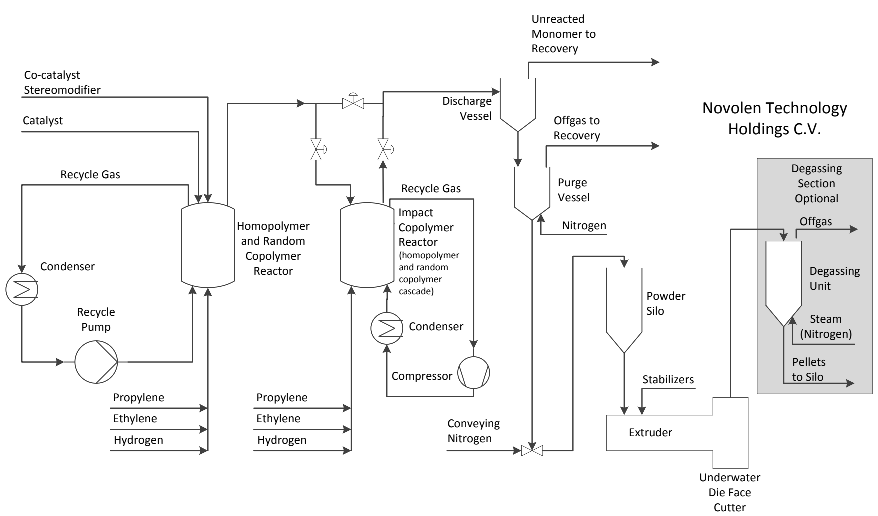
The figure below shows a network of chemical plants on a production site. The nodes refer to plants such as a polyolefin plant, a catalytic cracker, and a steam reformer. The network linking these plants is structured into a steam network, a hydrogen network, and an olefin network. For example, the catalytic cracker node is a consumer of hydrogen and steam and a producer of olefin.



Integrated Production

The interconnection between different process plants and between the units of a single plant account for efficient energy integration and for a largely complete recycling of materials. BASF's Verbund concept, for example, implements a tight integration of all chemical plants at one site.

Such an integrated production site can be visualized by a set of nodes, each representing a chemical plant, connected to diverse networks. Each node of such a network forms a complex network itself, comprising units, sensors, controllers, and actuators, along with their material and information connections.



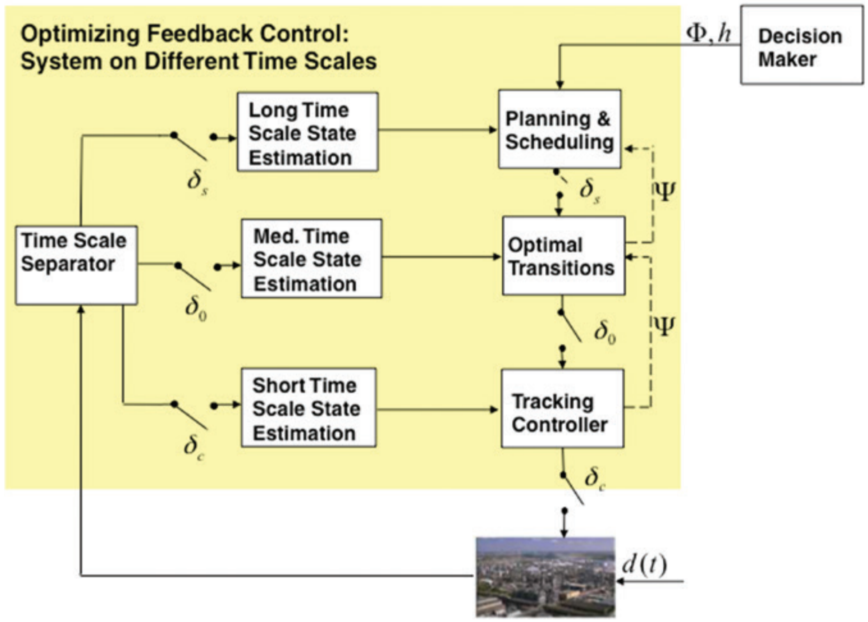
A Plant as a Network

At left is an example flow sheet for the polyolefin process. This complex network of a single plant is embedded into the network of plants on a production site. Altogether, a large-scale and strongly nonlinear hierarchical network control problem is formed, typically characterized by widely varying time scales, discrete-continuous dynamics, and a large number of controlled and manipulated variables.

Contributors: Wolfgang Marquardt and Kathrin Frankl, RWTH Aachen University, Germany

Challenges: Dynamic, Real-Time Optimization and Control

Optimization and control algorithms not only have to treat extremely large-scale, nonlinear, and nonconvex optimal control problems with widely varying time scales and long control horizons, but they also have to cope with discrete decisions to adjust the control strategy. Such algorithms must exploit the structure of the problem, which stems from the hierarchical nature of the network and the model structure of the individual units. Decomposition strategies are essential, but they must take into account the strong interactions between the units of the plant network and between the plants in the site network.

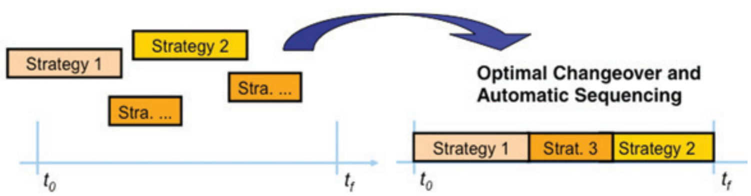


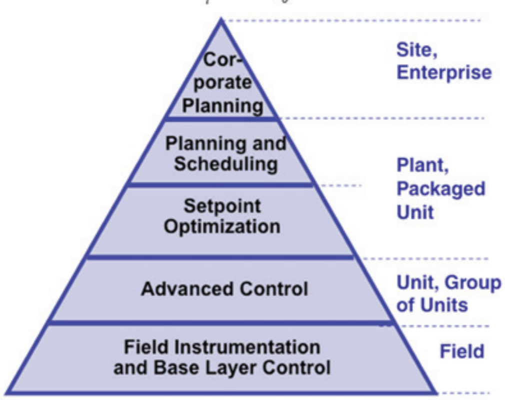
Challenges: Planning and Scheduling

Real-time business decisions relate to operational strategies such as the startup or shutdown of a unit in a plant, the production schedule of the diverse product types and grades, and the transitions between the resulting campaigns. The control and optimization schemes have to be robustly feasible and optimal despite the unavoidable uncertainty in the availability and prices of energy and raw materials, the prediction of time-varying demand of the different product types, and the usual disturbances.

Challenges: Modeling

Given the complexity of an integrated site, modeling represents the major challenge and bottleneck for the rollout of model-based control and real-time optimization techniques. The acquisition of process knowledge, casting it into hybrid first-principles/data-driven models, adjusting the models to the real plant, managing the unavoidable model uncertainty, and maintaining these models over time constitute the major challenges, not only from a technological but also from an organizational perspective. Obviously, modeling and the representational formalisms have to account for the functional separation in the different layers and their interrelations in the network hierarchy.



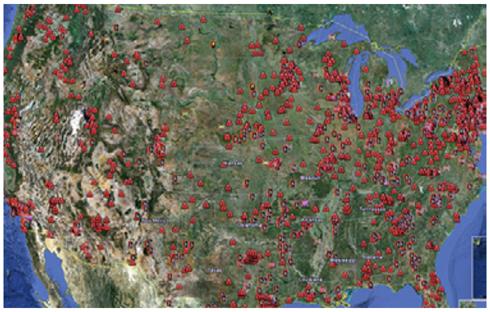


And There's Much More

Solutions to any of the challenges posed for the development of methodologies and algorithms for optimal operation of chemical process systems may be applied to any other hierarchical network problem. Prominent examples include freshwater supply or wastewater networks, gas distribution networks, and electric power networks, to mention only a few.

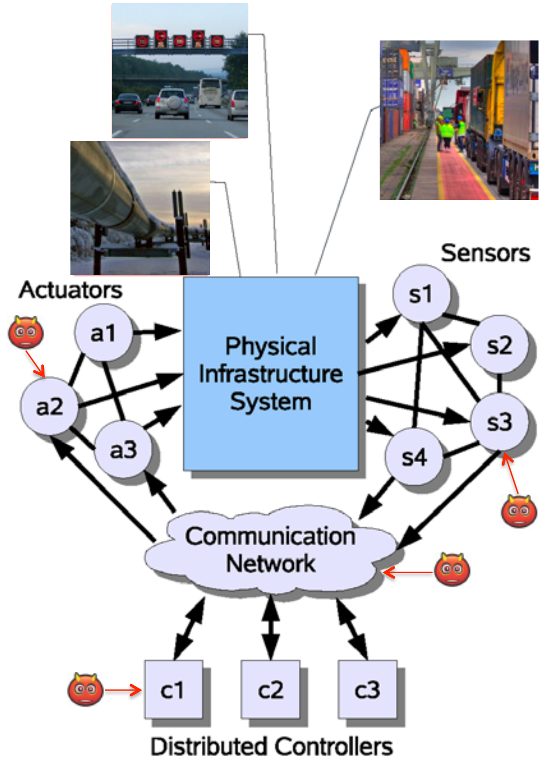
Resilient Cyberphysical Systems

Cyberphysical systems (CPSs) are being increasingly deployed in critical infrastructures such as electric power, water, transportation, and other networks. These deployments are facilitating real-time monitoring and closed-loop control by exploiting advances in wireless sensor-actuator networks, the Internet of Everything, data-driven analytics, and machine-to-machine interfaces. CPS operations depend on the synergy of computational and physical components. In addition, in many cases, CPSs also interact with human decision makers. Fundamentally, once we admit that CPS operations depend on actions of humans (albeit to different degrees), we also have to admit that malicious entities could take charge of CPS control by exploiting cyber insecurities or physical faults, or their combination. Therefore, to improve CPS resilience, we need diagnostic tools and automatic control algorithms that ensure survivability in the presence of both security attacks and random faults and include models of the incentives of human decision makers in the design process.



In December 2012, the U.S. Department of Homeland Security released a map of approximately 7,200 control system devices that appear to be directly linked to the Internet and are vulnerable to attack.

Cyberphysical Systems



Cyber attacks on CPS control elements can result in the loss of availability (denial-of-service attack) and/or integrity (deception attack) of safety-critical measurement and control data.

Cyber Vulnerabilities in CPS

Recent incidents (e.g., the Stuxnet attack) confirm that control systems for critical infrastructures are targets of highly motivated teams of attackers with access to ample financial and technical resources.

Cyber vulnerabilities arise in CPSs due to:

- Wider deployment of off-the-shelf information technology (IT) devices. CPSs inherit the vulnerabilities of these devices and thus are subject to software bugs and hardware failures.
- Replacement of proprietary protocols and closed networks with standard open Internet protocols and shared networks. Malicious attackers capable of exploiting protocol and network insecurities can target CPS operations.
- Generation, use, and modification of CPS data by multiple parties. This poses new challenges in access control and authorization among the strategic players, such as the operators, the IT vendors, and the end users.
- The presence of a large number of remotely accessible field devices. Thus, sensor-control data becomes prone to adversarial manipulation.

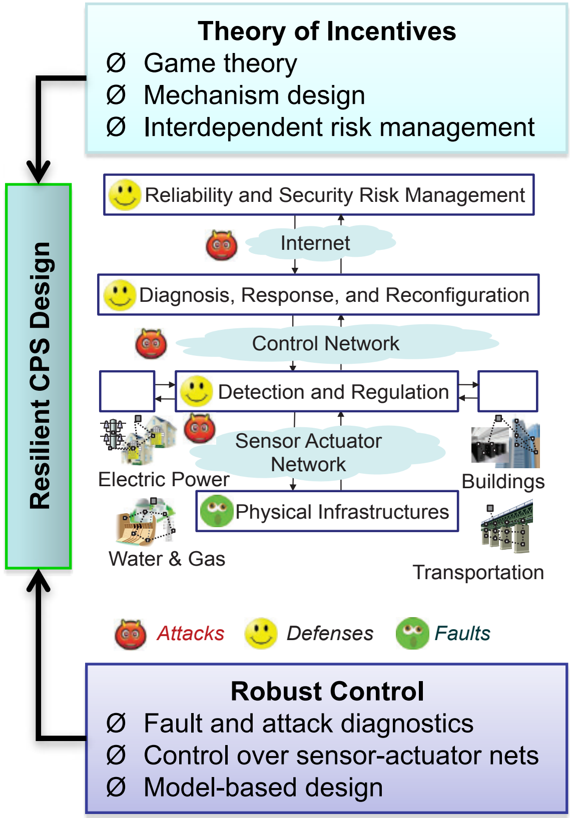
Contributors: Saurabh Amin, MIT, USA; Galina Schwartz and S. Shankar Sastry, University of California, Berkeley, USA

Control and Incentive Tools for Resilient CPS Operation

Resilient operation of CPSs requires the following high-confidence attributes: functional correctness (by design) for real-time operations, robustness to reliability failures (fault tolerance), and survivability even during successful attacks (operation through attacks). Designers and operators of CPSs currently lack comprehensive tools for resilient operation. Major challenges include: (1) spatiotemporal and hybrid dynamics of cyberphysical processes; (2) a large number of interactions with interdependencies; and (3) effects of public and private uncertainties. Notably, two distinct domains of tools have emerged to respond to these challenges:

- Robust control over networks: These tools primarily address safety and performance issues in closed-loop control over sensor-actuator networks.
- Theory of incentives: These tools provide ways to analyze and influence the strategic interactions of human decision makers.

To date, control and incentive tools have been designed and implemented separately. This separation was natural due to the lack of advanced CPS technologies in legacy supervisory control and data acquisition (SCADA) systems. Modern CPSs no longer permit such separation of control and incentive tools. The failure of loosely coupled tools in ensuring resilient operation of CPSs is evident in chronically unresolved design conflicts between efficiency and robustness against faults and attacks and the lack of proper incentive structures to enable private entities (or players) that operate the CPSs to maintain resilience. Consequently, control and incentive tools designed in isolation, or without cognizance of strategic interactions between private entities and interdependent processes in CPSs, are inadequate to maintain resilience.



The challenge of resilient CPS design: *The development of an integrated resilient design methodology necessitates a rigorous analytical framework to allow the co-design of control and incentive tools. This framework will enable designers and operators to build resilience into CPSs by maintaining synergistic integrations of human-centric elements with automated diagnostic and control processes.*

Emerging CPSs for Transportation and Electricity Infrastructures	Tools Based on Robust Control Theory	Tools Based on the Theory of Incentives
Active road traffic management	Distributed sensing and control	Congestion pricing and incentives
NextGen air traffic operations	Robust scheduling and routing	Strategic resource reallocation
Smart electricity transmission	Wide-area monitoring	Contractual design
Power markets, including nondispatchable generators	Risk-limiting dispatch	Market design
Smart electricity distribution	Distributed load control	Demand response schemes
Energy-efficient building operations	Predictive control of devices	Energy-saving incentives

For more information, visit http://ics-cert.us-cert.gov, http://www.gao.gov/products/GAO-12-926T, http://www.nist.gov/itl/cyberframework.cfm, http://www.nist.gov/el/smartgrid/cybersg.cfm, http://www.whitehouse.gov/the-press-office/2013/02/12/executive-order-improving-critical-infrastructure-cybersecurity, and http://en.wikipedia.org/wiki/Stuxnet.

Stair-Climbing Assistive Robots

A major drawback of assistive robotic devices for the mobility-impaired, such as wheelchairs or the Segway, is their inability to negotiate stairs and steps. Wheel-based systems are advantageous compared to other motion mechanisms because of their energy efficiency, relatively high velocity, and simple design, but enabling them to climb or descend stairs without manual support is an outstanding challenge for control engineers.

A single step as an obstacle for the wheelchair driver (Source: www.bsk-ev.org)

Manually overcoming stairs (Source: M.J. Lawn, Study of stair-climbing assistive mechanisms for the disabled, PhD thesis, Nagasaki Univ., 2002)

Stair-Climbing Wheelchairs

The iBOT wheelchair provides some degree of automated support for wheelchair drivers to overcome stairs. Shifting the center of gravity triggers the rotation of the iBOT's lower body. The wheelchair supports the driver by providing power during climbing; however, the wheelchair is not actively stabilized and the climbing cannot proceed autonomously. This limits the stair-climbing function for drivers with a high degree of disability or in cases where no handrails are present.

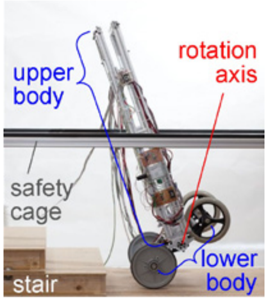
Stair climbing with the iBOT wheelchair (Source: D. Ding, R.A. Cooper, "Electric-powered wheelchairs—A review of current technology and insight into future directions," IEEE Control Systems Magazine, April, 2005)

Autonomous Stair-Climbing Device (SCD)

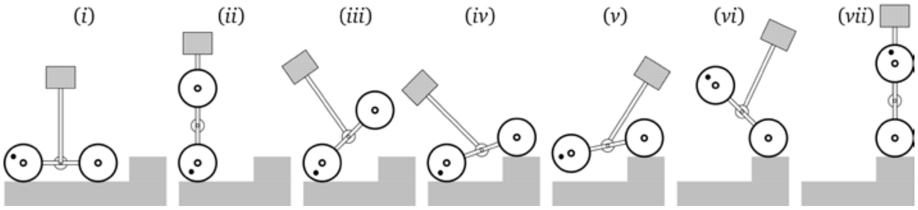
An autonomous stair-climbing device (SCD) could allow the mobility-impaired to manage stairs and steps without manual effort. One possible design is shown here. The mechanical system can be considered a double inverted pendulum. Motors are used to drive the wheels and to control the angle between the upper and lower body. The inclination angle can be estimated with high accuracy with an inclinometer and gyroscope. At least two different discrete states need to be considered:

- All wheels are in ground contact (see figure below, situations (i) and (iv)) and
- Two wheels are in ground contact (situations (ii) and (iii)).

Control action is needed to stabilize equilibrium points of the SCD in both states, enabling it to move on the ground in both states. Furthermore, stable state transitions must be achieved by the control to enable stair climbing.



SCD in balancing mode (Source: www.ims.tu-darmstadt.de)



A sequence of SCD configurations for stair climbing

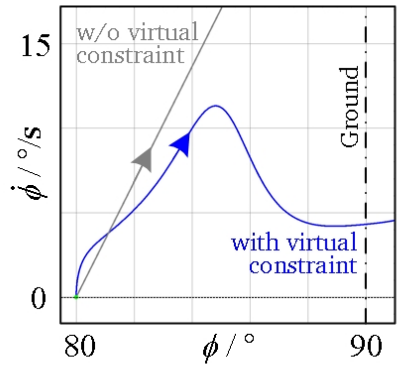
Contributors: Bruno Strah and Stephan Rinderknecht, Technische Universität Darmstadt, Germany

Control-Enabled Assistive Robotics

Control Characteristics

- **Hybrid nonlinear dynamics:** In a specific configuration, the SCD-like systems show nonlinear continuous-time dynamics with nonholonomic constraints due to wheel kinematics. In addition, discrete events such as the transition from two wheels to all wheels on the ground cause a discontinuous impact. Thus, the device exhibits hybrid nonlinear dynamics in its operation.
- **Unilateral constraints:** With the given mechanical structure, the wheel-ground contact exhibits a unilateral (non-negative) contact force, enabling wheel-ground contact activation and deactivation. These transitions should be realized by control actions that satisfy stability and performance criteria (e.g., soft landing). The degrees of freedom change during these transitions.
- **Underactuation:** Although the SCD is fully actuated when all its wheels are on the ground, this is not the case in configurations with two wheels in ground contact. Here the system is underactuated, with one more degree of freedom, namely, the inclination of the lower body.

Despite the wheel-versus-limb difference in motion mechanisms, these characteristics are shared by legged robots.



The transition from two wheels to all wheels in ground contact can be designed to achieve a soft landing using the "virtual constraints" framework. The upper-lower body angle (ρ, see figure below) and the lower body angle (ø) are coupled in the control system output function. The system zero dynamics of the SCD when falling to the ground are such that the upper-body motion causes a deceleration of the lower body as soon as the ground is close.

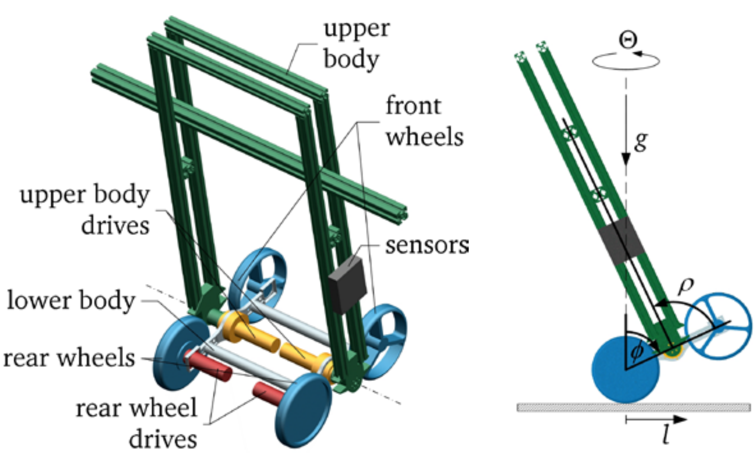
SCD Control Challenges

Stable motion and stable transition of configurations have been ensured using feedback linearization methods. In the underactuated state, the linearization can be achieved partially (input-output); in this case, the remaining internal dynamics have an unstable equilibrium, which can be stabilized with an additional linear full-state controller. A challenge for future research is to apply a full nonlinear control law, which could provide higher speed of motion across a broad amplitude range and disturbance compensation by (preferably) body motion instead of wheel motion.

Although the overall functionality has been achieved by considering motions in particular configurations and transitions, a further control challenge is to induce a limit cycle, causing a permanent lower-body rotation. Such a strategy could increase the stair-climbing velocity. A further advance would be to realize jumping to next steps. This capability could allow stair climbing in cases where the SCD lower-body and wheel geometry do not allow it. Furthermore, this control could mitigate the consequences of falling to the lower step in case of large disturbances.

Applications to Unmanned Service Robots

Stair climbing would be a useful capability not only for wheelchairs and other systems that mobility-impaired people can use for transport, but also for unmanned robots in service applications. Such robots are being developed for applications in homes, offices, museums, factories, and other facilities (www.doublerobotics.com).



CAD model of the SCD (above left) with main degrees of freedom (above right): position l, lower body angle (ø), upper-lower body angle (ρ), yaw angle (Θ), and gravitational acceleration g

Supply Chain as a Control Problem

Today's forces of interest for the supplier, manufacturer, and customer require ever-increasing levels of supply chain agility and inventory management to continuously improve operational efficiency. As these forces interact across the supply chains, further refinement of standards in the areas of sensing, measurement, communication, control, decision policy, organizational structure, practitioner responsibility, and implementation practices are required to move supply chain metrics of interest to new levels of performance and reliability.

Grand Control Challenge

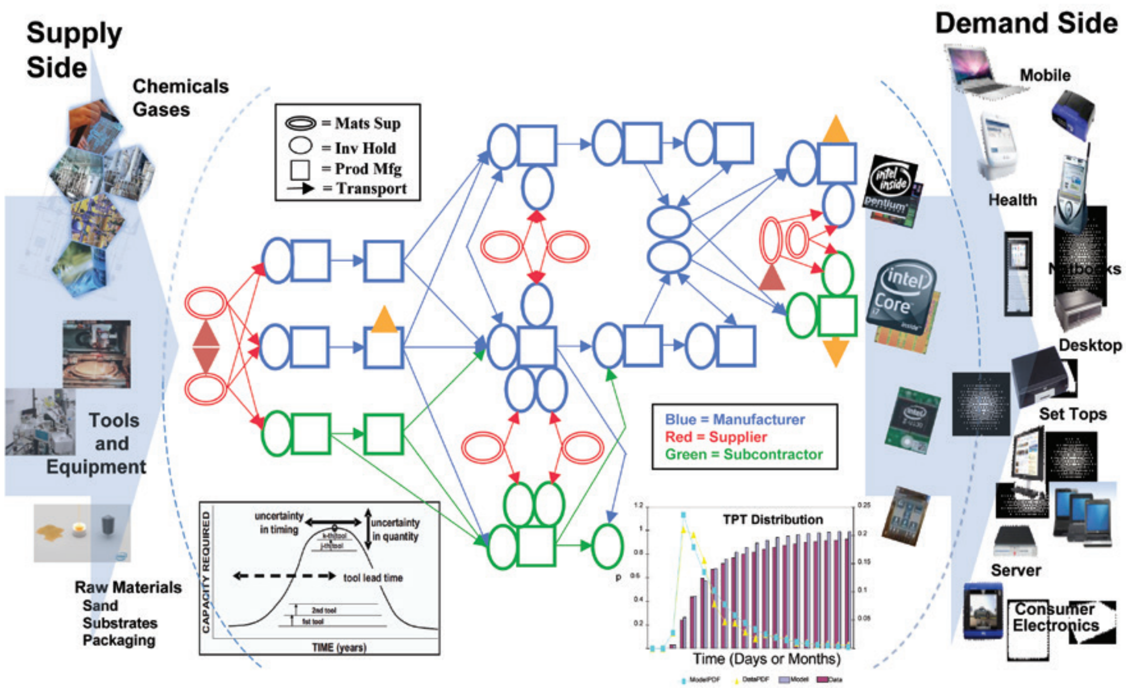
Well-controlled supply chains can deliver

- the right product,
- in the right quantity,
- from the right sources,
- to the right destinations,
- in the right quality/condition,
- at the right time,
- for the right cost

while

- reducing inventories,
- increasing supply chain agility,
- reducing operational cycle time,
- optimizing supply product mix relative to the demand mix, and
- enabling maximum business profitability.

Discrete Supply Chain Example



Contributors: Kirk D. Smith, Martin Braun, and Karl Kempf, Intel, USA; Joseph Lu and Duane Morningred, Honeywell, USA

Nature of Supply Chain

Multiple Ownership: A company's performance in the supply chain is affected by its suppliers, customers, suppliers' suppliers and customers' customers, and its collaboration with them. Each company has a vested interest in all the links in the supply chain, not just those of direct suppliers and customers.

Constant Evolution: Products and equipment in the supply chain may run their complete life cycles as the overall performance is being improved. Fast ramp-up and ramp-down of products, and their accompanying processes and toolsets, pose challenging transition problems.

Uncertain Dynamics: Supply chains are stochastic, nonlinear, and time varying. In addition to transport and throughput times being affected by "simple" logistical and manufacturing systems and related processing loads, they are affected by weather, politics, culture, innovations, contractual relationships, and other complex human interactions.

Risk Management: Common risk management measures, such as safety stock, contingency systems and procedures, customer and supplier agreements, and shipping time allowance, can greatly affect supply chain agility, maintainability, customer satisfaction, and of course, cost.

Present State of the Art

Modeling: Treating a segment of the supply chain as a network of inventories and specialization processes with preconfigured connections and estimated production dynamics.

Control: Using supply and demand forecasts to specify material processing and distribution rates that mitigate inventory control limit violations. Model predictive control (MPC), for example, has been successfully applied to several segments of the supply chain where traditional supply chain solutions have had difficulty (see figure below).

Optimization: Incorporation of economics and business logic to direct material to locations that maximize agility while minimizing unnecessary processing and shipping. Typically, solutions with longer time horizons and a greater model abstraction (such as planning and scheduling) are implemented as supervisory layers above the control layer.

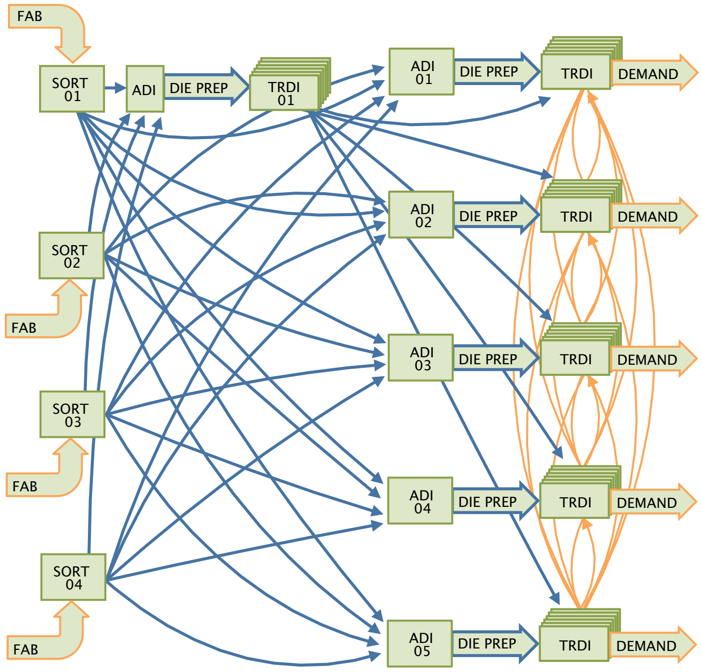
Challenges to be Solved

Crossing Company Boundaries: Incorporate pricing on the supply and demand side while maintaining local autonomy, sharing information, and creating a win-win-win approach among the suppliers, customers, and manufacturer. Moving beyond simple data exchange is deemed essential to executing local optimization in a collaborative manner and to thereby achieving the greater benefits of global optimization.

Demand/Supply Forecast: Accurate forecast models are essential for tight supply chain management. The challenge is how to forecast the supply and demand and quantify and account for uncertainties.

Managing Risks: How should strategic decisions such as system capacity allocations be derived and implemented with a tolerable investment and business risk? What are the optimal uncertainty buffers across the supply, demand, and manufacturing domains? Can options theory, for example, be used to support or even optimize both investment and production decisions?

Supply Chain Cost: How does a specific set of supply chain solutions take into account current business workflow and push the boundary of automating business processes beyond current practices to reduce the overall supply chain cost?



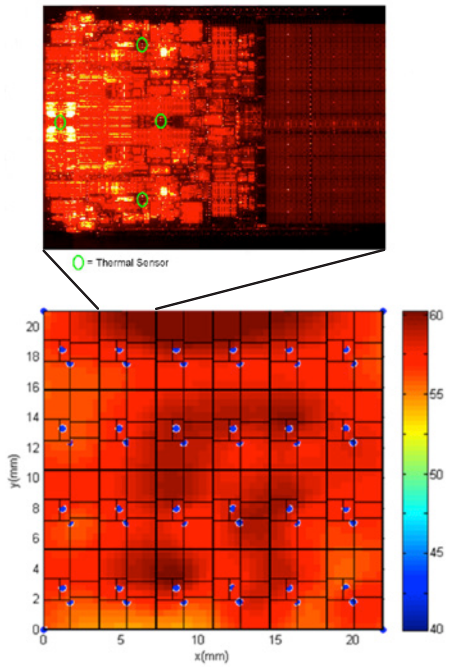
A concept for a model predictive control application for inventory and production management for semiconductor manufacturing. Wafers are sourced from fabrication facilities and go through stages/buffers for sorting, assembly die inventory (ADI), and tape and reel die inventory (TRDI) before finished dies are shipped to meet demand at manufacturing sites. Manipulated variables are shown in blue; feedforward signals in orange. Several such controllers can be coordinated in one application for managing shared inventory space, such as warehouse floor space. Versions of this concept have been implemented with individual controllers manipulating and controlling hundreds of variables each.

Thermal Control of Manycore and Multicore Processors

Today's high-end multicore and manycore CPUs are characterized by extreme power density and peak power consumption. The thermal dissipation systems for these processors are often designed with narrow, or even negative, margins for cost reasons. In addition, unexpected thermal emergencies may arise because of significant spatial and temporal variability of workloads, leading to nonuniform performance, power consumption, and temperature distribution. Hot-spot areas age faster since degradation effects are exponentially accelerated by high temperatures. This in turn can lead to chip damage or failure.

We are in an era of thermally limited computing. Hot-spot and thermal-runaway prevention based solely on worst-case thermal design is now unaffordable. Significant effort is thus being devoted to techniques that dynamically control the core power dissipation in a temperature-aware fashion, i.e., aiming to enforce a safe working temperature across the die surface. Today's multiprocessors include hardware support for dynamic power and thermal management, based on introspective monitors (i.e., per-core thermal/performance sensors and chipwide power gauges) and performance knobs. This infrastructure provides the sensors and the actuators for feedback control policies.

—European Research Council "Multitherman" Project

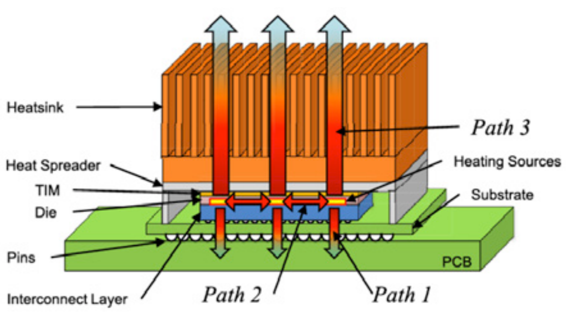


Top : Intel® Dual-Core Merom thermal map (Source: Proc. Int. Solid-State Circuits Conf., 2007); Bottom: Thermal map of the Intel single-chip cloud computer under full utilization (temperature scale in degrees Celsius)

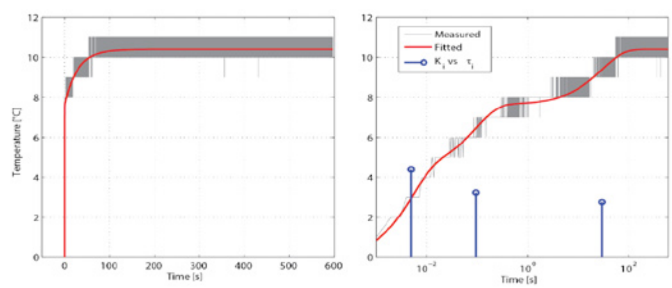
Thermal Control Challenges

Modern electronic devices have billions of transistors clocked at subnanosecond speed. Local on-die thermal transients have time constants of microseconds, whereas at the package and board level we see complex, nonlinear dynamics unfolding in seconds to minutes. A single chip can have hundreds of thermal domains that vary greatly in workload and intrinsic power density. Power consumption and heat generation are thus spatially and temporally heterogeneous, with nonlinear temperature dependency caused by leakage. In addition, the heat dissipation path is composed of different materials that lead to a multimodal time-domain response.

Accurate on-chip temperature sensors have high area cost and are affected by significant systematic and random noise. In addition, to keep post-manufacturing testing costs low, not all the sensors are accurately calibrated. Hence, manycore thermal management is a large-scale, hybrid, nonlinear multivariable control problem, affected by significant sensor, actuator, and process noise.



Typical layered chip package section and main thermal path direction and model



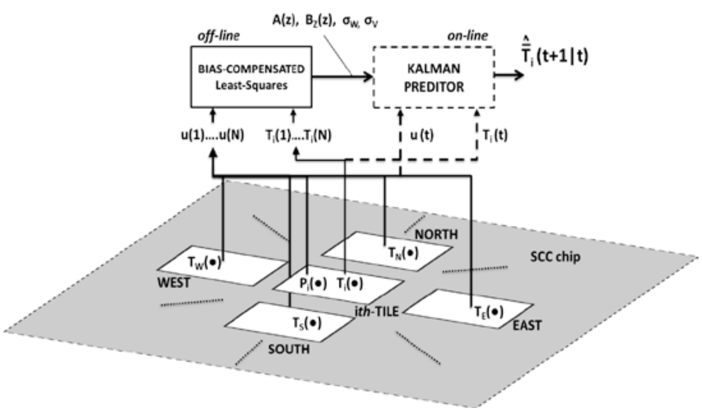
Linear (left) and logarithmic (right) transient analysis plots. The temperature is well approximated by a third-order exponential curve. The stem plot (right) shows the three time constants (4.9 ms, 94.6 ms, and 29.2 s).

Contributors: Andrea Bartolini, Andrea Tilli, and Luca Benini, Università di Bologna, Italy

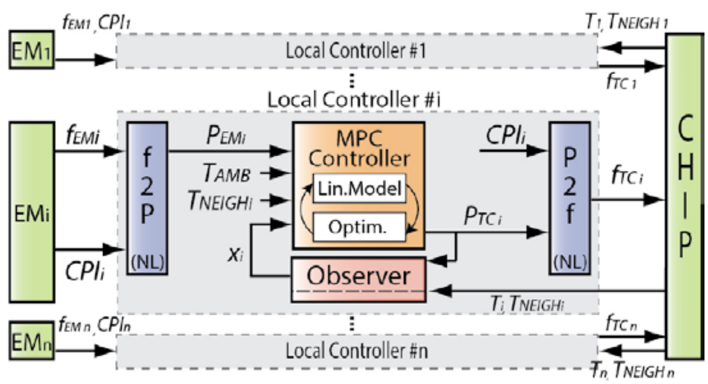
Scalable, Optimal Thermal Control for Manycore Systems

The previously highlighted challenges call for optimal control features:

- *Optimality:* The thermal controller acts (by voltage and clock scaling and shutdown) to reduce power consumption, but it must strive to minimize performance degradation while limiting temperatures to below a safe threshold across the entire silicon die. A model predictive control (MPC) approach can reduce performance degradation with respect to simple threshold-based control.
- *Predictability:* Regular workload phases can be exploited by a thermal model to predict future temperatures. This calls for system-level thermal models that relate different functional units and hardware macro block activity to the thermal map evolution.
- *Adaptability:* Fluctuations of process variations and ambient conditions (temperature, heat sink occlusion, etc.) may change the thermal behavior over the lifetime of a component. Model recalibration strategies and online system identification algorithms are required.
- *Robustness:* Thermal sensor readings are affected by significant output noise. System identification and controller design approaches are needed that are robust to measurement and process noise.
- *Scalability:* The trend toward massively parallel (100+) cores and hardware accelerators integrated on 3-D stacked dies calls for scalable control algorithms running in a few microseconds. Distributed control algorithms are needed that leverage the spatial localization of heat exchange and can exploit parallel hardware.
- *Modularity:* Thermal control not only happens at the hardware level, but it must interact with software layers such as the workload dispatcher and task scheduler.

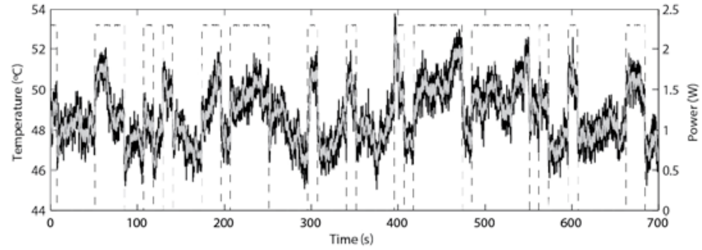


Distributed and robust thermal model learning strategy based on ARX-plus-noise system identification. The model takes as input the core power consumption and neighbors' core temperatures. It estimates the noise variance and the model parameters suitable for the Kalman predictor. The complexity of each single model is constant as the number of cores increases. Inputs to the Kalman predictor are also the current core power consumption and neighbors' core temperatures.



Distributed model predictive thermal controller. Each core executes its local control with temperature information from neighbor cores. Target frequency (f_{EM}) requests are generated by the energy manager (EM). Target power consumption (P_{EM}) is derived from f_{EM} and application properties (the nonlinear f2P function). The MPC exploits the thermal model to find the minimal power reduction that keeps the predicted temperature below a safe threshold. This value (P_{TC}) is then converted to a frequency setting (f_{TC}) through the P2f inverse function.

Model validation: input power (dashed), measured temperature (black), and one-step-ahead predicted temperature (gray). The learned model can be used effectively at runtime to estimate the actual silicon temperature.



For more information, visit http://www-micrel.deis.unibo.it/multitherman/.

Toward Verifiably Correct Control Implementations

Bugs may be introduced into control applications at all levels, starting from the high-level mathematical control laws to the actual machine code, complete with device drivers and multitasking. An important scientific and technical challenge for the controls and real-time software communities is to design analysis methods at these various levels of abstraction, along with verified compilation and synthesis tools.

State of the Art

- The CompCert compiler from INRIA and University of Rennes-I compiles C to a variety of popular targets (PowerPC, ARM, x86). The compiler has been proven correct mathematically with a machine-checkable proof. The assembly programs produced thus provably preserve the semantics of the source C code.
- The Astrée analyzer can verify many control system implementations in C if they are fairly static—excluding parallelism, dynamic scheduling, dynamic data structures, virtual methods, etc.

Safer, More Powerful Compilation

Compilers are software and as such may contain bugs. A bug in a compiler may result in the introduction of bugs in the object code the compiler generates, and thus in the program as it is executed in the embedded systems. Such bugs may be difficult to find, and thus for certain safety-critical systems, object code must be matched to source code for inspection, ruling out code optimization. However, disabling optimization leads to inefficient object code, requiring higher CPU performance or limitations in functionality.

Although progress has been made in safe compilation for programs written in C, an outstanding challenge remains for compilers for high-level specifications such as Simulink—a preferred formalism for many control systems—or complex languages such as C++.

In turn, this implies that the high-level specification language should have reasonable and unambiguous semantics.

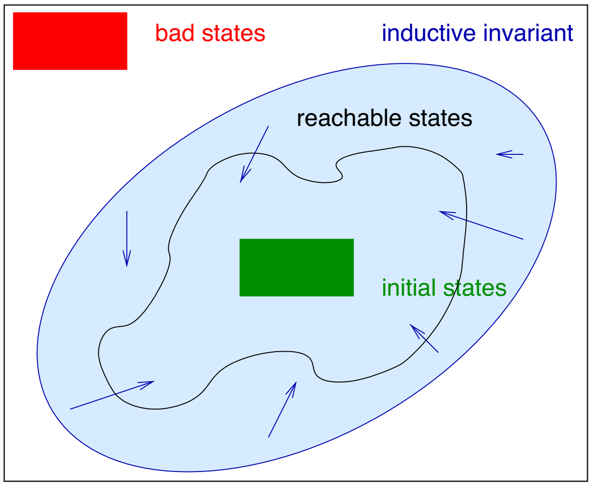
Enlarging the Scope of Static Program Analysis

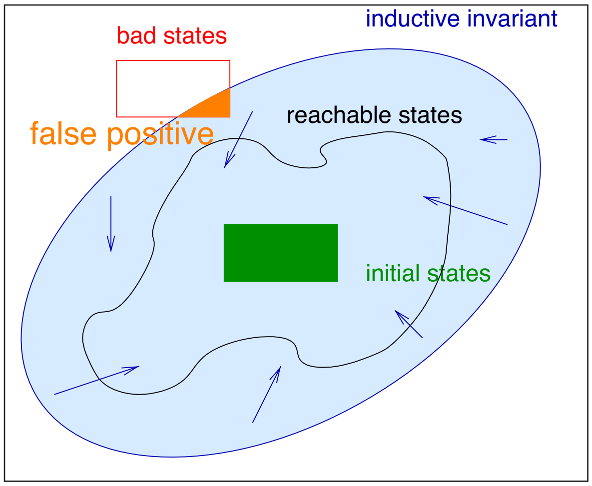
Static program analysis refers to the automated analysis of computer programs without actually executing the programs. Despite the recent availability of industrial-strength program analysis tools, considerable challenges remain.

- The spectrum of applications needs to be increased. Fewer restrictions on programming styles and technologies should be imposed while keeping the likelihood of false alarms low.
- As with compilers, analysis tool implementations should be formally proven correct with machine-checkable proofs. This is especially important if, for critical systems, some testing is replaced by static analysis.
- Speed and automation need to be enhanced. Tools should be able to prove desired properties with minimal user intervention and to provide counterexamples in case properties are not verified.

During the Ariane 5 rocket's maiden flight in 1996, flight control software malfunctioned and the rocket had to be destroyed by remote command early in its trajectory. The malfunction was the result of improper reuse of Ariane 4 software. Hardware redundancy was no help since both computers ran the same incorrect software.

Contributor: David Monniaux, VERIMAG, CNRS, France





A static software analyzer finds an inductive invariant, which includes but may overapproximate the initial states and all possible reachable states. If this invariant excludes bad states (as in the left graphic), the analyzer proves the absence of errors in any possible execution of the software. If the bad states intersect the inductive invariant but not the reachable states (right graphic), a false positive results.

Toward a Trusted Development Chain

High-level specification typically considers idealized mathematical computations. In reality, differential equations are discretized—for example, real numbers are implemented using a floating-point or fixed-point arithmetic; multiple clock domains may be used; mathematical functions may be approximated; and programs are split among different tasks or machines, which may not be in perfect synchronization.

Some of these transformations are automated, but many are still performed by hand, most of the time with no mathematical proof of their correctness. Tools are needed that automate these transformations or at least provide meaningful feedback to implementers.

Control applications increasingly run on multicore processors, including for critical embedded systems. Manual programming for parallel systems is notoriously error-prone. Shared memory implementations require careful placement of locking mechanisms—too few of them and *data races* may occur, but too many of them and *deadlocks* may freeze the system. Automated synthesis or verification of the parallel or distributed implementation, possibly with a formal proof of correctness, becomes increasingly desirable.

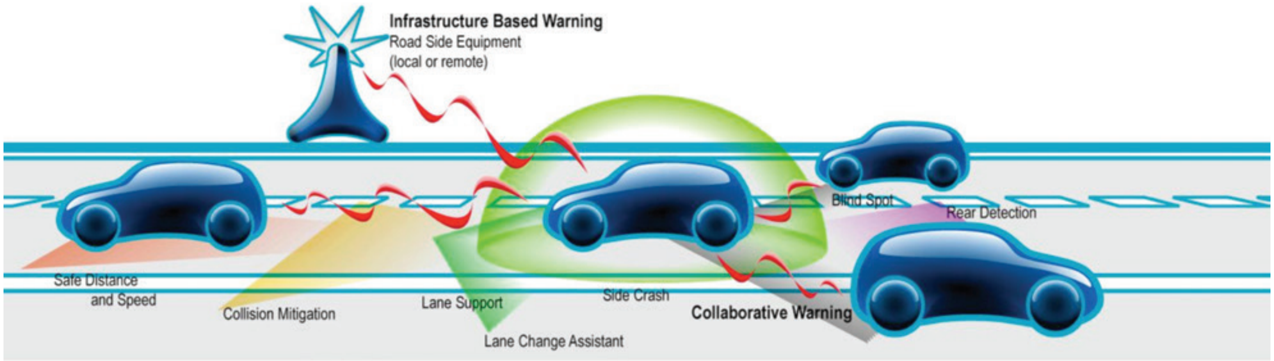
Communication protocols, especially on modern buses, are hard to get right; this is even truer when security properties are involved (e.g., resistance to eavesdropping or intrusion). Implementations of such protocols should be based on reusable, well-tested, or even formally proven libraries, and nonreusable parts should be synthesized from specifications. Doing this effectively and safely remains a research challenge.

```
int main() {
  int x = 0;
  int y = 0;

  while (1) {
   /* invariant:
   102 + -y + -x >= 0
   -y + x >= 0
   y >= 0
   */
   if (x <= 50) y++;
   else y--;

   if (y < 0) break;
   x++;
  }
}
```

Static analyzers, in this case the experimental tool Pagal, may display loop and function invariants. This helps developers understand what is going on in their software so it can be debugged more efficiently.

For more information, see the companion flyer on "Verification of Control System Software" in the Success Stories section of this volume.

Vehicle-to-Vehicle/Vehicle-to-Infrastructure Control



Source: SAFESPOT Project funded by the European Commission

Problems related to the single, isolated automotive vehicle and its subsystems are challenging enough, but the research community is also exploring the "big picture" of intelligent road transportation—the system, or system of systems, consisting of many vehicles and their drivers interacting on roads. Two related topics are included in this vision:

- Vehicle-to-infrastructure (V2I) interaction
- Vehicle-to-vehicle (V2V) interaction

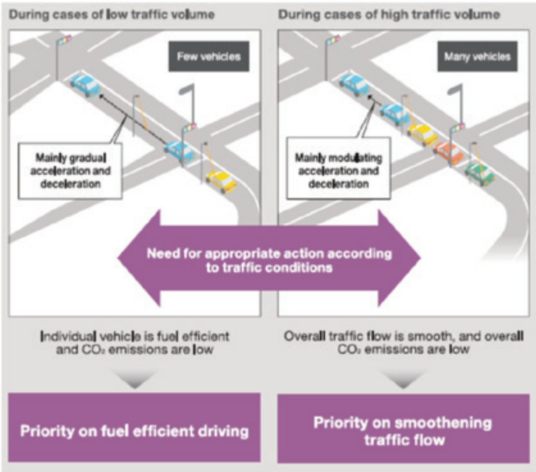
V2I and V2V promise revolutionary improvements in transportation—greater energy efficiency, less road construction, reduced collisions, and safety of vehicle occupants as well as pedestrians and bicyclists. Control is a key contributing discipline for both topics.

Vehicle-to-Infrastructure Control

In V2I, the infrastructure plays a coordination role by gathering global or local information on traffic and road conditions and then suggesting or imposing certain behaviors on a group of vehicles. One example is ramp metering, already widely used, which requires limited sensors and actuators (measurement of traffic density on a highway and traffic lights on ramps).

In a more sophisticated scenario, the velocities and accelerations of vehicles and intervehicle distances would be suggested by the infrastructure on the basis of traffic conditions, with the goal of optimizing overall emissions, fuel consumption, and traffic velocities. Suggestions to vehicles could be broadcast to drivers via road displays or directly to vehicles via wireless connections. Looking further ahead, in some cases suggestions could be integrated into the vehicle controls and implemented semiautomatically (always taking into account the restrictions on automatic vehicle driving imposed by the Vienna Convention on Road Traffic, discussed on the next page).

Some experts predict that the first V2I systems may be developed and deployed in the 2015–2020 time frame.



Source: Toyota USA

The figure at left shows two different traffic situations. In the left panel, traffic density is low and the central infrastructure-based controller acts to improve fuel efficiency and reduce emissions of individual vehicles, smoothing accelerations and decelerations; in the right panel, due to greater congestion, the infrastructure control is primarily concerned with depleting queues at intersections with an eye toward global fuel economy and emissions reduction.

Contributor: Luigi Glielmo, Università del Sannio, Italy

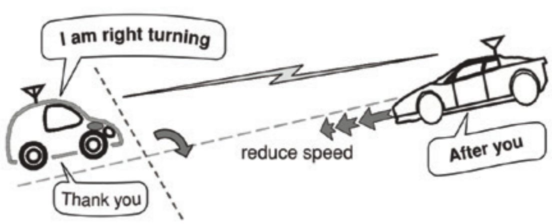
The Vienna Convention on Road Traffic

This international treaty, designed to facilitate international road traffic and increase road safety, was agreed upon at the United Nations Economic and Social Council's Conference on Road Traffic in 1968 and came into force on May 21, 1977 (http://www.unece.org/trans/conventn/crt1968e.pdf). The convention states that "Every driver shall at all times be able to control his vehicle," which conflicts somewhat with the automatic control concept. Systems such as anti-lock braking systems or electronic stability programs are acceptable because they do not take full control of the vehicle but rather help the driver to follow a desired path, possibly in situations where control of the vehicle has already been lost. Wider use of technological advances, however, will require amendment of the convention.

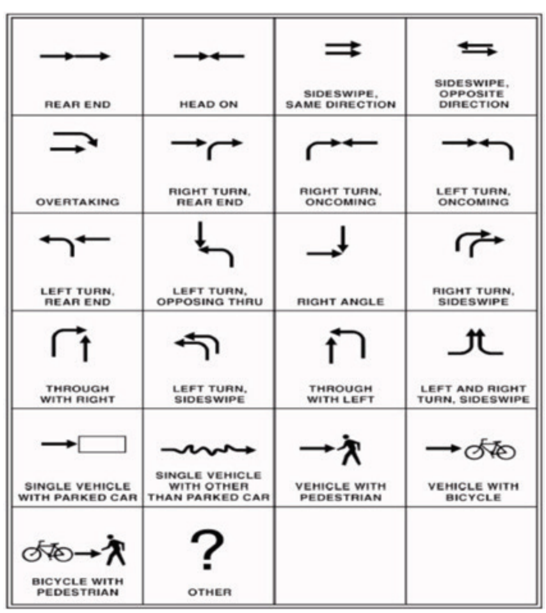
Vehicle-to-Vehicle Control

V2V, more difficult to realize because of its decentralized structure, aims at organizing the interaction among vehicles and possibly developing collaborations among them. At this level, information is interchanged and decisions are made on a "local" basis (e.g., among a group of vehicles in proximity to each other). The introduction of such information interchange requires an agreement among car manufacturers and suppliers in terms of communication technology, protocols, and the like, and efforts are under way in this direction (the CAR2CAR Consortium). The communication technology is based on IEEE 802.11, also known as Wireless LAN. A frequency spectrum in the 5.9-GHz range has been allocated on a harmonized basis in Europe in line with similar allocations in the U.S. (although the systems are not yet compatible).

In the V2V concept, when two or more vehicles or roadside stations are in radio communication range, they connect automatically and establish an ad hoc network enabling the sharing of position, speed, and direction data. Every vehicle is also a router and allows sending messages over multihop to more distant vehicles and roadside stations. The routing algorithm is based on the position of the vehicles and is able to handle fast changes of the network topology. Control technology comes into play at local and higher layers of the architecture. Uncertainties, delays, partial measurements, safety and performance objectives, and other aspects must be considered, and the system must be capable of making automatic or semiautomatic decisions, providing warnings/information and potentially effecting actions.



A V2V example (Source: N. Hashimoto, S. Kato, and S. Tsugawa, A cooperative assistance system between vehicles for elderly drivers, IATSS Research, vol. 33, no. 1, 2009, pp. 35-41)



A taxonomy of possible accidents illustrating the variety of situations that must be detected and handled optimally and robustly to avoid possibly dangerous situations (Source: SAFESPOT Project)