The Impact of Control Technology

OVERVIEW, SUCCESS STORIES, AND RESEARCH CHALLENGES

EDITED BY:

Tariq Samad and Anuradha Annaswamy

SPONSORS:

IEEE Control Systems Society

Institute for Advanced Study, Technical University of Munich

CoTeSys, Technical University of Munich

Deutsche Forschungs-Gemeinschaft

U.S. National Science Foundation

The Impact of Control Technology

Overview, Success Stories, and Research Challenges

Edited by:

Tariq Samad and Anuradha Annaswamy

Published by:

IEEE Control Systems Society

February 2011

Sponsors:

IEEE Control Systems Society
Institute for Advanced Study, Technical University of Munich
CoTeSys, Technical University of Munich
Deutsche Forschungs-Gemeinschaft
U.S. National Science Foundation

This report is available online at http://ieeecss.org/main/IoCT-report.

Foreword

Control systems research has a long history of mathematical rigor, with application to diverse branches of science and engineering. The methods, algorithms, and tools developed by control researchers have been widely used by generations of engineers to solve problems of practical importance with enormous impact on society. Control concepts have been crucial in the design and development of high-performance airplanes, fuel-efficient automobiles, industrial process plants, manufacturing enterprises, smart phones, planetary rovers, communication networks, and many other applications across various sectors of industry. In these and other complex engineering systems, control theory and its technological artifacts are also widely used to ensure reliable, efficient, and cost-effective operations. The tremendous variety of control applications, however, makes it difficult for control technologists working in one domain to be aware of recent developments in other areas. It is even more difficult for researchers and decision makers outside the control discipline to fully appreciate the contributions that control technology has made to modern societies or its potential for future impact.

Funding agencies in the U.S., Europe, Japan, and Australia, among others, have invested in control systems research over the last 40 years. The funded research has partly been on fundamental questions and challenges such as robustness, stability, and adaptability, and partly on cross-disciplinary endeavors in areas such as complexity, wireless sensor networks, real-time systems and platforms, and cognitive systems. As with other branches of engineering and science, however, recent funding trends in control point toward applied rather than basic research and toward the pursuit of application challenges. The new model for research in the "Innovation Economy" targets collaborations between academia and industry on a global scale, where the competition for limited funding resources is on the rise.

The National Academy of Engineering in the United States has identified 14 grand challenges primarily dealing with energy, environment, transportation, and healthcare [1]. The European Commission's R&D and innovation programs focus on similar objectives [2,3]. New developments in mathematical systems and control theory, algorithms, methods, and tools are needed to meet many of these challenges today and in the future. As a result, the control community is increasingly engaged in collaborative projects dealing with emerging concepts and themes—examples include cyber-physical systems and systems of systems—and in applications of these research fields in areas such as transportation networks and systems, medical devices, factories of the future, energy conservation and efficiency, and renewable energy integration for the power grid. For example, control systems researchers are teaming with computer scientists in using new hardware and software platforms to develop a new systems science addressing issues of sustainability, security, and cyber-enabled reconfiguration of engineering systems. In the last five years, radical developments have taken place in network science, a branch of complexity theory that seeks to establish universal laws and principles of networks, ranging from links between brain cells to the structure of the Internet [4]. Although their role is often unheralded, control engineers and scientists have been at the core of numerous innovations in this and other areas.

The Impact of Control Technology: Overview, Success Stories, and Research Challenges seeks to identify and recognize recent accomplishments in control systems research, to highlight new research directions and challenge problems for the field, and to communicate the contributions and potential of control within and beyond the boundaries of the discipline. In addition to articles reviewing the application of control in different domains and articles discussing new research frontiers, the report also includes a novel feature: 40-plus case studies, each in a graphical two-page format, illustrate "success stories" and "grand challenges" that will serve to enlighten and inspire the control community as well as its stakeholders. The product of a unique outreach experiment, this report conveys the essential concepts, ideas, and impacts of the control discipline. *The Impact of Control Technology* is also an excellent example of international collaboration, and academic-industry collaboration, in the control field.

We consider this report a milestone for the promotion of control systems research. The content of the report argues strongly for the importance of control in preparing the next generation of scientists and engineers. This endeavor is of value to students and faculty, as well as to R&D leaders and decision makers in academia, industry, and government. In these challenging times of rapid technology developments and reduced funding for basic research, each research community needs to reinvent itself and make a compelling case to justify investment. In addition, by highlighting the impact of control on society, this report will be instrumental for outreach to the broader public.

The Impact of Control Technology articulates the value of control, for today and for the future. In this regard, the report is a unique document, and we hope it will serve as a foundation for further such efforts as well.

RADHAKISAN BAHETI
NATIONAL SCIENCE FOUNDATION
ARLINGTON, VA, U.S.A.

ALKIS KONSTANTELLOS
EUROPEAN COMMISSION
BRUSSELS, BELGIUM

References

[1] National Academy of Engineering. Grand Challenges for Engineering [Online], 2008. Available at http://www.engineeringchallenges.org/.

[2] European Commission. Interim Evaluation of the Seventh Framework Programme (FP7), Report of the Expert Group [Online], 12 November 2010. Available at http://ec.europa.eu/research/evaluations/index_en.cfm?pg=fp7.

[3] European Commission. Facing the Future, EFP Brief No. 179 [Online], September 2010. Available at http://www.foresight-platform.eu/category/foresight-briefs/time-horizon/.

[4] R. Ogle. *Smart World: Breakthrough Creativity and New Science of Ideas*. Boston: Harvard Business School Press, 2007.

Preface

In October 2009, we helped organize a workshop entitled "Impact of Control: Past, Present, Future," in Berchtesgaden, Germany. Our co-organizers were Martin Buss and Patrick Dewilde of the Technical University of Munich and Gary Balas of the University of Minnesota. The principal sponsors of the workshop were the IEEE Control Systems Society, the Institute of Advanced Study (IAS) and the Cognitive Technical Systems (CoTeSyS) program at the Technical University of Munich, the German Federal Research Agency (Deutsche Forschungs-Gemeinschaft), and the U.S. National Science Foundation.

At the conclusion of the workshop, there was consensus among the participants that the presentations and discussions needed to be documented in the form of a report. In taking up this challenge, our original intent was to compile a report based entirely on the workshop proceedings. As we reflected on the "impact" theme, however, our plans for the report expanded. Thus, in addition to including sections derived from the workshop sessions, we decided to showcase, in the form of two-page flyers, a number of examples of recent successes and future opportunities for control. Our solicitations to the controls community to contribute "success stories" and "grand challenges" resulted in numerous responses, of which 40-plus are included in Parts 2 and 4 here. Other material was also added that was not directly reflected in the workshop agenda.

All told, this report is the result of well over 50 contributors from across the globe, representing numerous theoretical and application areas of control science and engineering. We have been delighted by the interest shown in this project by the contributors and are grateful for their efforts. A special thanks to the lead authors for the sections in Parts 1 and 3, who graciously undertook multiple rounds of revisions at our behest: Kishan Baheti, Ken Butts, Eduardo Camacho, Ian Craig, Munther Dahleh, Frank Doyle, Luigi Glielmo, Sandra Hirche, Christian Philippe, Mark Spong, and Greg Stewart.

The leadership of Gary Balas, Martin Buss, and Patrick Dewilde for the workshop was instrumental to its success and thus paved the way for this report. Kishan Baheti of the U.S. National Science Foundation and Alkis Konstantellos of the European Commission repeatedly emphasized to us the importance of this report for the controls community and gave us guidance on various occasions; without their continued encouragement, this report would have been a much less ambitious undertaking. We would also like to acknowledge the generous support of the Institute for Advanced Study at the Technische Universität München, which was crucial for the organization of the workshop and the preparation of this report.

Our thanks are also due to Barb Field, who copyedited and compiled the report, to Lindsay Fend of Waldbillig & Besteman, who was responsible for the design and production of the flyers, and to Judy Scharmann of Conference Catalysts, who set up the website for the report.

The first editor is grateful for the support provided to him by Honeywell for this project. The second editor is grateful to IAS for sponsoring her sabbatical, during which the workshop was conceived and held.

TARIQ SAMAD — MINNEAPOLIS, U.S.A.

ANURADHA ANNASWAMY — CAMBRIDGE, U.S.A.

JANUARY 2011

Contents

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Part 3: Cross-Cutting Research Directions

Part 4: Grand Challenges for Control Research

Appendices

Introduction and Summary

Tariq Samad and Anuradha Annaswamy

The Impact of Control Technology

Control is everywhere. Aircraft and spacecraft, process plants and factories, homes and buildings, automobiles and trains, cellular telephones and networks . . . these and other complex systems are testament to the ubiquity of control technology. Some artifacts of modern times would simply not be possible without control. And for many others, substantial, even revolutionary, advances in their performance, safety, reliability, and affordability have been achieved as a result of the ingenuity and effort of control engineers and scientists.

The realized impact of control technology is matched—indeed, overmatched—by its anticipated future impact. Decades of successful applications have hardly exhausted the potential or vitality of the field. The number and size of control conferences and journals continue to grow, new societal imperatives highlight the importance of control, and investments in control technology and technologists are taking place in old and new industrial sectors. Control is not only considered instrumental for evolutionary improvements in today's products, solutions, and systems; it is also considered a fundamental enabling technology for realizing future visions and ambitions in emerging areas such as biomedicine, renewable energy, and critical infrastructures.

The increasing complexity of technological systems demands inter- and cross-disciplinary research and development. Collaborations between control and other fields have been consistently productive. In particular, in the course of these collaborations, a widening appreciation of the principles of control has been apparent. Wherever dynamics and feedback are involved—and they are increasingly recognized as pervasive properties of complex systems—control expertise is regarded as crucial. Control is also seen as the paragon of rigor and the systems perspective by experts in other disciplines, distinctions that are being exploited as larger-scale, safety-critical, and mission-critical systems are developed or envisioned.

But if control is everywhere in reality, it is also nowhere in perception. Despite its accomplishments and promise, control remains a "hidden technology" [1]. We can offer hypotheses for why this is the case—the value of control is in intangible algorithms, its rigor and formality are intimidating to the uninitiated, its applications and research are distributed across many scientific and engineering disciplines—but regardless, the fact remains that the breadth and scale of the impact of control are largely unknown and unheralded. Remarkably, this is true not only outside the controls community, but even within it!

Our motivation in preparing this report is to help rectify this lack of awareness by highlighting the to-date and to-come impact of control technology. To this end, we have included overviews of the applications of control systems to a number of domains, discussions of emerging research areas for the field, and summaries of specific control-related successes and challenges in industry and government. Each of these topics is covered in a separate part of the report.

Although the report includes sections on research directions for control theory, our principal focus is on the accomplishments and promise of control technology and not on laying out the theory road map for

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

the field. In this sense, we consider this report as complementing, and not superseding, previous publications such as [2], which remains an excellent resource for the control research community.

Part 1: Application Domains for Control

The fundamentals of control science are universal, but the impact of control technology results from the combination of these fundamentals with application-specific considerations. The first part of the report consists of discussions of the role of control in both traditional and emerging application domains:

- Aerospace (C. Philippe et al.)
- Process industries (I. Craig et al.)
- Automotive (L. Glielmo et al.)
- Robotics (M. Spong and M. Fujita)
- Biological systems (F. Doyle et al.)
- Renewable energy and smart grids (E. Camacho et al.)

The material in these sections includes historical references, highlights from successful applications, economic and market information, and current and future challenges. Recommendations for research are also provided. We note that the above list of domains is hardly comprehensive; other important domains include buildings, railways, telecommunications, disk drives, and others.

Part 1 also includes two brief sections that offer integrative perspectives on control applications: K. Butts and A. Varga discuss control development processes and related tools and platforms, and G. Stewart and T. Samad discuss application-specific requirements and factors that are important for "real-world" control implementations.

Some general considerations stand out from the material in this part of the report:

- A few decades ago, discussion of the impact of control technology would have been limited to a few industries: aerospace systems, process plants, and homes and buildings. Today these traditional domains of control have been supplemented with a litany of others. The application domains represented here include ones where control has historically been prominent, ones that have embraced control technology relatively recently, and emerging domains that will, we expect, provide new opportunities for the field.

- In addition to the broadening of industry scope, control has also developed into a highly scalable technology. In multiple application areas, we have seen control principles initially being applied to individual sensors and actuators, then on multivariable systems, and even at plantwide scales. Ambitions now reach enterprise and "system of systems" levels.

- Successful applications of control are not the result of control expertise alone. In-depth domain knowledge has always been necessary. As the applications of control have broadened, the connections with traditional and new domains have been established and strengthened.

- Furthermore, technological prerequisites must be satisfied before control, especially advanced control, can be applied. Several new application areas have become viable for control as a result of developments in novel sensors and actuators, for example.

- Quantifying the impact of control technology is difficult. A control algorithm doesn't solve any problem in and of itself; the control innovation is linked with ancillary developments. In cases where economic or other societal benefits have been estimated, the results point to tremendous scale of impact [3].

Part 2: Success Stories in Control

After the broad application-oriented discussions in Part 1, the next part of the report highlights significant specific accomplishments of the field in a form intended for communication both within the controls community and with its stakeholders. The two-page flyers on "success stories" featured here were solicited from the controls community worldwide. The flyers include some technical details, but we have attempted, as far as possible without risking superficiality of treatment, to keep them accessible to a non-controls audience. Some documentation of societal/industrial benefit is included in all cases.

Examples from the content of Part 2 are:

- Mobile telephones rely on control—to the tune of billions of feedback loops across the globe.
- With antilock brakes and stability and traction control, automotive safety has been revolutionized by control technology.
- A mechanical control invention for automotive suspensions resulted in a win on its first use in Formula One.
- Advanced control is now widely implemented in devices like printers and copiers.
- Collision avoidance systems are well established in air traffic management and rely on estimation and control algorithms.
- Optimization and control technology implemented in railroads is reducing fuel consumption by tens of thousands of gallons per year, per locomotive.
- Paper machines manufacture paper whose thickness is controlled to within microns—over reels of paper that are often 40 km in length.
- Hundreds of ethylene processes are dynamically optimized with model predictive control techniques, resulting in over $1 million of increased production annually per plant.
- Warehouse operations are autonomously controlled by hundreds of mobile robots.
- Improved audio reproduction technology derived from control theory enhances perceptual quality by over 30% and is implemented in over 15 million integrated circuits.

These and the other success stories included are a small fraction of what has been achieved with control. Nevertheless, the significance and variety of these contributions is an indication of how extensive the true footprint of control is!

Part 3: Cross-Cutting Research Directions

We next move to the forward-looking content of the report. Part 3 consists of discussions of four topics of current research that are gaining increasing interest, both in the research community and for government investment. The topics covered are:

- Networked decision systems (M. Dahleh and M. Rinehart)
- Cyberphysical systems (R. Baheti and H. Gill)
- Cognitive control (M. Buss, S. Hirche, and T. Samad)
- Systems of systems (T. Samad and T. Parisini)

A few points of commonality among these topics are worth noting:

- The research required is interdisciplinary and multidisciplinary, an observation that emphasizes the critical need for the controls community to collaborate with other fields. Connections with computer science, other algorithmic fields such as information theory, and relatively new sources of inspiration such as cognitive science have been and are being established and need to be further promoted.

- Networks are pervasive. The term may only appear in one of the section titles, but it is implicit in the others. Centralized approaches are seen as being untenable for several reasons. Solution methods and architectures are increasingly distributed, decentralized, coordinated, and collaborative.

- In many situations, subsystems have high degrees of autonomy and heterogeneity. A continuing research imperative is to figure out how we can realize system-level goals for performance, predictability, stability, and other properties through appropriate system designs and subsystem interactions.

- These research directions are cross-cutting in the sense that each is relevant for a variety of challenges that are engaging industry, society, and government. Examples include smart grids, intelligent transportation systems, complex infrastructures, and emergency response teams. These and other examples are mentioned in multiple sections. Some overlap of concerns is unavoidable, but these themes highlight different facets of these complex needs.

- Although control systems have never been isolated components, the interconnections between control and other areas have not been fully explored in the past. This limitation is now being overcome. Interconnections and integration with real-time platforms, with humans as users and in other roles, and with other systems are points of focus.

Complexity is an overarching feature in this part of the report, but the themes provide color and specificity to this buzzword. The networked aspect is central, but it is a substrate and a metaphor. The complexity of control research is manifested in the integration and synthesis among controllers and optimizers, hardware and software components, humans and engineered intelligent agents, in more or less cooperative environments, with hierarchical and heterarchical structures, across physical domains that span most fields of engineering, representing spatial and temporal time scales ranging from the nano and micro to the mega and macro. The articles in Part 3 reflect these trends in control research.

Part 4: Grand Challenges for Control

Whereas Part 3 focuses on cross-application research themes, Part 4 outlines a set of exciting research opportunities in control that target domain-related challenges.

Analogous to Part 2, this part of the report also consists of two-page illustrated briefs. The focus here, however, is not on past accomplishments but on future opportunities. The close to 20 featured here demonstrate the expanding scope and scale of control. Examples of the featured challenges, for all of which control technology is a critical need, are:

- An artificial pancreas for treatment of diabetes is under development—control scientists are leading the effort.
- Control-enabled high-altitude wind energy devices have been demonstrated and promise efficiencies that are substantially higher than for today's turbines.
- Feedback and dynamics are essential for the development of smart grids—the smart grid can be considered an end-to-end optimization and control problem.
- Active control of unstable combustion phenomena will be essential for realizing higher efficiency and reliability and lower emissions in turbine engines.
- Next-generation air traffic control approaches are being developed with the objective of substantially reducing the energy use associated with air transportation.
- With successes in applying control to process units and even plantwide, industries are seeking to close the loop around entire supply chains.
- In automotive systems, vehicle-to-vehicle and vehicle-to-infrastructure coordination is projected to improve the efficiency and safety of road transportation.
- Advanced control is increasingly recognized as critical for achieving dramatic reductions in energy consumption in buildings.
- The ability of atomic force microscopes to image and manipulate matter at the nanometer scale is entirely dependent on the use of feedback loops.

Appendices

This report also includes two appendices: a brief account of the Berchtesgaden workshop and affiliations and e-mail addresses for the principal authors of the sections and flyers in this report.

Concluding Remarks

There's more to control technology than is typically appreciated—whether by its exponents, its beneficiaries, or others directly or indirectly associated with the field. Control has played an instrumental, if often behind the scenes, role in the development of engineering solutions to outstanding problems, resulting in substantial societal and industry impact. The report discusses the role of control in a number of prominent application domains and illustrates successes achieved in these domains and in others.

The report also highlights the fact that there is no dearth of opportunity for research in control. Control is flourishing as a research field unto itself and even more so as a keystone discipline for addressing multidisciplinary challenges. Evolving from and strengthened by a mature core focused on single systems, the new network-centric centers of gravity of control research, through productive interactions with other scientific disciplines and with an increasing number of application domains as targets, are demonstrating the power and advantage of the systems and control approach in new arenas.

We hope this report will help lift the veil on the "hidden technology" that control often seems to be. But this is not to say that we have exposed the impact of control in its entirety. Any artifact such as this report—a snapshot of success and opportunity in a dynamic and vibrant field—is inevitably incomplete. We suggest that as a Web-based resource, this report itself can be dynamic. . . . Additional success stories and grand challenges can conveniently be integrated, and new sections discussing application domains and research directions can be incorporated as well. We invite volunteers from the controls community to contribute to, and help lead, this effort.

References

[1] K.J. Åström. "Automatic control—the hidden technology," in *Advances in Control - Highlights of ECC'99*, P. Frank (ed.). London: Springer Verlag, 1999.

[2] R. Murray et al. Control in an Information Rich World: Report of the Panel on Future Directions in Control, Dynamics, and Systems [Online], 2002. Available at http://resolver.caltech.edu/CaltechBOOK:2002.001.

[3] J. Pereira (ed.). Monitoring and Control: Today's Market and Its Evolution till 2020. Luxembourg: Office for Official Publications of the European Communities, 2009.

Part 1

SELECTED APPLICATION DOMAINS FOR CONTROL

Aerospace Control

Christian Philippe, Anuradha Annaswamy, Gary Balas, Johann Bals, Sanjay Garg, Alexander Knoll, Kalmanje Krishnakumar, Massimo Maroni, Robert Osterhuber, and Yin C. Yeh

Apollo as the Catalyst for Control Technology

Advanced control technology played a fundamental role in putting the first man on the moon. To meet the challenging lunar descent and landing requirements, time and fuel optimal nonlinear control laws and variable Kalman-filter-based state estimators were developed and implemented into the Apollo lunar module first-generation digital flight computer. Remarkably, the memoryless thrust vector control law was the first application of the minimum time control law for a third-order plant. The success of the Apollo program also paved the way for embedded software, online reconfiguration software, concurrent control design and software engineering processes, man-machine interfaces, and digital fly-by-wire technologies.

> *Advanced control technology played a fundamental role in putting the first man on the moon.*

Control technology developed during the Apollo program was a catalyst for safer and more efficient aircraft. In the late 1960s, engineers at NASA Flight Research Center (now NASA Dryden) proposed replacing bulky mechanical flight-control systems on aircraft with much lighter weight and more reliable analog fly-by-wire technology. As the Apollo program came to completion in the early 1970s and following Neil Amstrong's recommendation, NASA Dryden engineers developed a digital fly-by-wire (DFBW) solution using the specialized software and hardware developed for Apollo. On 25 May 1972, the successful testing of the world's first-ever DFBW technology on a modified F-8 Crusader jet fighter precipitated a revolution in aircraft design and performance [1].

For military aircraft, the deployment of DFBW control systems allowed the development of highly maneuverable fighter aircraft and the improvement of their "carefree handling" performance and combat survivability by preventing stalling, spinning, and actuator hydraulic failures. In the commercial airline market, Airbus introduced full-authority fly-by-wire controls in 1988 with the A320 series, followed by Boeing with their B-777 in 1995. The primary benefits were (1) a reduction of the airframe weight through the use of smaller, lighter aerodynamic control surfaces and (2) increased aircraft safety and reliability.

> *Fly-by-wire control systems enabled the development of highly maneuverable fighter aircraft and improvements in handling performance and combat survivability.*

Nowadays, DFBW control systems are commonly implemented in high-performance jet fighters and aboard commercial airliners.

With regard to space applications, control-enabled solutions have guaranteed access to space through the successful development of launchers and space transportation systems, bringing many benefits to society. For instance, the successful deployment of interplanetary probes and space-based observatories such as Pioneer, Voyager, Cassini-Huygens, and the Hubble Space Telescope has allowed the exploration of our solar system—Venus, Mars, Jupiter, and Saturn's moon, Titan—and a greater knowledge of the universe. Thanks to space-based data from remote sensing and meteorological satellites, a better understanding of the earth, its climate, and its changing environment has been made possible. For example, the Franco-American mission Topex-Poseidon has shown through space altimetry that the oceans have been rising over the past decade; it has also provided unexpected information for monitoring oceanic phenomena such as variations in ocean circulation on the level of the 1997-1998 El Nino event.

Finally, since the launch of the first telecommunication satellites in the sixties, control technology has continued to play an important role in the successful deployment of more powerful satcoms featuring large flexible deployable antenna and solar arrays. Today telecommunication and navigation satellites are part of everyone's life: Internet, tele-education, telemedicine, videoconferencing, mobile communications, digital broadcasting, search and rescue, and traffic management.

Successful Applications and Demonstrations

Both aeronautics (commercial and military aircraft and unmanned aerial vehicles (UAVs)) and space (launchers, manned and unmanned space transportation vehicles, satellite and planetary rovers) application fields share common and specific control-relevant requirements. These requirements are listed in Fig. 1.

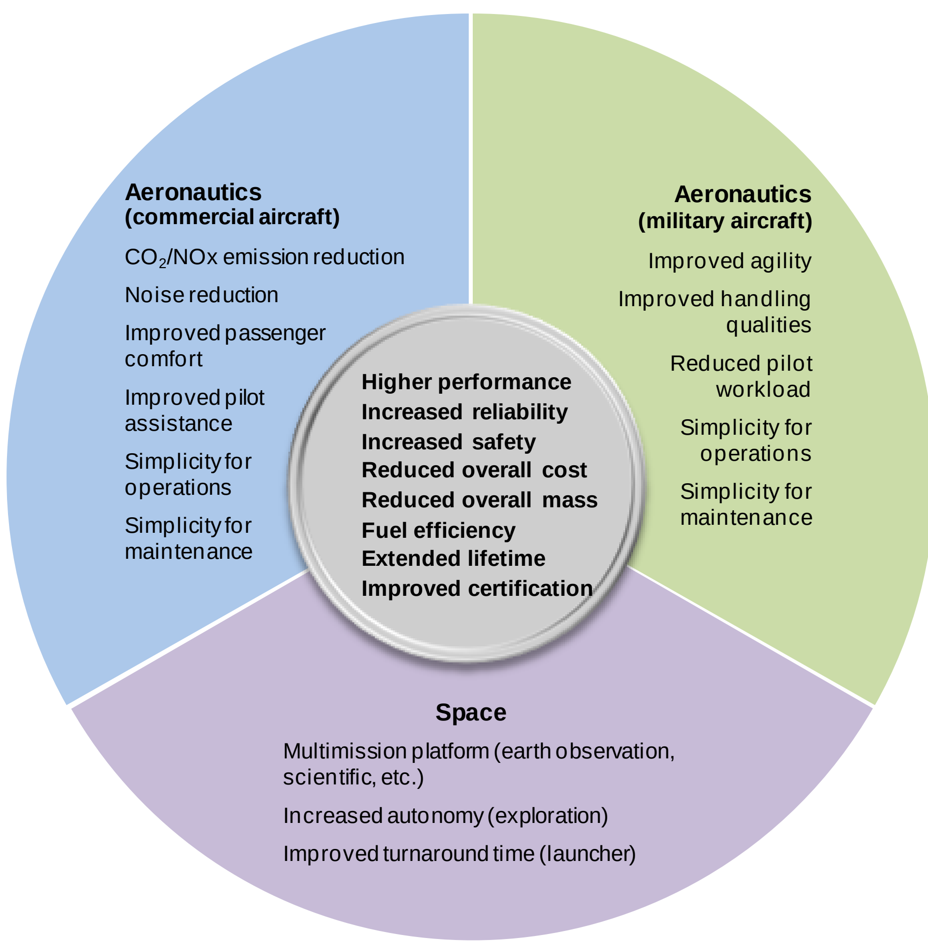


Figure 1. Common and specific control-relevant requirements for aeronautics and space.

Over the last 50 years, both application fields have seen the successful deployment of control technologies for satisfying the above control-relevant requirements. Both application fields require rigorous engineering processes, including standards such as Radio Technical Commission for Aeronautics/Design Objective-178B for aeronautics and European Cooperation for Space Standardization, and they both demand efficient and rigorous (model-based) design, development, and validation methods and tools.

Since at the conceptual level the list of control techniques investigated by academia, agencies, research organizations, and industries is lengthy and the techniques are often applied without taking into account the specific needs and constraints (implementation, validation, certification, financial) associated with the applications, the enumeration of successful control applications is limited to those that have been successfully deployed by the aerospace industries or investigated by research organizations. A clear assessment of the impact of advanced control technologies on past and present operational projects cannot be made due to information restrictions or confidentiality resulting from military applications or industry competitiveness concerns.

Commercial Aircraft

In addition to digital fly-by-wire control technology, which has reduced the operating cost of commercial airplanes, NASA Dryden Flight Research Center has initiated the development of propulsion controlled aircraft (PCA) technology with the main goal of reducing the aircraft accident rate by a factor of 10 within 20 years. The PCA is a computer-assisted engine control system that enables a pilot to land a plane safely when its normal control surfaces are disabled. The first successful demonstration of the PCA technology on an airliner took place in 1995. Although the technology is proven, it has not yet been incorporated into future aircraft designs. A further extension to DFBW flight control systems is to implement functions capable of compensating for aircraft damage and failure during flight, such as automatically using engine thrust and other avionics to compensate for severe failures—loss of hydraulics, loss of rudder, loss of ailerons, or loss of an engine. This new generation of DFBW flight control systems is called *intelligent flight control systems* (IFCS).

As a result of the miniaturization of sensor technologies, increasing actuator performance capabilities, and increasing processing resources, integrated flight-structural control technologies are being investigated that should further improve the safety and environmental performance of the aircraft as well as the comfort of passengers. For instance, one potential “green” aviation technology is *active wing shape control,* which holds promise for improved aerodynamic efficiency, lower emissions, reduced noise, and minimized carbon footprint. This control technology consists of shaping the wing structure in flight by actively controlling the washout twist distribution and wing deflection so as to affect local angles of attack in a favorable manner that leads to lower drag and higher lift. Another example is active load control technology, which could reduce structural weight considerably by reducing aerodynamic peak and fatigue loads at critical locations in the airframe structure. The associated functions are realized by control allocation and coordination, affecting distribution of aerodynamic loads over the airframe, as well as by active damping of airframe structural modes.

Military Aircraft and UAVs

To respond to the continuous demand for increased performance in military aircraft, the deployment of active control technologies has been mandatory. For example, the following functions are currently implemented onboard fighter aircraft:

- Carefree handling by providing angle-of-attack control and angle-of-sideslip suppression, which lead to automatic protection against stall and departure;

- Carefree handling by the automatic limiting of normal acceleration and roll rate to avoid over-stressing of the airframe;

- Automatic controller reconfiguration, allowing mission continuation or safe recovery following system failures or battle damage;

- Automatic terrain-following functions using information from the radar altimeter or digital terrain elevation database, aiming at holding the aircraft at a constant distance above ground level;

- Advanced autopilots, providing significant reductions in pilot workload and weapon system performance benefits.

Along with the increase in aircraft performance, specific safety functions are now implemented to protect the pilot, such as the pilot-initiated *spatial disorientation* automatic recovery mode from both nose high and low situations and automatic *g-loc* (g-force-induced loss of consciousness) recovery mode.

Aircraft Engines

With the increased emphasis on aircraft safety, enhanced performance and affordability, and the need to reduce the environmental impact of aircraft, corresponding progress needs to be made in the area of aircraft propulsion systems. Over the years, considerable improvements have been made in engines, with control playing a significant role. One such example is the work being carried out at NASA Glenn Research Center in partnership with the U.S. aerospace industry and academia to develop advanced controls and health management technologies through the concept of an *intelligent engine.* Turbine engine manufacturers such as Siemens-Westinghouse, Rolls-Royce, and United Technologies have successfully employed control principles in improving efficiencies and performance. In most cases, passive control methodologies have entered the production phase, with active control successes demonstrated in academia and research laboratories. The key enabling technologies for an intelligent engine are the increased efficiencies of components through active control of inlets, compressors, and combustors, advanced diagnostics and prognostics integrated with intelligent engine control to enhance component life, and distributed control with smart sensors and actuators in an adaptive fault-tolerant architecture.

Notable recent successes include:

- Development of life-extending control through intelligent modification of the engine acceleration schedule to minimize thermomechanical fatigue for each takeoff-to-landing cycle. Demonstrated 20% improvement in “on-wing” engine life through real-time engine/control simulation.

- Successful demonstration of control of thermoacoustic instability in combustors in gas turbine engines by modulating the fuel entering the engine using servo-valves and control strategies.

- Flight demonstration of high-stability engine control, which allows operation of engines with reduced stall margins during cruise, thus increasing fuel efficiency by up to 3%. The technology works through estimation of inlet distortion effects on stall margin using pressure sensors on the fan circumference and coordinating fuel flow and nozzle area control to maintain a desired stall margin.

Space

Robust control techniques such as H_∞/H_2 have been successfully applied to deal with complex architectures such as large flexible appendages (solar arrays and deployable reflectors) and requirements such as tight pointing stability performance, while reducing development cost and schedule.

For instance, the Linear Quadratic Gaussian (LQG) controller used for the atmospheric flight phase of the Ariane 5 launcher was replaced by a H_∞-based controller for the Ariane 5 Evolution [2]. This change was deemed necessary to optimize the control design tradeoff between the low-frequency performance requirements, such as load reduction and tracking of the attitude setpoint, and the attenuation of the low-frequency structural bending and fuel sloshing modes. For telecommunication satellites, the introduction of a robust control approach through a loop-shaping H_∞ design has allowed a 10% reduction in propellant mass consumption during station-keeping maneuvers.

Nowadays, increasing computing capability allows multidisciplinary modeling and simulation, which are essential for the development of robust controllers for complex uncertain systems. Recent progress in multidisciplinary requirements and integrated design processes, advanced analysis tools, commercial automatic production code generators, automatic advanced formal verification and test case generation tools, and the like, has reduced the development time and cost of embedded flight control systems.

Whatever the application field, decision makers rely on already proven technical solutions. This is especially true for space applications, as solutions cannot be tested beforehand due to the difficulties of reproducing space-representative conditions on Earth. Thus, for critical space control technologies or new control system concepts, dedicated precursor missions, usually named Pathfinder or X-vehicle, are typically implemented before deployment on the full-fledged mission. This approach is necessary for decreasing the technical risk and cost of the overall mission.

Furthermore, the gap between new control techniques and associated certification processes, including tools, methods, and standards, cannot be too large, otherwise the control technologies cannot be operationally deployed. Finally, despite the potential advantages afforded by advanced control techniques, they usually add complexity in the design, analysis, and tuning of the flight control system, thus requiring more skillful control engineers.

Market Sizes and Investments

Both Boeing and Airbus estimate that new aircraft demand will average around 1,300 per year over the next 20 years (2009-2028). This corresponds to a commercial airline market value of around $3 trillion. For UAVs, Teal Group's 2009 market study estimates that spending will almost double over the next decade from current worldwide UAV expenditures of $4.4 billion annually to $8.7 billion within a decade, totaling just over $62 billion [3]. The most significant catalyst to the UAV market is the enormous growth in interest by the defense sector.

Over the last 15 years, the average European space industry sales, including commercial and institutional programs, is around €4.5 billion annually. In 2007, space industry sales amounted to €6 billion [4]. Euroconsult estimates that around 1,200 satellites, excluding microsatellites (weighing less than 40 kg) and classified military satellites, mainly from United States and Russia, will be built and launched worldwide over the next decade (2008-2018), an increase of 50% compared to the previous decade [5]. Market revenues generated from the manufacturing and launch of these satellites are forecast to grow at the same rate, reaching $178 billion. Earth observation (EO) is emerging as the largest application with a total of 230 satellites, reflecting the priority given by governments to the challenges of global warming and climate change.

Depending on the application type—launcher, EO satellite, satcom, scientific satellite—the cost of the avionics system, including the embedded GNC software and equipment, represents around 8-15% of the overall cost of the satellite. For aircraft, the average value of the avionics system is 12%. Therefore, the market for control technology over the next decade can be conservatively estimated at not less than $25 billion for civil, military, and governmental space applications, $225 billion for both commercial and military aviation, and $5 billion for unmanned systems (Fig. 2). (Engine control systems, cabin environmental control systems, and other "embedded" applications of control are additional to avionics.)

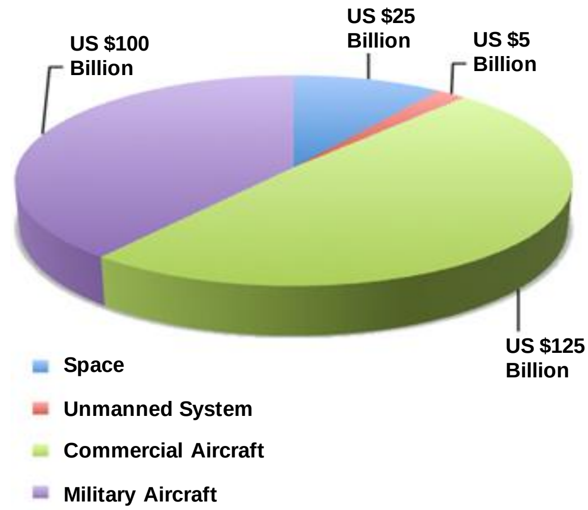


Figure 2. Estimated market for commercial and military aircraft, unmanned systems, and space (civil, military, and governmental) applications over the period 2009-2019.

Future Challenges

Air traffic demand is predicted to double in the next 10 to 15 years and to triple in 20 years' time. This growth cannot be sustained without a complete overhaul of the air traffic control infrastructure to optimize air routes and eliminate congestion. As a result, the U.S. Federal Aviation Administration (FAA) has initiated NextGEN (Next Generation Air Transportation System), whose primary goals are to provide new capabilities that make air transportation safer and more reliable, improve the capacity of the National Airspace System (NAS), and reduce aviation's impact on our environment [6]. A sister initiative in Europe called SESAR (Single European Sky ATM Research) aims to increase air transport safety and reduce greenhouse gas emissions by 10% per flight. The projected SESAR program cost is €50 billion at completion in 2020 [7].

In addition to air traffic management, aeronautics research investment priorities in Europe are to develop safer, greener, and smarter transport systems. Air travel is the fastest-growing source of greenhouse gas emissions in the world, and air transportation system energy requirements are expected to more than double in the next three decades. Each long-distance flight of a 747 adds about 400 tons of CO_2 to the atmosphere (about the same amount a typical European uses for heating and electricity in a year). Aviation now consumes about 13% of transportation-related energy, and this percentage is growing rapidly. Emissions at altitude are estimated to have two to four times greater impact, relative to terrestrial emissions, than reflected by the percentage of carbon emissions. Energy efficiency in air traffic management, in flight control of individual aircraft, and in engine control systems are all priority research needs.

The European Commission's CleanSky effort will amount to €1.6 billion (2008-2013), whereas the research effort of NASA's aeronautics programs is $2.6 billion (2009-2014). NASA's research programs will focus on the following technologies [8]:

- *Integrated vehicle health management* technology will enable nearly continuous onboard situational awareness of the vehicle state for use by the flight crew. It will also improve the safety and reliability of the aircraft by performing self-diagnosis and self-correction of in-flight anomalies.

- *Integrated intelligent flight deck* technology will allow robust detection of external hazards with sufficient time-to-alarm for safe maneuvering to avoid the hazards. It will also support new pilot tasks consisting of collaboration and negotiation with other aircraft and air traffic controllers.

- *Integrated resilient aircraft control* technology aims at enabling the aircraft to automatically detect, mitigate, and safely recover from an off-nominal condition that could lead to a loss of control in flight.

- *Advanced validation and verification of flight-critical systems* will provide methods for rigorous and systematic high-level validation of system safety properties and requirements from initial design through implementation, maintenance, and modification, as well as understanding of tradeoffs between complexity and verification in distributed systems. In addition, tools will be developed for analysis and testing of systems-of-systems capabilities.

In addition, the research programs will start to address the technical and regulatory issues related to the integration of unmanned aircraft systems in NAS. In support of NextGEN, enabling control optimization technologies will be developed for traffic scheduling and route planning, as well as balanced allocation of resources to maximize airspace productivity in response to arrival, departure, and traffic demands.

For military aircraft, the following control technologies are being investigated that should have safety, financial, and environmental benefits:

- Damage-tolerant flight control should automatically reconfigure the aircraft flight controls after significant loss of control due to battle damage.

- Automatic collision avoidance should reduce the risk of ground and midair collisions.

- Autonomous formation flight should provide a 5-10% reduction in fuel consumption by a trailing airplane during cruise.

- Autonomous midair refueling should allow unmanned air systems to significantly increase their mission times and operational range.

For future stealth aircraft, advanced air data systems will be required because external measurement devices need to be minimized. Moreover, the unusual shaping of such aircraft and the need to reduce the number and size of control surfaces for low observability, the possible reliance on thrust vectoring, and the development of novel control methods such as nose suction/blowing, are likely to lead to highly nonlinear aerodynamic characteristics that will require advances in the development of robust flight controllers. Finally, for some specific missions, combat UAVs will become the preferred weapons platform. The introduction of such technologies and systems will present flight control system engineers with interesting design, development, and certification challenges.

Engines are also an active topic for research in aerospace controls. Propulsion subsystems such as combustors and compressors have the potential to exhibit improvements in performance, reliability, and reduced emissions by integrating control into their design. Specific research initiatives that are under way include the following:

- Distributed, fault-tolerant engine control is being explored for enhanced reliability, reduced weight, and optimal performance, even with deterioration in the system and its components. The use of smart sensors and actuators together with advanced robust and adaptive control methods is being explored.

- Advanced health management technologies for self-diagnostics and prognostics are yet another example where controls are playing an increasing role. In problems such as life-usage monitoring and prediction, data fusion from multiple sensors and model-based information are being explored.

- Control of flows at the inlet are being investigated so as to circumvent separation as well as stall. Current research areas are focused on the development of microactuators that can provide a distributed multitude of inputs such as pressure, velocity, and fuel-to-air mixture; arrays of pressure and velocity sensors; models that capture the underlying spatiotemporal complexity with the available computational resources; and the corresponding distributed control strategies that can guarantee robust and optimal performance.

The exploitation of future space systems for civil, commercial, scientific, and space exploration also gives rise to a set of challenges and opportunities in the area of control. With the rapid advances in computing, communications, and sensing technology, three main categories of guidance, navigation, and control (GNC) systems can be defined:

- Low-end (recurring) GNC systems are often incorporated in existing multimission (EO applications) or commercial platforms (telecom applications). Industrial competition in the global space market drives the need for permanent reduction of production cost and schedule. This low-end GNC system might require some level of innovation in the development of certain operational modes and vigorous research effort in improving the verification and validation process.

- High-end GNC systems are generally required for satisfying challenging control performance requirements (such as pointing accuracy, pointing stability, safe precision landing, space object interception). Future space missions requiring such GNC systems are listed in Table 1. The high-end GNC systems often rely on innovative designs in the area of navigation, guidance, and control technologies. In some cases, increasing levels of autonomy are required in order to meet mission requirements.

- Safety-critical GNC systems include mainly launchers and manned space transportation systems. For example, the Automated Transfer Vehicle (ATV) GNC system falls in this category due to proximity operations and docking with the International Space Station (ISS).

Table 1. Proposed GNC System Classification for Space Applications

GNC System Class	Past/Present Missions	Ongoing Missions	Future Missions
Low-end (recurring)	Earth Observation (multi-mission platform) Telecommunication (commercial platform)	Navigation (Galileo) Small Telecom Satellite (SmallGEO)	Affordable low-earth orbit (LEO) platform Agile small LEO platform (300-kg class)
High-end	Earth's gravity field (GOCE) Comet rendezvous and lander deployment on the surface (Rosetta) Astronomy (Hubble Space Telescope)	Astrometry (GAIA) Astronomy (James Webb Space Telescope) Fundamental Physics (LISA) Planetary Entry Descent and Landing System (Mars Science Laboratory) Planetary rover	Jovian mission Interferometry mission (formation flying) Sample return mission (moon, Mars, asteroid) Solar power satellites
Safety-critical	Launcher (Ariane 5, Delta, Proton) Shuttle Resupply cargo (ATV)	Launcher (Vega, Ariane 5, etc.)	Next-generation launcher (AR6, Vega Evolution) Moon cargo lander Space tourism (suborbital and orbital) Commercial in-orbit servicing

For each GNC system class, Fig. 3 provides some examples of enabling control technologies. Synergy with other terrestrial applications is also indicated.

Opportunities for Research

Recommendation No. 1

Development of advanced analysis, verification, and validation technologies (theory, methods, and engineering tools) for supporting the certification of autonomous aerospace systems and systems of systems (SoS) and for reducing the "time to market" and associated development effort. The focus shall be on but is not limited to:

- Development of new worst-caseanalysis techniques for hybrid and nonlinear systems
- Enhancement of statistical approaches
- Improvement of transparent robust control design methods
- Development of "trouble-shooting" control techniques
- "Smart" interpretation and presentation of results

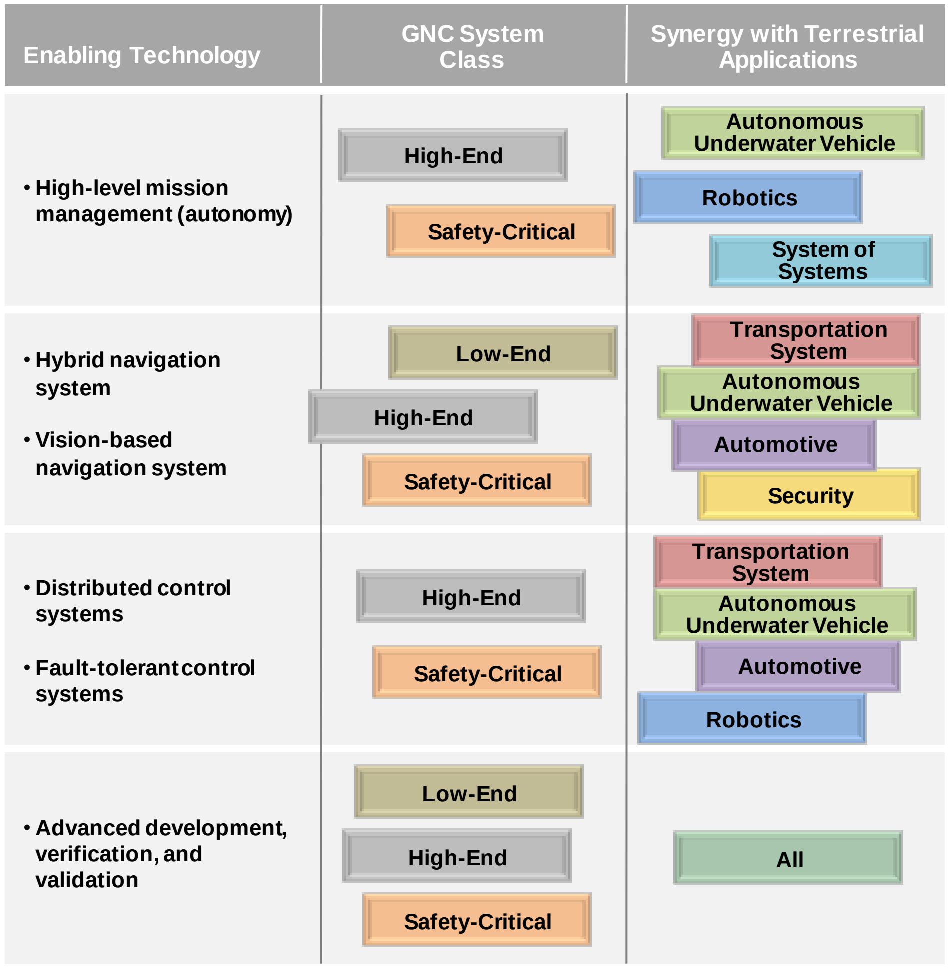


Figure 3. Examples of enabling control technologies for each GNC system class.

Recommendation No. 2

With the rapid trends toward autonomy (space exploration, UAVs, and virtual co-pilots), revolutionary control solutions need to be developed in order to deliver (high-performance) robust outer loops and to support advanced capabilities such as situation awareness and avoidance. The technology focus shall be on but is not limited to:

- Sensor fusion
- On-line trajectory and optimal path planning
- On-line system identification
- Robust fault detection, diagnosis and prognosis
- Decision making
- Adaptive reconfiguration control
- On-line planning and executive decision making

Recommendation No. 3

With emerging system-of-systems applications such as air traffic control, space interferometer, and swarms of UAVs, transformational control technologies are required to meet the new challenges:

- Numerical modeling of complex multisystems and their validation
- Information transmission over networks
- Decentralized control and decision making
- Subliminal control
- 4D trajectory planning
- Self-separation
- Conflict detection and resolution

Conclusions

The development of commercial and military aircraft and space vehicles is impossible today without flight control systems or guidance, navigation, and control systems. Industrial competition in the global aerospace market drives the need for continuous improvement of capabilities as well as reducing development and production cost. Control technologies will continue to have an important role for the successful realization of the next generation air transportation systems, including air traffic management and the vehicles that operate in this system. As our quest for knowledge continues to grow, control technologies will also play an important role in pushing back the frontiers of space exploration and protecting and securing the environment by gathering more accurate satellite data.

Acknowledgments

The authors wish to acknowledge John T. Bosworth, R. Lozano, and R. Tempo.

Selected recommendations for research in aerospace control:

- Autonomy is a key trend; for its promise to be realized, new control system architectures, high-performance robust outer-loop control solutions, and situation awareness and avoidance technologies must be developed.
- Advanced analysis, verification, and validation technologies for supporting certification and reducing development effort are essential for industrywide deployment of advanced control.
- Several “systems-of-systems” opportunities are emerging in aerospace that require transformational control technologies to be developed.

References

[1] James E. Tomayko. "Computers Take Flight: A History of NASA's Pioneering Digital Fly-By-Wire Project," NASA-SP-2000-4224. [Online}, 2000. Available at http://www.nasa.gov/centers/dryden/pdf/182985main_DFBW_rev1.pdf.

[2] G. Pigne. "Ariane 5 and Ariane 5 Evolution GN&C Overview," 34th COSPAR Scientific Assembly, The Second World Space Congress, Houston, TX, 10-19 October 2002.

[3] Teal Group. World Unmanned Aerial Vehicle Systems—2009: Market Profile and Forecast. Farifax, VA: Teal Group Corporation, 2010, www.tealgroup.com.

[4] ADS-EUROSPACE. "The European Space Industry in 2008: facts and figures" [Online], 2009. Available at http://eurospace.pagesperso-orange.fr/ffdata2008_web.pdf.

[5] Euroconsult. "Satellites to be Built & Launched by 2019, World Market Survey" [Online], 2010. Available at http://www.euroconsult-ec.com/research-reports/space-industry-reports/satellites-to-be-built-launched-by-2019-38-29.html.

[6] NextGen. "FAA's NextGen – Implementation Plan" [Online], 2010. Available at http://www.faa.gov/about/initiatives/nextgen/media/NGIP_3-2010.pdf.

[7] SESAR. "Today's Partners for Tomorrow's aviation" [Online], 2009. Available at http://www.sesarju.eu/sites/default/files/documents/reports/SESAR_Bochure.pdf.

[8] NASA. FY2011 Budget Estimate by Section: Aeronautics Research [Online], 2010. Available at http://www.nasa.gov/news/budget/index.html.

Related Content

The Impact of Control Technology report also includes more than 40 flyers describing specific "success stories" and "grand challenges" in control engineering and science, covering a variety of application domains. The ones below are closely related to the topic of this section.

Success Stories

- Automated Collision Avoidance Systems – *C. Tomlin and H. Erzberger*
- Control of the Flexible Ares I-X Launch Vehicle – *M. Whorton*
- Digital Fly-by-Wire Technology – *C. Philippe*
- H-infinity Control for Telecommunication Satellites – *C. Philippe*
- Nonlinear Multivariable Flight Control – *J. Bosworth and D. Enns*
- Robust Adaptive Control for the Joint Direct Attack Munition – *K.A. Wise and E. Lavretsky*

Grand Challenges

- Control of Combustion Instability – *A. Banaszuk, A. Annaswamy, and S. Garg*
- Energy-Efficient Air Transportation – *J. Alonso et al.*
- Verification, Validation, and Certification Challenges for Control Systems – *C. Philippe*

These flyers—and all other report content—are available at http://ieeecss.org/main/IoCT-report.

Control in the Process Industries

Ian Craig, Chris Aldrich, Richard Braatz, Francesco Cuzzola, Elom Domlan, Sebastian Engell, Juergen Hahn, Vladimir Havlena, Alexander Horch, Biao Huang, Maryam Khanbaghi, Alkis Konstantellos, Wolfgang Marquardt, Tom McAvoy, Thomas Parisini, Stratos Pistikopoulos, Tariq Samad, Sigurd Skogestad, Nina Thornhill, and Jie Yu

Introduction

Process control[1] is in many respects a mature technology serving mature industries.[2] It has gone through the emerging phase, the growth phase, and some would argue that it has also gone through the mature phase and is now in decline. The shares of companies operating in industries where process control is widely used, such as the petroleum industry, show typical signs of maturity—high dividend yields and low price-earnings ratios that reflect limited growth prospects.

The maturity of process control technology is also borne out by the decline in research funding for this area over the last decade or so, especially in the U.S. Paradoxically, this decline has occurred precisely because process control research has been so successful in addressing industry concerns. Although PID control been the king of the regulatory control loop for many decades, advanced process control has over the last few decades moved beyond the laboratory to become a standard in several industries. Many vendors now routinely offer advanced solutions such as model predictive control (MPC) technology, with its ability to economically optimize multivariable, constrained processes. Although there is always room to improve upon existing control solutions, it becomes harder to make an argument for research funding if vendors can adequately address most of their customers' control problems.

In contrast to the situation described above, government funding for process control research is readily available in Europe, mostly in the form of industrial-academic collaboration. For example, all German chemical companies have grown their process control departments considerably in recent years. The process control market will remain significant, and process control researchers still have much to offer the process industries. However, researchers will have to get out of their comfort zones for research funding levels to be maintained or increased. For example, there is much room for traditional tools to

[1] Process control refers to the technologies required to design and implement control systems in the process industries. The goal of process control is to bring about and maintain the conditions of a process at desired or optimal values. Process control technologies include physical and empirical modeling, computer simulation and optimization, automation hardware and software (such as actuators, measuring instruments, implementation platforms, plant communication infrastructure), control structure design, advanced control strategies and related technologies such as process monitoring/diagnosis, and planning/scheduling solutions.

[2] Process industries are those in which raw materials are physically or chemically transformed or where material and energy streams may interact and transform each other. These include continuous, batch, or sequential processes and can refer to process units, whole plants, and enterprises. Specific industries include biological/biochemical/bio-fuels enterprises, cement, chemical, electrochemical, glass/ceramics, heating, ventilation, and air conditioning (HVAC), minerals and metals, petrochemical/refining, pharmaceuticals, power generation, pulp and paper, and water systems. Process components are also prevalent in other industries such as automotive, green buildings, microelectronics, and nuclear power.

be applied in nontraditional industries such as the biological/biochemical and pharmaceutical industries. Conversely, traditional industries increasingly require agile and dynamic enterprise-wide solutions. Comfort zones are stretched, as forays into nontraditional fields and enterprise-wide solutions both require domain knowledge that often does not form part of the traditional training of a process control researcher. For those willing to make the effort the potential rewards are huge. In nontraditional industries, one can find plenty of low-hanging fruit where simply applying basic control tools can have substantial impact. Similarly, the economic benefit of control at the enterprise level could dwarf that obtained from improved unit process control.

Successful Applications of Control

The impact of process control can be viewed from two main vantage points: technical and economic. The main technical impact occurs in cases where operation would not be possible without control, such as when control is required to stabilize an unstable process. Few such examples currently exist in process control (for example, level and anti-slug control), but the number may increase with the advance of process intensification efforts. For stable processes, the technical impact of process control is to improve dynamic response and reject disturbances. Once the basic control infrastructure is in place, advanced process control is one of the technologies most often used for improving economic performance. The ideal situation, however, is where the plant is designed to be easy to control, for example, by providing for adequate actuator authority to deal with dynamic disturbances. Unfortunately, plants are often designed taking primarily a steady-state view. Once the plant is designed and working, process control provides the means for maximizing production and product quality.

> *Advanced control is one of the technologies most often used for improving the economic performance of a process plant.*

The economic performance improvements resulting from advanced process control are often divided into various subcategories and expressed in term of percentages. For example, improvements achieved in various industries by a major vendor are:

- Increased throughput 3-5%
- Reduced fuel consumption 3-5%
- Reduced emission levels 3-5%
- Reduced electricity consumption 3-5%
- Reduced quality variability 10-20%
- Reduced refractory consumption 10-20%

Single-loop PID control structures based on hierarchical time decomposition (cascades) still dominate most process control applications. PID controllers are successful because they work reasonably well in most applications. They are generally implemented without the need for a process model and are relatively easy to tune according to well-established tuning rules, either manually or automatically, using software supplied by all the major vendors of process control platforms. The main impact of PID control is technical.

Advanced control that optimizes PID loop setpoints is usually required to make a significant economic impact. The most successful advanced process control technology has been MPC. In the petrochemical industry, for example, MPC is often combined with online optimization of the setpoints on the basis of large, rigorous nonlinear stationary plant models. This practice is known as real-time optimization (RTO). One major vendor has reported more than 1,000 MPC implementations, and since its commercialization in the early 1980s MPC has become a standard in various industries.

> *Since the 1980s, model predictive control has become a standard in various industries.*

Other high-impact process control success stories include inferential sensing using, for example, Kalman filtering; process modeling and identification tools for the estimation of process variables that are too difficult/expensive or impossible to measure online; automatic fault detection and diagnosis and statistical performance monitoring using multivariate statistical methods—very important where the number of control loops per control engineer has escalated rapidly over the last two decades; modeling and process simulation tools for rapid prototyping; and systems methodologies and analysis tools to deal with highly complex processes.

An industry breakdown of successful process control applications is given below. Some of the application highlights and their impact are:

- Industrial energy control for CO_2 footprint reduction (for example, \$4M in savings for one particular application).
- Refining and petrochemicals (typical cost reductions of 10-20%).
- Steel rolling mill tension control (for example, €100K savings just in plant building costs).
- Boiler startup optimization using nonlinear MPC (reported to save around 15% of energy at each startup in numerous installations).

Pharmaceuticals

Standard feedback control (such as PID control) and multivariable statistical methods have been applied in most steps of the pharmaceutical manufacturing process, but not routinely and usually not using more advanced methods. Applications of multivariable statistics have seen huge growth in the pharmaceutical industry, and the growth of feedback control applications has been extensive, but starting from a low level.

Robust nonlinear feedback control has been applied in the crystallization of pharmaceuticals in many companies and has been well documented in the literature. Financial, safety, and environmental quantification of impact is difficult because companies are purposely secretive about this information.

Mining, Minerals, and Metals

Several control technologies besides PID have been experimented with in the metals industry. Some are now considered established technologies with significant advantages over classical PID control, such as:

- MPC: automatic flatness control, temperature control based on finite-element heating models.
- Optimal Linear Quadratic Gaussian (LQG) control: tension control based on online section measurements, hot strip mill combined control of looper and stand.

- Kalman filtering: automatic gauge control (AGC), automatic camber control.
- Neural-networks-based control: waterbox control for steel microstructure identification as a function of casting properties and rolling temperature.

Although it is difficult to quantify the impact of these applications in general terms, some examples provide some insight:

- Tension control in a rolling mill for long products: Advanced multivariable control exploiting section-measuring devices, beyond performance and dimensional quality effects, allows for a much more compact mill, saving more than €100K just in building costs for the plant.
- Control of a reheating furnace: Advanced control results in a 5% reduction in gas usage (a typical plant produces 500,000 t/year with a gas consumption of approximately 65 Nm^3/t).
- A copper concentrator scheduling solution running successfully since 2005 has reportedly increased throughput by 1-2%.
- The application of automated plantwide process monitoring schemes on a mineral processing concentrator circuit has decreased response times from as much as six weeks to a maximum of three days. This has led to substantial savings in terms of reducing product losses.

Chemical/Petrochemical

Key to the success of advanced control technologies in large-scale continuous processes such as petroleum refining is the ability to "model" and "optimize" (online or offline) the process and then build suitable MPC strategies around this optimized model. Typical studies lead to improvements on all fronts, including cost reductions of 10-20%, increased safety margins, and reduction of emissions (up to 70%) under varying conditions (guaranteed operability envelopes).

Advanced control seems to have the greatest impact in refining and petrochemicals because the margins can be low and every last bit of performance has to be squeezed out. Less advanced control has been applied upstream on oil platforms because there the driver is throughput.

Long-term applications of MPC in the chemical and petrochemical industry include crude towers, other distillation columns, gas separations, fluid catalytic cracking units, hydrocrackers, and polymerization reactors. The economic savings have been in the hundreds of millions of dollars. Other process control technologies that have been employed successfully are RTO, multivariate statistical process control (SPC), controller performance assessment and monitoring, plantwide loop oscillation detection, closed-loop identification, and soft sensors.

Extended Kalman filters are used extensively to construct observers that monitor and control batch processes, especially polymerizations.

Other

MPC applications are expanding from chemical/refining plants to industrial energy and public power generation utilities. One example is a new Advanced Energy Solutions product developed by Honeywell and applied in plants in Europe, Africa, and Asia.

Smith predictors, MPCs, Kalman filters, fuzzy logic controllers, and nonlinear control in the form of state feedback linearization have been implemented in the glass and ceramic industries. The first immediate and obvious impact was process variability reduction. In two cases where the financial impact had to be calculated, the cost reduction was a few million dollars.

LQG, robust, and H_∞ control trials have been performed in experimental fusion main plasma control (in the international ITER project).

Standard feedback control (such as PID control) has been applied routinely in electrochemical processes, such as lithium-ion batteries, for many years. More advanced methods for feedback control and for integrated design and control still need to be employed.

Applications of multivariable decoupling control solutions have become standard in many pulp and paper mills. Plantwide model-based optimization of paper mills has also been reported. In an application of robust multivariable control design to the cross-direction control of paper machines, an approximately 80% reduction in control tuning time and up to 50% higher performance have been reported.

Model-based control and optimization solutions in cement production have provided significant savings in more than 300 installations worldwide.

Additional emerging industries for the process control community include renewable energy, some types of biological processes, molecular and nanotechnology, and megascale processes.

Investment in advanced process control (APC) in industries where it has not been used previously could proceed as follows. Controller performance monitoring tools could be used as a door opener for APC by showing that the reality is not as good as plant managers believe or claim. Once installed, only a few highly successful APC installations would be required to potentially open up a whole industry for APC, which could then go on to become an industry standard. Sound economic performance assessment methods play a key role in justifying investment in APC, especially in groundbreaking implementations.

Market Sizes and Investment

Modern process industries cannot operate without process control. The process control market size is therefore a percentage of the process industries' market size. To understand the significance of this percentage, consider that the investment in instrumentation and control of a greenfields project is about 5-8% of the overall plant. In addition, a significant install base of regulatory and advanced controllers must be maintained to achieve production targets. The remaining component making up the total process control market size is the implementation of new (advanced) controllers on existing plants.

One way of estimating the process control market size would be to examine the turnover of the major process control vendors. A difficulty with this approach is that these vendors often also serve non-process industry markets but report turnover figures at the holding company level. The process control market size should therefore be estimated by stripping out the non-process industry contributions if reported separately.

A recent estimate of the world market for process control can be obtained from a report published in 2009 by the European Union [1]. According to this report, the 2007 world market for "Monitoring and Control" was about €188B, with Europe's share estimated at €62B. The 2007 world market for process-control-related industries was about €26B, with Europe's share estimated at €10B. Of the world market for the process industries, €5B was for equipment, €4B for software, and €17B for services. Growth was projected at 6.9% per annum.

Research Funding and Employment Summary

Government

Government funding for process control has been low, especially in the U.S. The exception has been in applications to new industries such as semiconductors or fuel cells, but such programs are usually of limited duration. The low governmental funding level is a bit surprising in light of the relatively high industrial interest and importance.

The European Union (EU) has set aside €10M/year specifically for control research; about 33% of this figure is for process control. Approximately ten times this amount is available for process control research through other EU grant programs. Government funding is available in Europe, and particularly in Germany, mostly in schemes with industrial-academic collaboration (industry's share is partly supported; the academic partners are fully subsidized).

In the UK, convincing funding councils to give any priority to industrial process control is difficult because most research funding is geared toward topical issues such as systems biology or CO_2 capture. The situation is different in Sweden (for example, the new Process Industries Centers at Lund and Linköping) and in Canada, especially in Alberta, where sustainability and responsible use of tar sands are very pressing concerns.

Employment

Of concern is that a decline in process control research funding at universities, as is happening in the U.S., inevitably leads to a decline in the hiring of academic staff and hence also graduate students working in this field. This in turn leads to a decline in the number of APC advocates in the process industries. Due to lack of knowledge about what APC can achieve, many plants are likely to operate suboptimally. Conversely, it takes only a few highly successful and well-publicized implementations for APC to become an industry standard, as has occurred in the cement industry.

Pharmaceuticals

Pharmaceutical companies are investing heavily in modeling and control technology within their companies but are only employing and supporting a small number of control engineers in academia. Investment in control has been increasing and is expected to continue to do so.

The market size for the pharmaceutical industry is many billions of dollars, and the growth rate for the industry has been between 10% and 20% per year. Thus, pharmaceuticals is a large potential market for control technologies. Progress has long been hindered by the need for recertification if changes in the control regime were made—once a production process for a medicinal product was validated and licensed, it was effectively “locked”—but this has come to an end as a result of the process analytical technologies (PAT) initiative of the U.S. Food and Drug Administration (FDA).

Mining, Minerals, and Metals

Companies investing in applied control research is a key factor in the global competitiveness scenario. For example, some companies producing turnkey steel-making plants invest more than 8% of their overall budget in research, with control technology representing 30% of the research budget. Sensing devices, mathematical models, prediction systems, and simulation environments account for most of this investment. A conservative estimate of the process control market size in the steel-making domain is €1B.

As for funding agencies, the example of Regione Friuli Venezia Giulia in Italy is worth noting. In the last six years, the commission funded eight research projects in the steel-making domain where control and automation were the main subjects, with total funding of more than €1.5M.

An increasing number of mining companies are seeing process control as a strategic investment, particularly as a means of becoming more energy-efficient and improving their returns on investment in capital equipment. In the platinum industry, for example, both Anglo Platinum and Lonmin have recently increased their investment in research and development in plant automation and process control substantially. These investments are likely to grow, provided they can realize tangible benefits for these companies over the relatively short term.

Other

Little investment has been made in advanced control research for electrochemical processes such as lithium-ion batteries, and very little funding has been available for academic researchers in this area. Given the size of the industry, the investment could grow rapidly if a control engineer found a “killer application” of advanced control that resulted in a major improvement in economics. The potential market for control for new electrochemical products is huge. Some countries are trying to obtain up to 30% of their total energy needs through renewable energy, and that cannot happen unless advances are made in solar cell manufacture and battery designs (to handle the increased need for load balancing when wind and solar power are used).

The IChemE held an industry-academia event in 2008 that addressed some issues raised here. An abbreviated event report is available [2].

Future Challenges

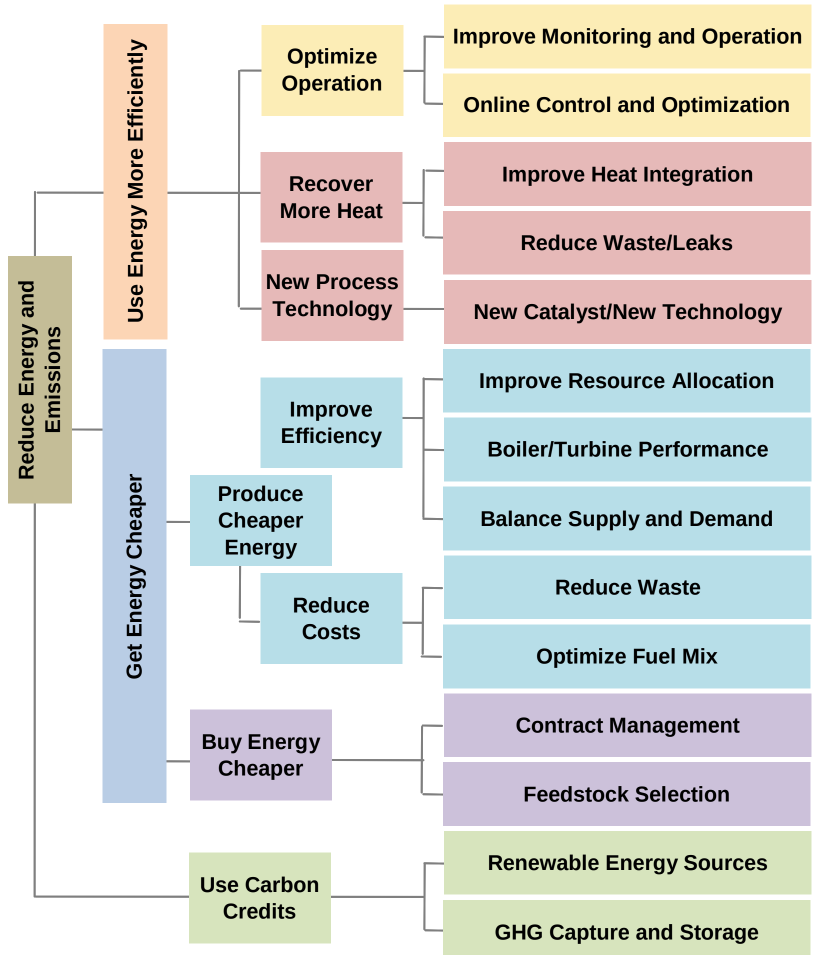
Energy Efficiency in Industrial Processes

Mitigating climate change is perhaps *the* grand challenge of the 21st century as highlighted by the 2009 Copenhagen climate change summit. The process industries are major users of energy as well as major emitters of greenhouse gases (GHG). Improving energy efficiency in industrial processes will be key to reducing GHG emissions and will become the major focus of the process industries. The manufacturing optimization and energy consumption reduction that will be required can open new doors to control.

Energy efficiency has always been an important consideration in advanced process control and is often explicitly included in optimization and control objective functions. The nexus with GHG emissions has further elevated its importance. More holistic perspectives on energy use and emissions reduction in industrial processes are being sought. As one example, Fig. 1 depicts a classification of automation and process technology measures primarily focused on refineries. Control and related technologies are well represented.

Very-Large-Scale Integrated Process Control (VLSIPC)

The use of economic-performance-optimizing MPCs (in contrast to setpoint tracking) is considered a strong trend for the future. Such control, known as dynamic real-time optimization (D-RTO), is applied to process and energy systems that are typically modeled by a large number of nonlinear differential-algebraic equations. D-RTO can be viewed as a variant of nonlinear MPC with an economic objective. Such integrated optimization-based control systems will be implemented in a complicated multiloop, hierarchical, and decentralized architecture to cope effectively with the network character of such systems. To implement VLSIPC systems successfully, the process control community must take a broader view. Its target should be the economic performance of a technical system that is implemented by the plant, the monitoring and automation system, and the human operators and decision makers in the face of process and model uncertainty.



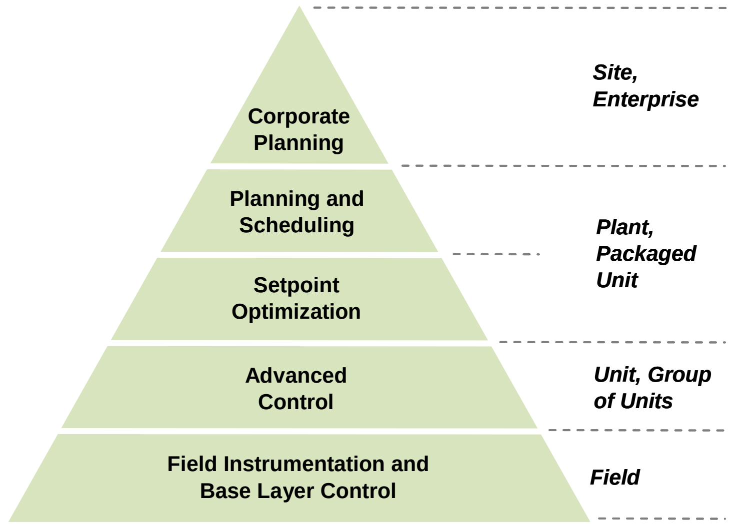
Source: Brendan Sheehan, Honeywell Process Solutions

Figure 1. An analysis of measures for reducing energy and emissions in refineries. Advanced optimization, control, and monitoring technologies are crucial in several cases.

The ISA S95 and ISA S88 standards provide a framework that makes it easier to integrate enterprise resource planning (ERP) systems, manufacturing execution systems (MESs), and distributed control systems (DCSs), paving the way for automating entire businesses. Such systems will be better able to deal with increasing volatility in production, energy, and raw material availability/requests/pricing by helping to make the business more agile. The integration of areas that are currently operated quite independently, such as operation and maintenance, scheduling and control, and energy management and production, is becoming a distinct possibility. A related challenge is to design such integrated systems to be easy to maintain through, for example, monitoring solutions that provide meaningful and easy-to-understand recommendations.

Classical approaches to DCS design and deployment (including all the layers from instrumentation and regulatory control to the advanced layer covering APC/RTO and the business layer covering MES and ERP; see Fig. 2) have reached their limits. Deploying and maintaining the next-generation DCS solutions that can respond to a volatile economic environment in real time will require that most of the activities be automated. Current solutions, however, are still static. DCS management tools for consistent cross-layer real-time responsiveness to changes in, for example, process topology and the business plan should become part of the system.

Related theoretical and practical challenges are methods for structuring automation solutions into layers and decentralized control tasks in order to solve large-scale online optimization and control problems in the presence of plant-model mismatch. Tools are required for multiscale modeling and control systems design, processes with many orders of magnitude ranges in time and spatial scales, “controlling” the sensor-to-enterprise-to-supply-chain, and integrating process control, design, and operations.



Source: Wolfgang Marquardt, RWTH Aachen

Figure 2. Layers in a process enterprise. Cross-layer integrated control solutions are needed.

One example of a “mega” industrial process that requires the development of larger, more complex control systems and solutions is the Shell Pearl gas-to-liquids (GTL) project in Qatar [3]. This integrated project comprises two offshore platforms, each with 11 production wells, two multiphase pipelines, and an onshore processing complex. When complete, Pearl will produce multiple finished products, including enough fuel to fill over 160,000 cars a day and enough synthetic base oil each year to make lubricants for more than 225 million cars. The control system for Pearl is correspondingly large. The control room comprises almost 1,000 control cabinets hosting 179 servers, programmed with 12 million lines of software code. Almost 6,000 km of control wiring extends throughout the plant. Advanced control solutions will need to be correspondingly larger scale and more complex. For such mega projects, the total number of sensors and actuators will be in the hundreds of thousands.

Other Application Challenges

New possibilities for control exist in areas where manual control is still heavily used. One possible application is drilling for natural gas or oil, which is extremely costly, especially when done offshore.

A promising application related to energy savings and carbon emissions reduction is climate control in buildings. Although this application is not strictly speaking process control, members of the process control community perhaps possess the most relevant knowledge.

Most processes in the electrochemical industries offer promising opportunities for a control engineer to make an impact. An example would be lithium-ion batteries, which have huge built-in inefficiencies that limit their widespread application in the automotive industry unless government subsidies are available. Other examples would be micro fuel cells and high-efficiency solar cells, where a major limitation has been the inability to manufacture reliable products (low yield). The potential impact of control in these areas could be significant.

Several problems exist in the steel-making domain where the use of advanced control methodologies may have a significant impact, such as looperless tension control in rolling mills for long products and electric arc furnace model-based control.

Process intensification efforts, such as Eastman Chemical collapsing an entire chemical plant into one highly integrated process unit, could introduce intentional unsteady-state operation, providing a research opportunity for the process control community.

Opportunities for Research

The application challenges discussed above provide excellent opportunities for research. Some of these research needs are discussed below.

Generating Good Process Models at Low Cost

This pursuit may sound like an oxymoron, but many consider it to be the holy grail of control—especially in the process industries where fundamental models are often extremely complex and expensive to obtain. A method of developing good low-cost process models could open up a host of new applications for APC. New plants are increasingly using training simulators that include faithful models which are valid over large operation ranges. Deriving "low-cost" control-relevant models from these simulated models is a distinct possibility.

Controller Design for Models Described by Nonlinear Partial Differential Equations (PDEs) and Integro-Partial Differential Algebraic Equations (IPDAEs)

A key theoretical challenge in nearly all process industries is the lack of nonconservative methods for the design of robust controllers for models described by nonlinear PDEs and IPDAEs. Although the finite-time case has been largely solved in recent years, at least for processes with reliable numerical simulation schemes, the infinite-time case is wide open. Much progress has been made in robust control analysis and design for infinite-time linear PDEs, but the amount of conservatism in terms of performance suboptimality and robustness margins has not been thoroughly characterized for many classes of control design methods.

The steel-making domain requires very complicated and detailed models that describe complex physical phenomena, such as steel tandem rolling. The behavior of the dimensions during rolling depends in a very complex way on many aspects/parameters of the process. Mill models typically merge PDEs with look-up tables and nonlinear algebraic relationships. The many approximations that are currently introduced to control the process clearly limit what can be achieved in the control phase. The impossibility of tuning the physical models using indirect information such as forces and torques makes the models unsuitable, and operators tune the controller by trial-and-error.

Control Structure Adaptation

An optimal control strategy structure based on the process model and available inputs and outputs should be employed at all times. Determining the optimal structure for a control strategy is currently part of the "art" of process control, based mostly on experience. Tools are required for solution configuration, a higher level of autonomy, and automated recovery from faults (such as loss of manipulated variables). Control structure adaptation has been used with significant impact in aerospace (reconfigurable flight control with redundancy of control surfaces/actuators). The role of similar investigations in process control must be explored.

New Sensing Technologies

Video cameras and related image-processing techniques look very promising, and applications have been reported by industries that have traditionally lacked online sensors, such as food processing, minerals processing, and pulp and paper. Video sensors can be used for both fault detection and real-time quality control.

Other Remarks

This section focuses on barriers to the application of APC in the process industries.

Technical Barriers

The process control infrastructure in a plant is often insufficient for the implementation of APC. Examples could include insufficient CPU capacity and lack of measured variables, actuators, and actuator authority. Process limitations include too many unknown parameters or a process that is too nonlinear for a "standard" APC solution to be applied.

Obtaining good process models for control remains a significant challenge, especially for processes with high complexity and dimensionality. A significant gap still exists between the capabilities and tools for developing high-fidelity models (in dynamic simulation tools, for example) and the ability to derive "intelligent" advanced model-based controllers from such models in a seamless way. Modeling of processes that contain huge numbers of solid particles in the presence of a liquid, as found in the pharmaceutical industry, are usually much more challenging to model and control than processes that only involve liquids and gases. Such models require a deep understanding of transport phenomena, physical chemistry, and nonlinear PDEs.

Other technical barriers include the lack of robustness/adaptability of control solutions, APC solutions that are difficult to engineer and maintain, difficulty in justifying/measuring economic improvements, poor infrastructure in developing countries, and long distances between the site and the technical office.

Workforce Barriers

The workforce is divided into process control engineers who design, commission, and maintain APC systems and process engineers and operators who are the users of process control solutions.

Barriers for the application of APC among process control engineers include lack of domain knowledge and the inability of most to apply the latest control technology. Domain knowledge is often lacking in the control community, which limits the ability to generate high-impact applications of feedback control theory. Rigorous process modeling relies to a great extent on domain knowledge, which requires process control researchers to get out of their comfort zones. Stationary simulations, which are of limited use for controller design, are commonly used in process design, with dynamic modeling and simulation only done if needed. This practice leads to a lack of skilled personnel.

Only a very limited portion of the control community is able to apply the latest control technology, partly because the level of rigor in technical education is continually decreasing. Applications of model-based control require multidisciplinary skills (control and domain knowledge) and are still a mixture of both art and science. Despite the difficulty of replacing required expertise with tools, most process control vendors invest heavily in tool development, and these tools are becoming a major differentiator among the APC solution providers (Honeywell, ASPEN, and others). Completing huge software engineering projects that adhere to a high standard seems to be a limiting factor in getting applications implemented. The problem is partly due to human resource factors, as when skilled personnel leave and carry away knowledge of the details in their heads.

Barriers for the application of APC among process engineers and operators include lack of advanced control knowledge. The lack of advanced control knowledge among plant personnel is pervasive and results in suboptimal operation and maintenance of APC solutions. The reasons are manifold and include the retirement of skilled staff, high staff turnover, and the employment of new “unskilled” replacement staff. This problem can partly be addressed by making advanced control solutions easier to engineer, operate, and maintain, and by process sites allowing remote monitoring and maintenance by centrally located skilled staff. Industry could live with 95% optimality but not with five different optimization tools in one plant.

Education and knowledge transfer for process control personnel are important for sustaining the success of control applications and for identifying new opportunities for control. The introduction of template solutions with grey-box ID-based tools facilitates knowledge transfer from the development team to the application engineers. New opportunities for traditional APC solutions can be found in upstream processes in the pharmaceutical industry, namely, the organic chemical reactors and separators prevalent in the chemical industry, and process engineers with sufficient control knowledge will be able to identify such opportunities.

Cultural Barriers

Many control engineers incorrectly perceive that their control toolbox is generic enough that the best control design involves simply selecting the right tool from the toolbox. This misperception often leads to a control solution looking for an application, instead of the reverse. Generic control tools usually cannot be applied to the most challenging control problems (it could be argued that this lack of ability to apply generic tools is what defines challenging control problems). In that case, a deep understanding of processes is required to produce a control design method that is robust and reliable for all members of a particular class of process. This cultural problem can be seen in control engineers no matter what their engineering discipline, even though they are supposed to be experts in the processes of their discipline.

Another cultural problem is that many control engineers in academia typically do not formulate feedback control algorithms with the degree of robustness to uncertainties and insensitivity to disturbances needed for implementation in the pharmaceutical and biomedical industries. A hiccup in their processes can result in injuries, casualties, and hundreds of millions of dollars in recalls, lawsuits, and process redesigns. A feedback control algorithm that performs well only 99.99% of the time or requires occasional manual retuning to function 100% of the time is useless in those industries.

Control is still widely viewed as a “service” activity rather than a “core” activity and thus is often not fully appreciated by management. Having a few good reference implementations of advanced control makes selling new projects much easier. Economic justification is crucial for new applications of advanced control.

Cultural differences exist all over the world. In Europe, plants run well already and the awareness of optimization is widespread. However, no investment will be made without good justification. In greenfields plants, it is easier to “add” advanced solutions as long as there is sufficient justification for doing so. Again, having good references makes it much easier to convince clients to implement advanced control.

A significant disconnect often exists between academia and industry. Unfortunately, control at many universities has become more of an applied mathematics subject than an engineering subject. Students are often not prepared to get into the workforce and solve real control problems. Conversely, process control users often regard universities as specialized APC vendors and require a quick project turn-around that is not conducive to fundamental research. However, research activities in process control need to be more focused on addressing industrial problems rather than pure theoretical advances.

Regulatory Barriers

Stricter regulations play a key role in tightening product quality specifications, which normally lead to increasing demand for APC. Conversely, when a clear roadmap for the necessary legal and economic environment is lacking (for example, for CO_2 footprint reduction, energy efficiency, and life-cycle considerations), companies are often not compelled to invest in advanced control. Regulations that are too strict can, however, hinder progress, as was the case in the pharmaceutical industry before introduction of the PAT initiative by the U.S. FDA.

Conclusions

The process industries have historically been a major beneficiary of advanced control solutions. PID auto-tuners, model predictive control, and real-time optimization have all had a substantial impact on the cost, efficiency, and safety of process plant operations. Ironically, though, the success of these technologies is leading (or in some cases has led) to a perception of commoditization. If off-the-shelf solutions are now available, why is more R&D investment required?

In fact, process control remains a vital area of research with substantial opportunities for future impact:

- Although advanced control has been widely adopted in some process industry sectors, many other sectors are just starting to deploy such solutions—as evidenced by some of the applications noted in this section. All process industries have their individual characteristics, so methodologies and techniques must be tailored.

- Advances in control applications are driven not only by advances in control theory. Control solutions are enabled by other technologies, and developments in these areas open up new opportunities. Thus, new sensors (video cameras are a good example), wireless communications, broadband access to the Internet, and increasingly more powerful processors all present new opportunities for impact with control theories and algorithms.

- Control has gradually moved up the plant automation hierarchy from field solutions (PID), to multivariable systems, to higher level optimization. But opportunities for control do not stop there. Little work has been done in the area of integration with planning and scheduling, and especially with enterprise applications. Furthermore, the development of mega projects and the concept of a plant as a node in a larger supply-chain network suggest new horizons for the field.

Finally, while control scientists and engineers must continue to strive to overcome barriers, both technical and otherwise, increasing societal and industry demands for energy efficiency, reduced GHG emissions, more competitive operations, and greater automation and closed-loop responsiveness all promise increasing demand for advanced control and related technologies in the process industries.

Selected recommendations for research in control for the process industries:

- Methodologies, including algorithms, to develop control-relevant process models at low cost would open up application opportunities that are currently not cost-effective for advanced control.
- Determining the structure of a control strategy is an art today and needs to be developed to a science; rigorous control structure adaptation techniques are also needed.
- With increasing integration of process plants and their upstream and downstream connections, new research opportunities have emerged in wrapping closed-loop optimization and control loops around enterprises and supply chains.

References

[1] J. Pereira (ed.). Monitoring and Control: Today's Market and Its Evolution till 2020. Luxembourg: Office for Official Publications of the European Communities, 2009.

[2] Institution of Chemical Engineers, "The Chemical Engineer," p. 55, Feb. 2009.

[3] "Pearl GTL – An overview" [Online], 6 September 2010. Available at http://www.shell.com/home/content/aboutshell/our_strategy/major_projects_2/pearl/overview/.

Related Content

The Impact of Control Technology report also includes more than 40 flyers describing specific "success stories" and "grand challenges" in control engineering and science, covering a variety of application domains. The ones below are closely related to the topic of this section.

Success Stories

- Advanced Control for the Cement Industry – *E. Gallestey*
- Advanced Energy Solutions for Power Plants – *V. Havlena*
- Advanced Tension Control in Steel Rolling Mills – *T. Parisini and L. Ciani*
- Auto-tuners for PID Controllers – *K.J. Åström and T. Hägglund*
- Cross-Direction Control of Paper Machines – *G. Stewart*
- Ethylene Plantwide Control and Optimization – *J. Lu and R. Nath*
- Performance Monitoring for Mineral Processing – *C. Aldrich*

Grand Challenges

- Process Manufacturing Networks – *W. Marquardt and K. Frankl*
- Supply Chain as a Control Problem – *K.D. Smith et al.*

These flyers—and all other report content—are available at http://ieeecss.org/main/IoCT-report.

Automotive Control

Luigi Glielmo, Ken Butts, Carlos Canudas-de-Wit, Ilya Kolmanovsky,
Boris Lohmann, and Greg Stewart

Automotive: Control Systems in Millions of Copies

Here *automotive* refers to the application field of vehicles on tires, such as cars, trucks, and motorbikes and related infrastructures. (Although some undergrounds run on tires, such as in Paris, we did not include those vehicles and transportation systems! To avoid ambiguity, one could add the requirement of steerable wheels.) Because of the large number of consumers involved, and hence the economic significance of the entire supply chain, the automobile is the symbol of modern-era *homo sapiens,* certainly in developed countries and increasingly in emerging regions. Historically, the automotive industry was not a major user of advanced controls, but the situation began to change several decades ago with the advent of cheaper, smaller, and better embedded processors and other developments. Today control is pervasive in automobiles, and all major manufacturers and many of their suppliers have invested significantly in Ph.D.-level control engineers. Indeed, over the last decade or more, the automotive industry has become one of the foremost industry sectors in terms of the importance accorded to advanced control technology.

Over the last decade, the automotive industry has become one of the foremost industry sectors in terms of the importance accorded to advanced control technology.

On the one hand, because of the successes we will discuss later, mechanical engineering is now much more aware of the possibilities offered by combining mechanical design and control design than was the case just a few years ago. Hence the design of a new engine, or a new subsystem, is performed through dynamical simulations on powerful platforms where control algorithms can be included from the early phases. On the other hand, applying the control-based approach to complex systems is difficult, since control synthesis requires abstraction and usually simplifications that are not so obvious. Roughly speaking, one could say that a sound control-based innovation requires at least a good Ph.D.-level researcher working on it. To make things worse, we must consider that any development in the industry needs to be overseen from early conception to industrialization and maintenance over years; hence the control machinery, often captured only after adequate training, has to be somewhat translated and made understandable to all people in the workflow, a hard but crucial task for achieving widespread market penetration. Another way to put it, according to an automaker expert, is that model-based control design procedures are lengthy and difficult to include in a production schedule. Here are a few numbers to start with: for an engine management electronic control unit (ECU), more than 100 inputs and outputs need to be handled; some 100 system functions need to be implemented; some 1,000 pages of specifications need to be understood; some 10,000 parameters need to be calibrated (hundreds of kilobytes); some 100,000 lines of code need to be written (many megabytes) [1].

In addition, although the costs of sensors and actuators tend to decrease, the high volumes of production (millions of units) and the tight margins of this business suggest a careful evaluation of the return on investment before industrialization of a new control concept.

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

What are the societal goals in this sector? From a broad perspective, automotive transportation should be efficient, sustainable, and safe, partly because of environmental concerns and partly because of the large number of automobile-related fatalities and injuries worldwide. Economic considerations of manufacturers, suppliers, and consumers must also be taken into account. All aspects are then related to the vehicle itself and to the interrelation among vehicles and between vehicles and the transportation infrastructure.

As consequences of those goals, a list of top-level requirements can be delineated, such as higher engine performance, in terms of tradeoffs among power, fuel needs, and emissions reduction; increased reliability, safety, and passenger comfort; and a longer expected lifetime (strongly connected with a heavy reduction of maintenance and repair costs). Fulfilling these expectations also requires faster and cheaper vehicle development.

Successful Applications

The emissions reduction successes obtained in the automotive field are strictly related to control applications. For example, the operation of three-way catalytic converters, which managed to dramatically reduce emissions by spark-ignited engines, depends on the precision with which the mixture of fuel and air is close to stoichiometry. The ordinary mechanical carburetor could not achieve the necessary precision, and hence "electronic injection" was introduced, where the fuel is injected in precise quantities related to the amount of aspirated air. Typical control problems in this context are control of the injectors, an electromechanical device, and estimation of the air flow, which in early applications could not be directly measured. Interestingly, the idea of "injecting" fuel rather than letting it be aspirated is not a recent one; it was suggested by mechanical engineers at Bendix in the 1950s but failed commercially due to insufficiently robust components. The first solid-state systems were developed in the 1970s.

Now all new engine concepts, such as homogenous charge-compression ignition (HCCI), are mechatronic designs where the role of control is crucial. Another increasingly important control-based engine technology is variable valve actuation (VVA) or variable valve timing (VVT) (Fig. 1), which tends to detach engine valves closing and opening from the camshaft. This is crucial for cylinder-by-cylinder and stroke-by-stroke combustion control since it is possible to adapt the inflow and outflow of the air to the cylinders to the rotational velocity of the engine, the torque requested by the driver, and so on.

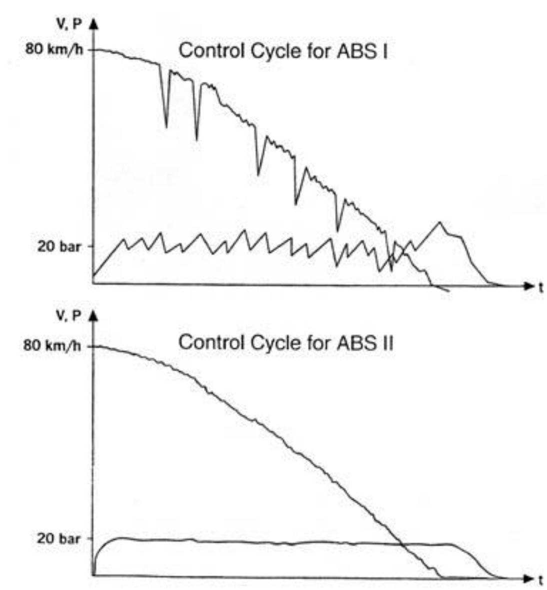
> *New engine concepts, such as homogeneous charge-compression ignition (HCCI), are mechatronic designs where the role of control is crucial.*

The control approach has gained wider visibility among nonexperts as a result of high-impact applications such as the antilock braking system (or *Antiblockier* system in German, ABS in any case) (Fig. 2), electronic stability control (ESC), and the automatic manual transmission. From another perspective, the success of control applications is apparent in that they become mandatory through specific legislation (as will soon happen for ESC) or, conversely, specific legislation calls for the development of control products to meet constraints, typically on emission levels.

Source: www.fiat.it

Figure 1. Fiat VVA valve opening schematic (MultiAir).

The X-by-wire concept, aimed at eliminating mechanical connections among components and facilitating the exchange of information among the various subsystems, has already resulted in successful applications or, in safety-critical areas where it cannot be totally applied, inspired improvements. For example, through drive-by-wire, the torque requested by the driver, originally implicit in the cabled throttle command, becomes a numerical value that can be transmitted to the engine ECU (as a reference value to be tracked) or the ESC (as a known disturbance input), thus enabling coordination of the subsystems and hence more effective operation of the entire system. The steer-by-wire idea is still considered audacious in commercial vehicles, but it can be found in simplified versions as active steering, where the mechanical connection between steering wheel and tires is kept but an electromechanical system enables additional turning of the wheels, possibly depending on the speed of the vehicle. Additional degrees of freedom can be added on four-wheel steering systems where, for example, the rear wheels can turn in the same or the opposite direction as the front wheels, depending on the kind of maneuver and the velocity. Power steering, once an ingenious hydraulic device, is now electromechanical. On some devices, for example, an algorithm controls the steering ratio; on others it modulates the assisting torque.



Source: www.bmwusa.com

Figure 2. Evolution of BMW ABS from version I to version II: Speed and braking pressure.

Note, however, that consumers do not always perceive the control part of the technology (often included in the more generic term *electronic*), although the word *control* has largely made its way to the general public, especially through driver assistance products. Obviously, success also depends on the introduction of new or cheaper sensors (for example, radars, lidars, and cameras for driver assistance).

As in other sectors, the control methodologies used have ranged from standard regulators to optimal control, in relation also to the availability of suitable models for the problem at hand. Sometimes the solution is reached through ad hoc, perhaps nonoptimal but effective, solutions and then improved with more sophisticated control methods. The list of applied control methods (from simple gain scheduling to proportional-integral-derivative (PID), or various forms of optimal control) has widened over the years with the dispersion of graduates in the industry. The list of new methods successfully developed in the automotive control field goes from hierarchical control structure (distributed on various layers) to gain-scheduled PIDs, passing through artificial intelligence control schemes such as neural networks and fuzzy-logic-based controls (to represent experts' knowledge). Virtual sensors (that is, subsystems and/or algorithms that exploit mathematical models to estimate process variables or operating conditions) play a large part in this scenario. Virtual sensors are widely used because they are cheaper than real ones (which often may not be physically feasible), can have a faster response compared to physical sensing devices, and can be more accurate (or better calibrated) than real sensors. Control methodologies such as Kalman filters, state observers, or online system parameter identification are successfully applied in designing virtual sensors.

Market Data and Socioeconomic Effects

Because control technology is hidden, collecting data on its market penetration is not straightforward, and often the data are extracted by taking a broader view. In a report by the European Commission, a relatively good close-up has been compiled for monitoring and control (M&C), therein defined as "the control of any system, device or network through automated procedures, managed by a control unit with or without the capability to display information" [2]. Eleven M&C application markets are defined: environment, critical infrastructures, manufacturing industries, process industries, buildings, logistics and transport, electric power and grid, vehicles, household appliances, healthcare, and home. The M&C market for vehicles "represents expenses by vehicle manufacturers for inside produced [vs. aftermarket products] vehicle embedded solutions. . . . World leaders are Bosch, Continental AG, Delphi, Denso, etc." The primary market for vehicles is represented by in-car systems, accounting for 95%; the remaining 5% represents vehicles such as aircraft, buses, trucks, and railways. The market's total world value exceeded €56 billion in 2007 (see Table 1), which is 28% of the total M&C world market (about €188 billion), and the European share is equal to about €17 billion, or 30% of the vehicle world market.

Table 1. World and European Vehicle M&C Markets (2007, in million Euros) [2]

Area	Hardware	Software	Services	Total
World	32489	2076	21842	**56407**
Europe	9873	631	6637	**17141**

Table 1 categorizes the market into hardware, software, and services. Tables 2, 3, and 4 illustrate how each of these groups is subsequently divided into solutions and list respective market values worldwide and in Europe. Note that the "control layer" is the largest value category. Aspects of control technology are also included in other categories.

Table 2. World and European Vehicle M&C Markets: Hardware (2007, in million Euros) [2]

Area	Control Layer	Interfaces Layer	Network	Computing Systems	OS and Drivers	Total
World	19534	2515	2515	2648	5277	**32489**
Europe	5936	764	764	805	1604	**9873**

Table 3. World and European Vehicle M&C Markets: Software
(2007, in million Euros) [2]

Area	Communication Software	Application and Visualization	Total
World	1017	1060	**2076**
Europe	309	322	**631**

Note: "Communication Software" and "Application and Visualization" do not seem related to core control engineering software [2].

Table 4. World and European Vehicle M&C Markets: Services
(2007, in million Euros) [2]

Area	Application Design	Integration, Installation, and Training	Communication and Networking	Maintenance, Repair, and Overall	Total
World	5148	5745	5720	5229	**21842**
Europe	1564	1746	1738	1589	**6637**

The report also suggested a market growth of 5.1% annually until 2020. This optimistic forecast was made in 2007, before the big financial crisis hit the car industry in 2008. Still, the numbers appear to be gigantic and suggest another way to look at future perspectives of automotive control applications: How many vehicles are on the road these days? *The Wall Street Journal* estimates the number at 800 million (counting cars and light trucks), compared to 650 million in 2000, and there are expected to be more than one billion by 2020—again, optimistic (or pessimistic for environmentalists!) [3]. Older cars have no electronic control units or lines of codes, but now even low-end cars can boast 30 to 50 ECUs governing windows, doors, dashboard, seats, and so on, in addition to powertrain and vehicle dynamics; luxury cars can mount more than 70 and as many as 100 ECUs. Analysts seem to agree that some 80% of all automotive innovations are driven by software. A recent article reports that "a modern premium-class automobile probably contains close to 100 million lines of software code" [4]. Some analysts disagree on the 100 million figure; others believe cars will require 200 to 300 million lines of software code in the near future. What seems undisputable is that the amount of software on a car is comparable with that on a civil aviation aircraft. Furthermore, according to the same article, the cost of electronics (hardware and software) in a vehicle now accounts for 15% of the total cost and can be estimated at 45% for hybrid electric vehicles (HEVs), where software plays a greater role—for example, the GMC Yukon hybrid automobile features a two-mode hybrid automatic transmission whose control software design took 70% of the total staff hours [4]. However, not all ECUs and software on board are related to control functions. One report estimates that about 36% of the automotive electronics market is not related to controls (but rather to security, driving information systems, and body). The part that is related to controls breaks down as follows: safety functions, 16%; chassis/suspension functions, 13%; and powertrain functions, 35% [5].

Another interesting way to quantify the impact of automotive controls relies on cost-benefit analyses, often performed by government agencies to support and justify legislation on safety and environmental protection. The *eImpact* project (funded by the European Commission within the broad objective of halving automotive-related fatalities) [6] aimed at assessing the socioeconomic effects of intelligent vehicle safety systems (IVSSs) and their impact on traffic, safety, and efficiency, focusing on 12 different technologies:

1. Electronic Stability Control (ESC)
2. Full Speed Range ACC (FSR)
3. Emergency Braking (EBR)
4. Pre-Crash Protection of Vulnerable Road Users (PCV)
5. Lane Change Assistant (Warning) (LCA)
6. Lane Keeping Support (LKS)
7. Night Vision Warn (NIW)
8. Driver Drowsiness Monitoring and Warning (DDM)
9. eCall (one-way communication) (ECA)
10. Intersection Safety (INS)
11. Wireless Local Danger Warning (WLD)
12. Speed Alert (SPE)

The assessment procedure followed the scheme in Fig. 3.

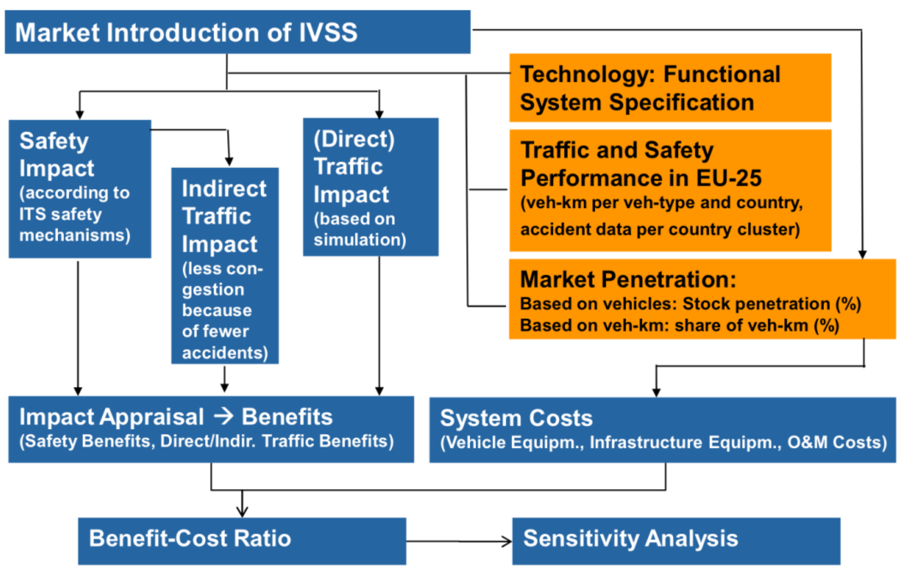


Figure 3. Cost-benefit assessment procedure [6].

For further details regarding the procedure, see [6]. Here we provide some of the conclusions from the report. First, Tables 5 and 6 show the estimated number of avoided fatalities, injuries, and accidents for each of the IVSS technologies for 2010 and 2020. Note that not all technologies were considered available in 2010.

Table 5. Number of Avoided Fatalities, Injuries, and Accidents for Each IVSS in Year 2010 [6]

	year 2010: number of avoided					
	Fatalities		Injuries		Accidents	
	low	high	low	high	low	high
ESC	1,914	2,240	32,792	38,265	24,594	28,698
FSR	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
EBR	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
PCV	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
LCA	2	11	264	1,189	198	892
LKS	56	149	1,420	3,784	1,065	2,838
NIW	2	10	87	367	66	275
DDM	4	13	153	367	114	275
ECA	*1,955*		*severe: 13,691* *slight: -15,647*		*0*	
INS	*n.a.*		*n.a.*		*n.a.*	
WLD	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
SPE	77	119	2,405	3,463	1,804	2,597
Base	33,895		1,409,415		1,081,627	

Note: "Low" and "high" refer to penetration extent, related to the absence or presence of incentives.

Table 6. Number of Avoided Fatalities, Injuries, and Accidents for Each IVSS in Year 2020 [6]

	year 2020: number of avoided					
	Fatalities		Injuries		Accidents	
	low	high	low	high	low	high
ESC	2,577	3,253	41,549	52,182	36,263	45,543
FSR	49	101	3,668	9,774	2,750	7,329
EBR	72	193	4,241	10,925	3,180	8,192
PCV	14	39	718	1,918	539	1,438
LCA	33	86	3,449	8,596	2,586	6,445
LKS	197	678	5,109	17,296	3,831	12,969
NIW	30	73	1,046	2,542	784	1,906
DDM	20	94	682	2,715	512	2,036
ECA	*1,199*		*severe: 8,398* *slight: -9,598*		*0*	
INS	*803*		*63,700*		*47,764*	
WLD	29	66	989	1,906	742	1,429
SPE	753	1,076	24,643	34,887	18,478	26,159
Base	20,791		873,695		798,808	

Note: "Low" and "high" refer to penetration extent, related to the absence or presence of incentives.

The above estimates are then used to compute benefit-cost ratios, with monetary values assigned to each type of event. Various other direct and indirect costs are also factored in, including indirect costs arising from the traffic congestion caused by an accident. The results of the cost-benefit analysis are reported in Table 7. Note that the cost-efficiency of a technology can increase, decrease, or remain unaffected by the penetration rate.

Table 7. Synopsis of Benefit-Cost Ratios [6]

	2010		***2020***	
	Low	**High**	**Low**	**High**
ESC	4.4	4.3	3.0	2.8
FSR	n.a.	n.a.	1.6	1.8
EBR	n.a.	n.a.	3.6	4.1
PCV	n.a.	n.a.	0.5	0.6
LCA	3.1	3.7	2.9	2.6
LKS	2.7	2.7	1.9	1.9
NIW	0.8	0.9	0.7	0.6
DDM	2.5	2.9	1.7	2.1
ECA	2.7		1.9	
INS	n.a.		0.2	
WLD	n.a.	n.a.	1.8	1.6
SPE	2.2	2.0	1.9	1.7

Note: "Low" and "high" refer to penetration extent, related to the absence or presence of incentives.

Challenges and Research Opportunities

Certainly, the automotive field remains interesting to control engineers due to an abundance of problems where the model-based approach can make a difference, provided we manage to find good models and suitable control design techniques.

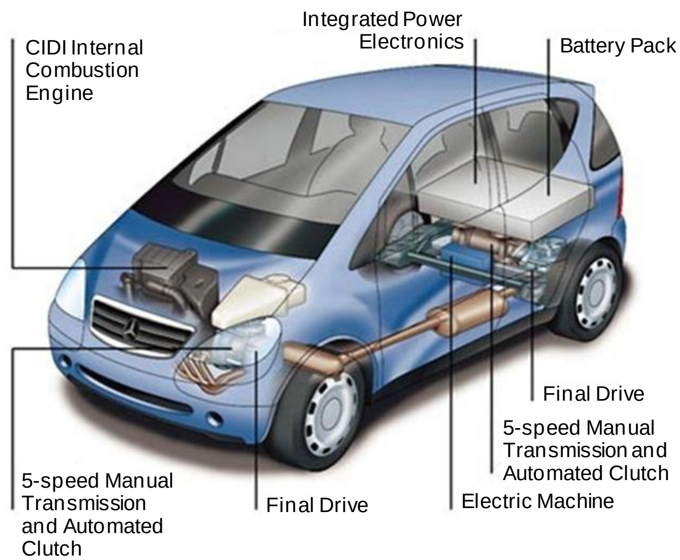
Electric vehicles (EVs) and hybrid electric vehicles (HEVs) are designed, prototyped, and produced in a variety of configurations. Almost every classical carmaker has an HEV project (even Ferrari has started its own "green" 599), and new companies like Tesla Motors aim their efforts directly at all-electric sport cars with incredible speed and mileage performances. The possible architectures of the vehicle (series or parallel configurations, four independently motored wheels, active differential, energy recovery with supercapacitors or flywheels, and so on) suggest a variety of problems, subproblems, and possibilities for innovation. Think of the general issue (sometimes called *energy source fragmentation*) of coordinating the different power sources (batteries, fuel, regenerative braking, solar panels) so as to trade off between autonomy and performance with the constraint of maintaining a reasonable state of charge (SOC) of the battery.

Source: Toyota

Figure 4. The Toyota Prius.

Performance goals can be cast in the form of power split control. Although engine development is pointing toward downsizing for minimizing fuel consumption, coupling it with an electrical motor greatly improves driveability and gives a very small engine the feel of a much larger one, appealing to the driver looking for low CO_2 emissions. Yet another variation on the theme is designing controllers that yield good fuel economy during unknown driving cycles without conflicting with driver aspirations in terms of driving fun; additional challenges come in the form of constraints on powertrain activity, such as Stop&Start functionality or dual-clutch gear shift.

The battery management system (BMS) is mandatory since damaged cells in a battery pack cannot simply be replaced with fresh ones, and this makes the battery one of the costliest and most delicate car components. The BMS objective is to maintain the health and safety of the battery pack through careful charge, discharge, measurement, and estimation to guarantee the affordability of the entire vehicle system. An interesting component of any BMS is a virtual sensor: the SOC estimator. Indeed, although batteries are ubiquitous as the core power source/storage system in small modern electronic devices ranging from cell phones to power tools, in large power grids, and in HEVs, one of the most difficult tasks in battery control applications is the correct estimation of the SOC. This is true for two reasons: first, battery charge and discharge states are definitions based more on manufacturer specifications than on an effective and universally accepted index; second, measuring such an index is difficult because of the highly nonlinear behavior of the battery during operation.



Source: Daimler

Figure 5. The Mercedes-Benz M-Class HyPer, a new hybrid concept vehicle.

Simple voltage-based charge gauges can be cost-effective for small toy rechargeable cells, but definitely not for the $30,000 battery pack of the Tesla Roadster, which requires a dedicated SOC estimator to accurately and reliably measure the health status of every single lithium-ion cell on board. The new-generation SOC virtual sensors are based on an electrochemical mathematical model of the single cell and extended Kalman filters for current/voltage feedback SOC estimation. In particular, for lithium-ion cells, the lithium concentration inside the two electrodes is of great interest not only because this value is closely related to the SOC, but also because the ability to estimate an excess or shortage of this concentration can avoid early aging and prevent battery malfunctions and safety hazards.

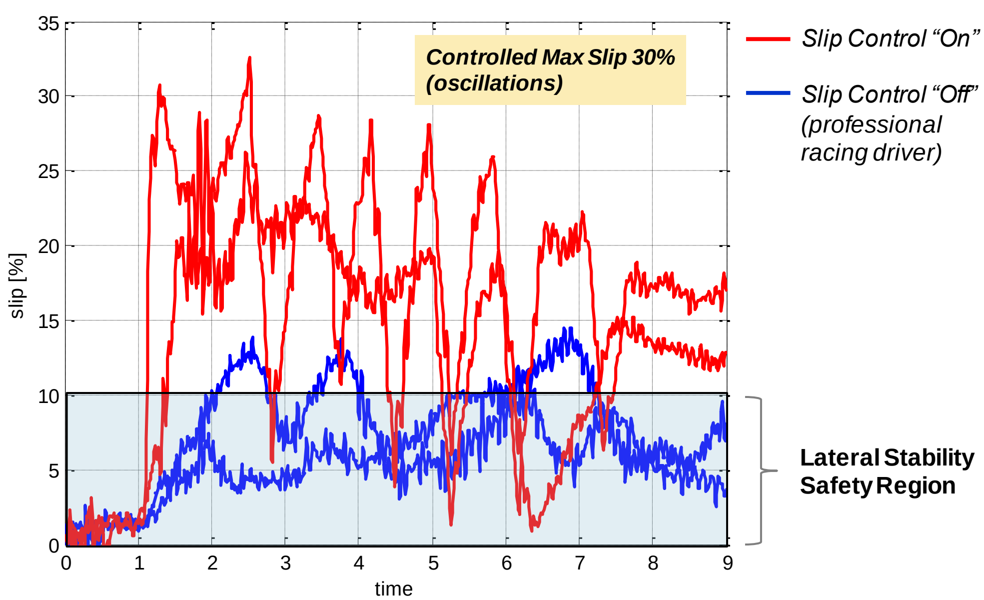
Battery packs composed of thousands of cells connected in series and in parallel have a global state of health equal to that of the weakest cell in the group because damaged cells cannot simply be replaced with new ones. Therefore, ensuring the equalization of the cell-to-cell SOC and maintaining the entire pack in good condition with respect to temperature, stress, and aging is one of the major challenges for hybrid vehicle control applications. Indeed, extended Kalman filtering of large-scale systems (derived from distributed parameter models) can be one of the most effective control tools for solving such a problem.

As mentioned earlier, the camshaft can be replaced with variable valve actuators, allowing for electronically controlled variable valve timing. This new generation of engines greatly improves on the conventional camshaft by better balancing the competing criteria of idle speed stability, fuel economy, and torque performance. VVAs are already used in production vehicles, but they need further improvement, for example, in the area of impact velocities between the valve, valve seat, and the actuator itself, which must be reduced to avoid excessive wear on the system and ensure acceptable noise levels. Another area of improvement is the opening and closing of the valves, which must be both fast and consistent with the strokes to avoid collision with the pistons and reduce variability in trapped mass.

Numerous control engineering challenges can be found working with the machinery on motorbikes or, more generally, on tilting vehicles. These vehicles (including electric versions), which are increasing in popularity because of their urban agility, are commonly found in two-wheeled versions but are now also commercially available in three-wheeled versions such as the Piaggio MP3 (see Fig. 6) and have been prototyped in four-wheeled versions. Estimation and control of the roll angle is a difficult problem, especially with low-cost inertial sensors; and with only two wheels (and sometimes only one in wheelie and stoppie maneuvers!), the problem of estimating the velocity of the bike is even more arduous so that traction control and ABS still provide challenging opportunities for improvement (see Fig. 7). According to specialists, though, the ultimate control problem is active yaw-roll control by coordination of brakes and traction control, which is extremely challenging since the system is not completely controllable.

Source: www.gizmag.com

Figure 6. A three-wheeled tilting vehicle.



Source: Politecnico di Milano

Figure 7. A motorbike TC in production features only raw limitations of slip peaks.

The development of new low-cost components (sensors, actuators, and microprocessors) will sustain control system market penetration and possibly the development of more sophisticated and effective control algorithms, typically characterized by significant computational loads. Nowadays the complexity of automotive control software is not related to the algorithm and its code but more often to its data, that is, the large and ever-growing number of (larger and larger) look-up tables used by gain-scheduled controllers. Further, these large tables have to be filled with numbers through lengthy calibration procedures. Thus, opportunities exist not only for reducing the use of look-up tables by means of different control algorithms, but also for devising better calibration/optimization algorithms and tools to fill the look-up tables, possibly online, during experiments on the test bench or the vehicle, rather than in the intervals between experiments, which is the current practice.

Another aspect to be considered is the need for validation and verification (V&V) procedures behind any control engineering achievement in the automotive field and the relative proportions of the various competencies required. According to one interview, "control engineering" accounts for only 25% of the production effort; software implementation and integration accounts for 30% and validation and testing for 45%—but the proportions of the latter two activities are expected to decrease whereas the proportion of effort devoted to control engineering is expected to increase. One specialist noted a preference for software being written by control engineers rather than by computer engineers, but this is seldom the case. The reason for this may be a (cultural) barrier on the part of control engineers, which can be partially overcome by a deeper awareness of the specific V&V tools now available, as well as by the availability of popular control design packages. On this same "software side" of control engineering, another challenge is presented by the rapidly growing complexity of the systems, which is exacerbated by the presence of legacy systems developed over the years. As a consequence, it is not unusual that new control algorithms, rather than being appropriately embedded into the existing code, are more or less added to it. To cope with this problem, efforts are under way in the automotive industry to establish an open and standardized automotive software architecture, notably AUTOSAR (AUTomotive Open System ARchitecture), that will "create a basis for industry collaboration on basic functions while providing a platform which continues to encourage competition on innovative functions", e.g. [1] and [7]. In Fig. 8, notice the boxes on sensors and actuators and think of the "application software component" as the piece of code containing the control algorithm.

Conclusions

The role of control technology in automotive vehicles and infrastructure will continue to widen as a necessary consequence of societal, economic, and environmental requirements. The application area will attract attention from control scientists and specialists not only for the difficult problems that need to be solved, but also because of the high volumes of production and the large number of players (from global automakers to local or specialized ones, suppliers, developers, and so on) in search of innovation and in competition for market share. However, again because of the large volumes, control experts will have to pay more attention to the entire software development cycle, since validation and verification, as well as calibration and maintenance, are crucial and sometimes very expensive items for this industry. An insufficient awareness of those aspects may slow penetration of our concepts and methods into this area, so integral to our current way of life.

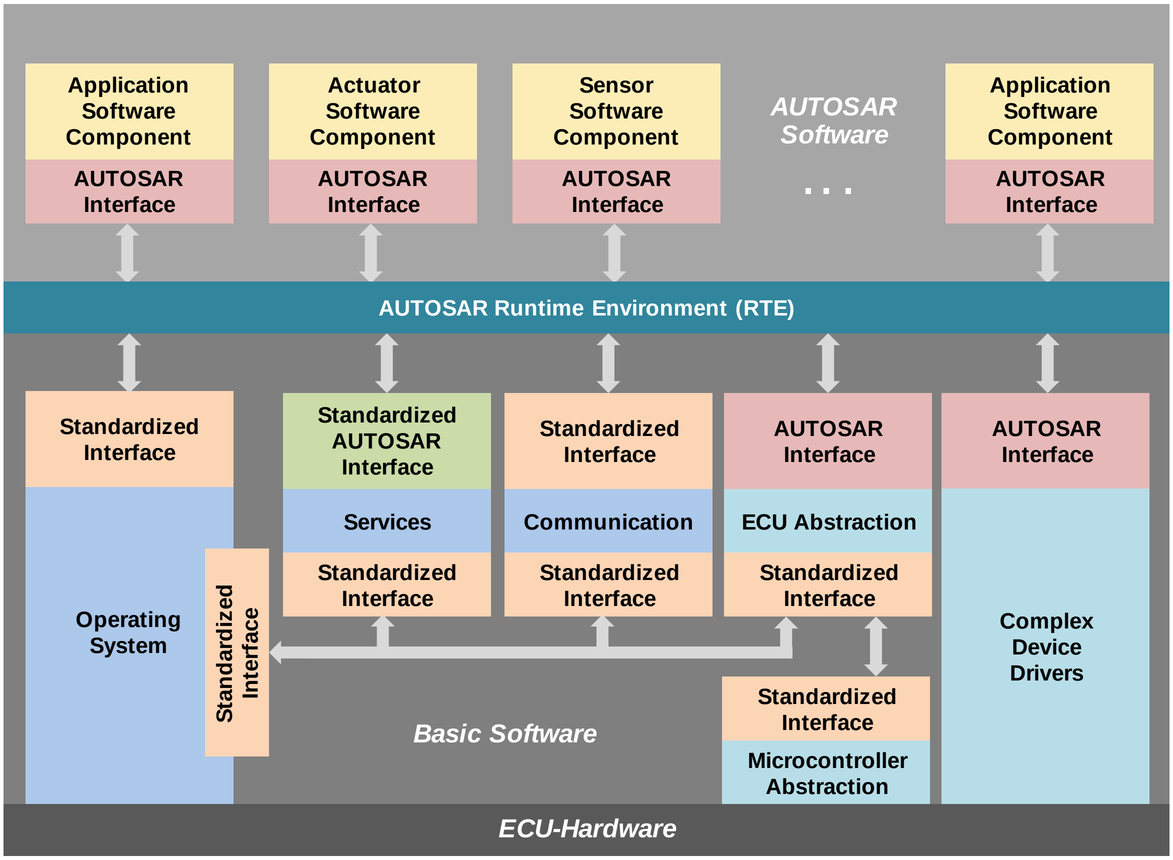


Figure 8. AUTOSAR software architecture.

Acknowledgments

Initial discussions that helped shape this report were held at the Berchtesgaden workshop with the following group members: Juergen Ackermann (DLR), Ken Butts (Toyota), Carlos Canudas-de-Wit (CNRS), Luigi Glielmo (Università del Sannio), Ilya Kolmanovsky (University of Michigan), Boris Lohmann (Technical University of Munich), and Greg Stewart (Honeywell). This work also includes direct and indirect contributions by many others besides those in the working group: Paolo Coeli (Centro Ricerche FIAT), Erik Coelingh (Volvo Car Corporation), Marika Di Benedetto (Università dell'Aquila), Paolo Falcone (Chalmers University), Raymond Freymann (BMW Group), Ruggero Frezza (M31, formerly at Università di Padova), Davor Hrovat (Ford Motor Co.), Mrdjan Jankovic (Ford Motor Co.), Jan-Ola Olsson (Volvo Car Corporation), A. Philips (Ford Motor Co.), Sergio Savaresi (Politecnico di Milano), Ralf Schwartz (AUDI), Malcolm Smith (University of Cambridge), C. Stephens (Ford Motor Co.), Tariq Samad (Honeywell), and E. Tseng (Ford Motor Co.). We wish to thank Luigi Iannelli (Università del Sannio at Benevento) and new Ph.D. graduates Angelo Palladino, Giovanni Palmieri, and Carmelo Speltino for useful discussions and reading.

Selected recommendations for research in automotive control:

- Powertrain architectures with multiple power sources are becoming increasingly popular; these will require sophisticated coordinated control approaches to manage the heterogeneous power sources.
- Correct estimation of the state of charge of a battery is one of the most difficult and important research needs in battery management systems for electric and hybrid-electric vehicles.
- Motorbikes and tilting vehicles represent an emerging and exciting opportunity for control technology, especially for active yaw-roll control.

References

[1] ETAS Group. The ECU Tool Company: Standardization and Partnerships in the Automotive Software Industry [Online], 2007. Available at http://www.imhtwk.de/resources/ETAS+Automotive+Standards+and+Partnerships.2.2-c.pdf.

[2] J. Pereira (ed.). "Monitoring and Control: Today's Market and Its Evolution till 2020, " Luxembourg: Office for Official Publications of the European Commission, 2009.

[3] J.B. White. "One billion cars," *The Wall Street Journal* [Online], April 17, 2006. Available at http://online.wsj.com/public/article/SB114487401909024337-ouwLdesvUMPaejrsk_WhxkaZzNU_20060516.html?mod=tff_main_tff_top.

[4] R.N. Charette. "This Car Runs on Code," *IEEE Spectrum,* February 2009.

[5] Taiwan External Trade Development Council. "Automobile Components & Auto Electronics Industry: Analysis & Investment Opportunities," January 2008.

[6] H. Baum, T. Geißler, U. Westerkamp, and C. Vitale. Cost-Benefit Analyses for Stand-Alone and Co-operative Intelligent Vehicle Safety Systems, version 2, report D6 within eImpact, Information Society and Media project funded by the European Commission, 2008.

[7] Autosar Technical Overview [Online]. Available at http://www.autosar.org/index.php?p=1&up=2&uup=0.

Related Content

The *Impact of Control Technology* report also includes more than 40 flyers describing specific "success stories" and "grand challenges" in control engineering and science, covering a variety of application domains. The ones below are closely related to the topic of this section.

Success Stories

- Active Safety Control for Automobiles – *L. Glielmo*
- Automated Manual Transmissions – *L. Iannelli*
- Control for Formula One! – *M. Smith*
- Coordinated Ramp Metering for Freeways – *M. Papageorgiou and I. Papamichail*

Grand Challenges

- Advanced Driver Assistance Systems Through Massive Sensor Fusion – *L. Glielmo*
- Vehicle-to-Vehicle/Vehicle-to-Infrastructure Control – *L. Glielmo*

These flyers—and other report content—are available at http://ieeecss.org/main/IoCT-report.

Control in Robotics

Mark W. Spong and Masayuki Fujita

Introduction

The interplay between robotics and control theory has a rich history extending back over half a century. We begin this section of the report by briefly reviewing the history of this interplay, focusing on fundamentals—how control theory has enabled solutions to fundamental problems in robotics and how problems in robotics have motivated the development of new control theory. We focus primarily on the early years, as the importance of new results often takes considerable time to be fully appreciated and to have an impact on practical applications. Progress in robotics has been especially rapid in the last decade or two, and the future continues to look bright.

Robotics was dominated early on by the machine tool industry. As such, the early philosophy in the design of robots was to design mechanisms to be as stiff as possible with each axis (joint) controlled independently as a single-input/single-output (SISO) linear system. Point-to-point control enabled simple tasks such as materials transfer and spot welding. Continuous-path tracking enabled more complex tasks such as arc welding and spray painting. Sensing of the external environment was limited or nonexistent.

Consideration of more advanced tasks such as assembly required regulation of contact forces and moments. Higher speed operation and higher payload-to-weight ratios required an increased understanding of the complex, interconnected nonlinear dynamics of robots. This requirement motivated the development of new theoretical results in nonlinear, robust, and adaptive control, which in turn enabled more sophisticated applications.

Today, robot control systems are highly advanced with integrated force and vision systems. Mobile robots, underwater and flying robots, robot networks, surgical robots, and others are playing increasing roles in society. Robots are also ubiquitous as educational tools in K-12 and college freshman experience courses.

The Early Years

The first industrial robot in the United States was the Unimate, which was installed in a General Motors plant in 1961 and used to move die castings from an assembly line and to weld these parts on auto bodies (Fig. 1). Full-scale production began in 1966. Another company with early robot products was Cincinnati Milacron, with companies in Japan and Europe also entering the market in the 1970s. Prior to the 1980s, robotics continued to be focused on manipulator arms and simple factory automation tasks: materials handling, welding, and painting.

From a control technology standpoint, the primary barriers to progress were the high cost of computation, a lack of good sensors, and a lack of fundamental understanding of robot dynamics. Given these barriers, it is not surprising that two factors were the primary drivers in the advancement of robot control in these early days. First, with the realization of the close connection between robot performance and automatic control, a community developed that focused on increasing fundamental understanding of dynamics, architecture, and system-level design. In retrospect, we can see that this

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

work had some significant limitations: control schemes were mostly based on approximate linear models and did not exploit knowledge of the natural dynamics of the robot, vision and force control were not well integrated into the overall motion control architecture, and mechanical design and control system design were separate.

The second factor was exogenous to both the controls and robotics communities, namely, Moore's Law. The increasing speed and decreasing cost of computation have been key enablers for the development and implementation of advanced, sensor-based control.

(Credit: George Devol)

Figure 1. Unimate, the first industrial robot.

At the forefront of research, both established control methods were explored in innovative applications for robots, and creative new ideas—some of which influenced control research more generally—were proposed. Especially worth noting is the early work on computed torque and inverse dynamics control [1]. As a sign of those times, it is interesting to note that until the mid-1980s, papers on robot control invariably included a calculation of the computational burden of the implementation.

Control of Manipulators

Robot manipulators have become a "standard" control application, and the synergies were widely recognized and exploited in research. The earlier research on computed torque and inverse dynamics control has been applied to numerous practical problems within and outside of robotics.

Beginning in the mid-1980s, robot manipulators became a "standard" control application, and the synergies were widely recognized and exploited in research. The earlier research on computed torque and inverse dynamics control [1], for example, helped motivate the differential geometric method of feedback linearization that has been applied to numerous practical problems within and outside of robotics [2]. For fully actuated rigid manipulators, the feedback linearization method was put on a firm theoretical foundation and shown to be equivalent to the inverse dynamics method [3]. The first nontrivial application of the feedback linearization method in robotics, in the sense that it requires a nonlinear coordinate transformation based on the solution of a set of PDEs, was to the problem of joint flexibility in robot manipulators [4]. Joint flexibility had previously been identified as the major limiting factor to manipulator performance, and it remains an important component of robot dynamics and control.

Another line of research pursued connections with robust control. Since feedback linearization relies on the exact cancellation of nonlinearities, the question of robustness to parameter uncertainty is immediately raised. Standard H_∞ control cannot adequately address this problem due to the persistent nature of the uncertainty. A solution for the special case of second-order systems, using the small-gain theorem, was worked out in [5], and the general case was presented in [6], which subsequently led to a new area of control now known as L_1-optimal control—a prime example of a robotics control contribution leading to new control theory. Several other methods of robust control, such as sliding modes and Lyapunov methods, have also been applied to the robust control problem for robot manipulators.

The mid-1980s were also a time of development in adaptive control, and again the connection with robotics was pursued. The fundamental breakthrough in the adaptive control of rigid manipulators was made by Slotine and Li [7]. The key to the solution of the adaptive control problem was the recognition of two important properties of Lagrangian dynamical systems: linearity in the inertia parameters and the skew-symmetry property of the robot inertia matrix [8].

Subsequently, the skew symmetry property was recognized as being related to the fundamental property of passivity. The term *passivity-based control* was introduced in the context of adaptive control of manipulators [9]. Passivity-based control has now become an important design method for a wide range of control engineering applications.

A final notable trend during this phase of the evolution of robot control was teleoperation—the control of robotic manipulators by possibly remotely located human operators. The obvious challenge that results is accommodating the delays involved, both for communication of sensory feedback and for transmission of the operator's command to the manipulator. That instability could be induced by time delays in so-called bilateral teleoperators, which involves feedback of sensed forces to the master, was recognized as a problem as early as the mid-1960s. Passivity-based control provided a breakthrough and enabled delay-independent stabilization of bilateral teleoperators [10], [11]. The key concept was to represent a master-slave teleoperator system as an interconnection of two-port networks and then encode the velocity and force signals as so-called scattering variables before transmitting them over the network. This approach renders the time-delay network element passive and the entire system stable independent of the time delay.

A state-of-the-art teleoperated robot is the Da Vinci surgical system from Intuitive Surgical, which integrates advances in micromanipulators, miniature cameras, and a master-slave control system to enable a surgeon to operate on a patient via a console with a 3-D video feed and foot and hand controls. However, neither force feedback nor remote operations are supported as yet; the surgeon's console is typically by the patient's side.

A state-of-the-art teleoperated robot is the Da Vinci surgical system from Intuitive Surgical, which integrates advances in micromanipulators, miniature cameras, and a master-slave control system to enable a surgeon to operate on a patient via a console with a 3-D video feed and foot and hand controls.

Mobile Robots

The problem of kinematic control of mobile robots received much attention starting in the 1980s as an application of differential geometric methods. The difficulty of the problem was dramatically revealed by Brockett's theorem, which showed that smooth time-invariant stabilizing control laws for such systems do not exist [12]. Brockett's theorem stimulated the development of alternative control methods , including hybrid switching control and time-varying approaches to stabilization of nonholonomic systems.

Mobile robots are now regularly used in many applications. One prominent application is aiding disaster recovery efforts in mines and after earthquakes. Military uses, such as for roadside bomb detection, form another broad category. Recently, products have been developed for consumer applications, such as the Roomba® and other robots from iRobot. Finally, wheeled mobile robots are exploring Mars and are poised to return to the moon.

Market Sizes and Investment

The robotics industry was slow getting started. Unimation did not show its first profit until 1975, almost a decade after it began full-scale production of its pioneering Unimate robot. Today, the Robotic Industries Association estimates that more than one million robots are in use worldwide; Japan has the largest deployment, with the United States having the second largest.

According to one recent market research report from Electronics.ca Publications, the global market for robotics was worth $17.3 billion in 2008 and is projected to increase to $21.4 billion in 2014, a compound annual growth rate (CAGR) of 4.0%. The largest segment of the market is industrial applications, worth $11.5 billion. Industrial robots, with their heavy reliance on the automotive industry, were especially hard hit with the recent global recession—2009 shipments were down 50% from year-ago levels, according to the Robotic Industry Association. Projected growth is lower for this segment than for professional service (market size of $3.3 billion in 2008) and military ($917 million) applications. Domestic services, security, and space applications constitute smaller segments, although the huge success of the Roomba floor-cleaning robot has demonstrated the enormous potential of consumer robotics.

Research Challenges

Underactuation

Underactuated robots have fewer control inputs than degrees of freedom and are a natural progression from flexible-joint and flexible-link robots. Underactuation leads naturally to a consideration of partial or output feedback linearization as opposed to full-state feedback linearization. Consideration of normal forms and zero dynamics is important in this context [13]. Energy/passivity methods are fundamental for the control of underactuated systems.

Visual Servo Control and Force Control

The idea of using imaging or video sensors for robot control is not new; it predates the availability of low-cost, high-quality digital cameras and advances in computational platforms enabling real-time processing of digital video signals. These latter developments have significantly increased interest in the topic.

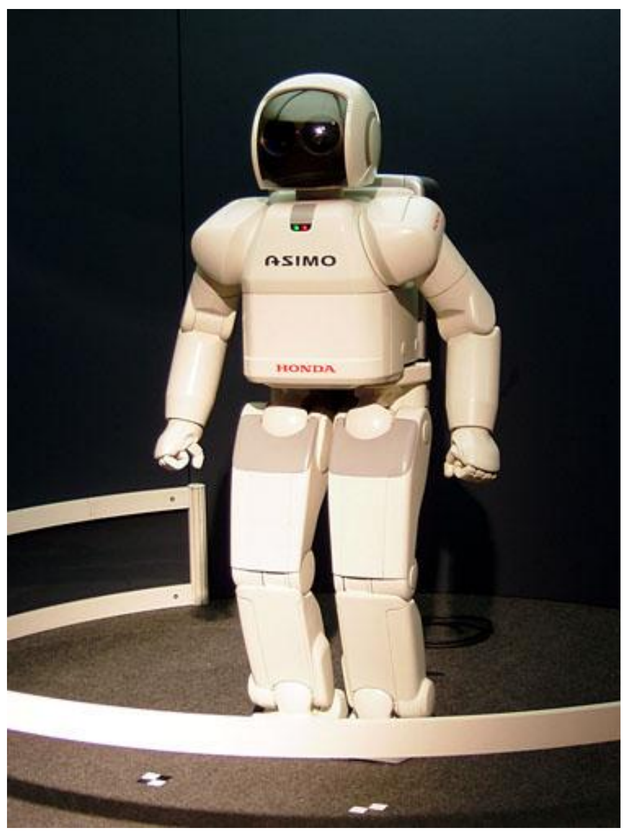
Visual servo control has traditionally used two methodologies, namely, position-based control and image-based control [14]. Position-based control uses vision to estimate the absolute position of the robot and uses the computed position error in the control algorithm. Image-based control, on the other hand, is based on computing the error directly in the image plane of the camera and avoids calculation of the robot position; thus, it is less sensitive to kinematic and calibration errors. Recently, both position-based and image-based methods have been incorporated into hybrid switching control strategies in order to take advantage of the strengths and avoid the weaknesses of both approaches.

Similar to vision-based control, force control in robotics has also traditionally been divided into two fundamental strategies, in this case, called hybrid position/force control and impedance control, respectively. Hybrid position/force control is based on the observation that one cannot simultaneously control both the position of a robot and the force it imparts to the environment. Thus, the task at hand can be decomposed into "directions" along which either position or force (but not both) is controlled. Conversely, impedance control does not attempt to control or track positions and forces. Rather the "mechanical impedance," which is the suitably defined Laplace transform of the velocity/force ratio, is the quantity to be controlled.

Locomotion

The development of legged robots is motivated by the fact that wheeled robots are not useful in rough terrain or in built structures. The number of legs involved is a free parameter in this research, with robots with as few as one (hopping robots) and as many as eight having been developed by multiple research groups. Bipedal robots are a particularly popular category, both for the anatomical similarity with their creators and because of the research challenges posed by their dynamic instability. An understanding of the dynamics and control of bipedal locomotion is also useful for the development of prosthetic and orthotic devices to aid humans with disabilities or missing limbs.

Readers who have seen videos of Honda's Asimov robots (Fig. 2) (readers who have not can check YouTube) or other humanoid robots may think that bipedal robots are "for real" now. The accomplishments of this research are indeed impressive. These robots can walk up and down ramps and stairs, counteract pushes and pulls, change gait, roll carts, play table tennis, and perform other functions. But the transition from research laboratory to commercial practice has not been made as yet. In particular, challenges remain for control engineers in the locomotion aspects specifically.



(Credit: Gnsin)

Figure 2. Honda's Asimov humanoid robot at Expo 2005 in Aichi, Japan.

Control of bipedal locomotion requires consideration of three difficult issues: hybrid nonlinear dynamics, unilateral constraints, and underactuation. The hybrid nature of the control problem results from impacts of the foot with the ground, which introduce discrete transitions between phases of continuous dynamic motion. Unilateral constraints arise from the fact that the foot can push but not pull on the ground and so the foot/ground reaction forces cannot change sign. Underactuation results again from the foot/ground interaction; there is no actuation torque between the foot and the ground. All these difficult issues require advanced methods of control to address them adequately. Energy/passivity methods, geometric nonlinear control, partial feedback linearization, zero dynamics, and hybrid control theory are all fundamental tools for designing rigorous control algorithms for walking [15], [16].

Multi-Agent Systems and Networked Control

Networked control systems and multi-agent systems are important recent application areas for robotics (Fig. 3). Synchronization, coordination, cooperative manipulation, flocking, and swarming combine graph theoretic methods with nonlinear control.

The emerging "hot topic" of cyber-physical systems is also closely related to networked control. Cyber-physical systems will get their functionality through massive networking. Sensors, actuators, processors, databases, and control software will work together without the need to be collocated.

Figure 3. Coordinated robots competing in the international RoboCup soccer competition in 2003. The Cornell team, led by controls researcher Raffaello D'Andrea, won the competition in 1999, 2000, 2002, and 2003.

Conclusions

Robotics today is a much richer field than even a decade or two ago, with far-ranging applications. Developments in miniaturization, in new sensors, and in increasing processing power have all opened new doors for robots.

As we reflect on the progress made in the field and the opportunities now lying ahead, it is clear that robotics is not a "closed" discipline. The definition of what constitutes a robot has broadened considerably, perhaps even leading to categorical confusion! A Roomba robot is a robot, but is a drone aircraft a robot or an airplane? And as increasingly many "robotic" features are added to automobiles—such as collision avoidance or steering feedback for lane departure warning—should we start thinking of our personal vehicles as robots too? Even in this report some of this redundancy or ambiguity exists. But the problems are similar in many respects, and these different communities have much to gain by building bridges, even nominal ones. Seeking out fundamental problems is the best way to make an impact.

Selected recommendations for research in robotics control:

- Approaches integrating position-based and image-based methods represent a promising research direction for solving the visual servo control problem.
- Control advances are needed for making legged robot locomotion practical; the problem is characterized by hybrid nonlinear dynamics, unilateral constraints, and underactuation.
- With the increasing interest in multivehicle robotics—under/in sea, on land, and in the air—multi-agent and networked control systems have become, and will continue to be, a key research area.

References

[1] B.R. Markiewicz. "Analysis of the Computed Torque Drive Method and Comparison with Conventional Position Servo for a Computer-Controlled Manipulator," NASA-JPL Technical Memo, 33-61, Mar. 1973.

[2] L.R. Hunt, R. Su, and G. Meyer. "Design for multi-input nonlinear systems," in *Differential Geometric Control Theory,* R.W. Brockett et al., eds. Boston-Basel-Stuttgart: Birkhauser, 1983, pp. 268-298.

[3] T. J. Tarn, A. K. Bejczy, A. Isidori, and Y. Chen. "Nonlinear feedback in robot arm control," in Proc. IEEE Conference on Decision and Control, Las Vegas, 1984, pp. 736-751.

[4] M.W. Spong. "Modeling and control of elastic joint robots," *Transactions of the ASME, J. Dynamic Systems, Measurement and Control,* vol. 109, pp. 310-319, Dec. 1987.

[5] M.W. Spong and M. Vidyasagar. "Robust linear compensator design for nonlinear robotic Control," *IEEE Journal of Robotics and Automation,* vol. RA 3, no. 4, pp. 345-350, Aug. 1987.

[6] M. Vidyasagar. "Optimal rejection of persistent bounded disturbances," *IEEE Trans. Auto. Control,* vol. AC-31, no. 6, pp. 527-534, June 1986.

[7] J.-J.E. Slotine and W. Li. "On the adaptive control of robot manipulators," *Int. J. Robotics Res.,* vol. 6, no. 3, pp. 147-157, Fall 1987.

[8] D. Koditschek. “Natural motion of robot arms,” in Proc. IEEE Conference on Decision and Control, Las Vegas, 1984.

[9] R. Ortega and M.W. Spong. “Adaptive control of robot manipulators: A tutorial,” *Automatica,* vol. 25, no. 6, pp. 877-888, 1989.

[10] R. Anderson and M. W. Spong. “Bilateral teleoperation with time delay,” *IEEE Trans. Aut. Cont.,* vol. 34, no. 5, pp. 494-501, May 1989.

[11] G. Niemeyer and J.-J.E. Slotine. “Stable adaptive teleoperation,” *IEEE Journal of Oceanographic Engineering,* vol. 16, no. 1, pp. 152–162, 1991.

[12] R. W. Brockett. “Asymptotic stability and feedback stabilization,” in *Differential Geometric Control Theory,* R.W. Brockett, R.S. Millman, and H. J. Sussmann, eds. Boston-Basel-Stuttgart: Birkhauser, 1983, pp. 181-191.

[13] A. Isidori and C.I. Byrnes. “Output regulation of nonlinear systems,” *IEEE Transactions on Automatic Control,* vol. 35, no. 2, pp. 131-140, 1990.

[14] S. A. Hutchinson, G.D. Hager, and P.I. Corke. “A tutorial on visual servo control,” *IEEE Transactions on Robotics and Automation,* vol. 12, no. 5, pp. 651-670, 1996.

[15] E.R. Westervelt, J.W. Grizzle, and D.E. Koditschek. “Hybrid zero dynamics of planar biped walkers,” *IEEE Transactions on Automatic Control,* vol. 48, no. 1, pp. 42-56, 2003.

[16] M.W. Spong and F. Bullo. “Controlled symmetries and passive walking,” *IEEE Transactions on Automatic Control,* vol. 50, no. 7, pp 1025-1031, 2005.

Related Content

The Impact of Control Technology report also includes more than 40 flyers describing specific “success stories” and “grand challenges” in control engineering and science, covering a variety of application domains. The ones below are closely related to the topic of this section.

Success Stories

- Dynamic Positioning System for Marine Vessels – *S.S. Ge, C.Y. Sang, and B.V.E. How*
- Mobile-Robot-Enabled Smart Warehouses – *R. D´Andrea*

Grand Challenges

- Control Challenges in High-Speed Atomic Force Microscopy – *S.O.R. Moheimani*
- Control for Offshore Oil and Gas Platforms – *S.S. Ge, C.Y. Sang, and B.V.E. How*

These flyers—and all other report content—are available at http://ieeecss.org/main/IoCT-report.

Control in Biological Systems

Francis J. Doyle III, B. Wayne Bequette, Rick Middleton, Babatunde Ogunnaike, Brad Paden, Robert S. Parker, and Mathukumalli Vidyasagar

Introduction

The field of control and systems has been connected to biological systems and biotechnology for many decades, going back to the work of Norbert Wiener on cybernetics in 1965, the work of Walter Cannon on homeostasis in 1929, and the early work of Claude Bernard on the *milieu interieur* in 1865. Nonetheless, the impact of control and systems on devices and applications in the field of biology has only emerged in recent years.

For this report, we will concentrate on the so-called *red* biotechnology, that is, the medical field of use, as opposed to *blue* biotechnology (aquatic use of biological technology), *green* biotechnology (agriculture and plant use), and *white* biotechnology (industrial applications).[1] For energy and process applications, the reader is referred to other sections of this report.

Hence, the emphasis in this section is on medical applications of control systems technology, which is very different from other areas in this study for multiple reasons:

- It is much less mature.
- It has far wider impact on human life.
- It is much less established.

This report is not meant to be a comprehensive review of all developments in biomedical control systems technology; instead the reader is referred to selected reviews, books, and tutorials on the topic [1]-[5].

Successful Applications of Control: Cardiovascular Systems and Endocrine Systems

As noted above, the field of biomedical control systems is relatively young compared to aerospace, automotive, and the chemical process fields. Nevertheless, some noteworthy recent developments have emerged in two key application areas: cardiovascular systems and endocrinology.

The field of biomedical control systems is relatively young compared to aerospace, automotive, and the chemical process fields. Nevertheless, some noteworthy recent developments have emerged in two key application areas: cardiovascular systems and endocrinology.

Cardiac Assist Devices

The area of cardiac assist devices has had a relatively long history of development, although advanced control theory and process modeling have only recently been applied to these devices [6]-[10]. In

[1] See, for example, the Wikipedia entry for *biotechnology.*

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

effect, cardiac assist devices are mechanical pumps that supplement endogenous cardiac output at an appropriate pressure to allow normal circulation through the patient's body. The control challenges include the changing demands for cardiac output as a function of the patient's "state" (for example, level of exercise, emotion, posture). The first such implantable device to receive approval by the FDA (1998) is the Baxter/Novacor left ventricular assist device (LVAD). Clearly, the ideal device would mimic the body's own mechanisms for maintaining cardiac output at target levels; however, the devices currently on the market are rather primitive in terms of automation, requiring the patient to adjust setpoints directly [6]. Recent developments for the pacemaker include real-time analysis and adaptive control [11]. Ventricular assist devices (VADs) are exploring feedback and model-based control to compensate for changes in patient needs (such as exercise) [12].

A more recent development is the use of magnetic levitation in the World Heart Inc. ventricular assist device called Levacor. World Heart recently received an FDA investigational device exemption (IDE) in preparation for clinical trials. The control system is a hybrid passive/active magnetic bearing where the active magnetic bearing employs a single active feedback loop designed by loop shaping. A key component of the technology is the high-reliability electronic design know-how transferred from aircraft control systems to this device.

In the cardiovascular area, another applied technology has been developed by Magnetecs: a magnetically guided catheter system for electrophysiology and other procedures. The control system is a combination of simple feedforward methods involving coordinate changes, feedback, and adaptive synthesis of visual models of the heart.

Blood Pressure Control

The IVAC Titrator was developed to regulate mean arterial pressure in hypertensive intensive care unit (ICU) patients by infusing sodium nitroprusside. The device received FDA approval in 1987 and was marketed for a short time, but was discontinued after a few years. The reasons for its failure in the marketplace include the following: (1) no consistent communication standards existed at the time, so the device had its own blood pressure sensor that was not particularly easy to set up; (2) the computer interface technology was not advanced; (3) the units were overpriced (IVAC chose to recoup R&D costs within a short time period); and (4) although studies showed less variability in blood pressure than with manual control, the effect of the reduced variability on patient outcomes was unclear [13]. Some studies suggested that patients were able to reduce hospital stays by a day. With new communication standards and advances in microprocessor-based pump technology, a closed-loop blood pressure system could probably succeed in the marketplace today.

Anesthesia Delivery

The effect of the intravenous anesthetic propofol is directly related to its concentration in the blood. Target-controlled infusion (TCI) is a model-based open-loop strategy designed to regulate the concentration of a drug in the blood by giving an initial intravenous bolus (shot), followed by time-dependent infusion. A commercial device, the Diprifusor (AstraZeneca Pharmaceuticals), has been available throughout much of the world since 1996 [14], [15], with millions of successful propofol infusions administered [16]. For a variety of reasons, no TCI device has received FDA approval in the United States [17]. Approval may be more likely if the infusion system incorporates a depth of anesthesia monitor, such as the bispectral index (BIS) manufactured by Aspect Medical Systems, to form a fully closed-loop system.

Other Applications

Beyond those highlighted here, a number of biomedical devices that have been successfully translated into commercial products using closed-loop technology include the implantable cardioverter defibrillator (ICD), the intracardiac electrogram (IEGM), and the oxygen saturation monitor. In other biomedical device areas, sensors are used to provide feedback to control and deliver electric signals that stimulate the brain to ease the tremors of Parkinson's disease and epilepsy by determining the extent and timing of stimulation. Additionally, closed-loop biomedical devices are used to treat peripheral vascular disease by using sensors to measure blood flow in a patient's limbs and determine the level of spinal cord or peripheral nerve stimulation required to improve blood flow, thereby reducing ischemic pain in the limbs. Closed-loop temperature control has been employed in ablation systems (such as the Atakr from Medtronic) with thermocouple feedback for safety.

Market Sizes and Investment

The potential market for the ventricular assist device is roughly 35,000 end-stage heart disease patients per year in the U.S. alone. The market capitalization of VAD companies exceeds $1B in the U.S. The pacemaker, with 250,000 implanted per year worldwide [11], is a ubiquitous biomedical device reliant on control algorithms to continue functioning. Catheter system companies have a collective market capitalization on the order of $0.5B.

The pacemaker, with 250,000 implanted per year worldwide, is a ubiquitous biomedical device reliant on control algorithms to continue functioning.

Approximately 17 million individuals in the U.S. are diagnosed diabetics, 5-10% of whom have type 1 and require insulin therapy. Similar incidence rates apply to other regions of the world. A 2005 estimate put the number of insulin pump users worldwide at 400,000 and growing by 10-12% per annum [18].

It is worth noting that regulatory factors and the cost of clinical trials often mean that market interest is less than patient demand. Regardless of the regulatory issues, however, the medical interest in developing tools that assist patients remains high because of the potential for impact at the patient level if a treatment intervention or device is successful.

Several government agencies are investing in research technology (including control systems) for the artificial pancreas (see below). The U.S. National Institutes of Health (NIH) recently announced a competition for the artificial pancreas ("Closed Loop Technologies: Clinical and Behavioral Approaches to Improve Type 1 Diabetes Outcomes," total of $5.5M funding). The EU sponsors multiple initiatives on the topic of the artificial pancreas, including "Development of a bio-artificial pancreas for type 1 diabetes therapy" and "AP@home." The NIH National Institute for Biomedical Imaging and Bioengineering (NIBIB) is a key player in research investment for biomedical devices. Several private foundations fund research in this area as well, including the Hillblom Foundation (endocrine and neurodegenerative disorders) and the Juvenile Diabetes Research Foundation (JDRF). The JDRF funds the Artificial Pancreas Consortium at a level of over $5.5M per year. A related topic is closed-loop control of blood glucose in the intensive care unit; several companies (such as Luminous Medical) are funding the development of sensors and closed-loop control algorithms for this application. Medical technology companies are hiring in this field, including Johnson & Johnson, Roche, Medtronic, and others. Small start-ups in this field have attracted venture capital (VC) funding at significant levels: World Heart received $30M in VC support in 2009, and Magnetecs has also attracted VC support.

Opportunities for New Applications and Research

"Red" biotechnology is an emerging and vibrant area for research in control systems. Below we discuss two topics of particular interest and then offer some general remarks on new research and development opportunities.

The Artificial Pancreas

In the area of endocrine systems, the most active area for control systems development has been the artificial pancreas for type 1 diabetes (Fig. 1). Such a device would be composed of a continuous glucose sensor, an insulin infusion pump, and an algorithm to regulate the insulin dosing in accordance with the measured glucose levels. Following is a brief summary of some of the key contributions, consisting primarily of the application of linear and nonlinear proportional-derivative (PD) algorithms to emulate the naturalistic biphasic insulin secretion profile. Some of the earliest work includes the glucose-controlled insulin infusion system (GCIIS) [19], which used some patient data (10-sec glucose sampling with a 4- to 5-min delay). The Biostator [20] also features a nonlinear PD algorithm, with the added nuance of a five-measurement window for filtering glucose measurements. It was implemented bedside and required specific patient customization. A nice review of the early algorithms is provided by Albisser [21], along with some patient data. Another detailed review is given by Broekhuyse et al. [22]. These reviews concluded that no controller was uniformly superior and that much more development was needed.

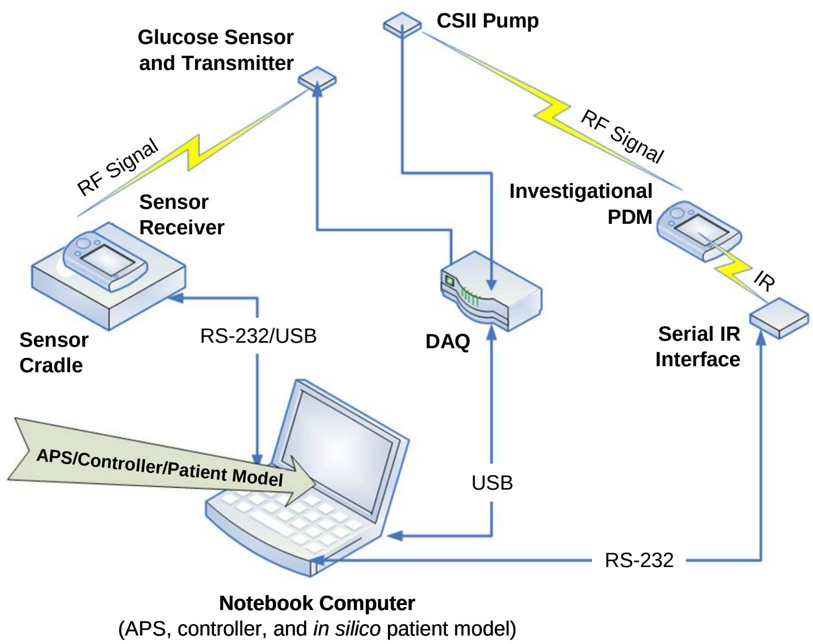


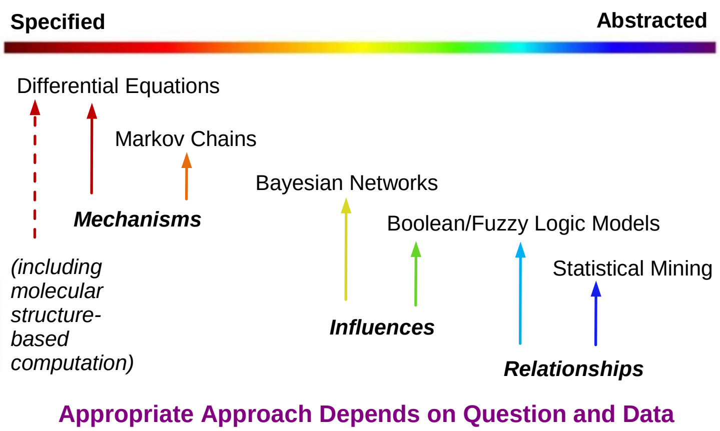
Figure 1. Components and communication protocols for the artificial pancreas [32].

More recently, advanced control technologies have been developed for the artificial pancreas, including variations on PID control [23], run-to-run control [24], and model predictive control [25]. In the last several years, clinical studies of advanced control methods have shown promise for future device developments [26]-[31]. Most of these trials use some degree of human intervention, for example, to input the size of a meal in advance of eating the meal.

To date, however, the state of the art in feedback control technology for insulin pumps and glucose sensors is limited mainly to bolus "wizards" and hypoglycemic alarming. The bolus wizards are effectively feedforward manual control algorithms that allow a patient to calculate an appropriate bolus of insulin to "cancel" the expected glucose rise from an anticipated meal or to recover from an elevated hyperglycemic state [33]. Hypoglycemic alarming refers to the prediction of low blood sugar in advance (say 30 minutes or more), generating an audible warning alarm so that the patient can take corrective action or, with newer products, shut off the insulin pump. The hypoglycemic alarming technology is appearing in European markets and is expected to appear in the U.S. soon.

Opportunities in the Field of Systems Biology

Here we briefly summarize some of the key technical issues in the area of systems biology. A broad spectrum of mathematical and analytical methods can be applied in the development of models for biomolecular regulatory networks and their subsequent analysis, with the twin goals of predicting their dynamics and generating conceptual insights about their operation. These methods range from highly abstracted (such as partial least-squares regression) to highly specified (for example, mass action kinetic differential equations, discrete stochastic and multiscale models). Fig. 2 illustrates only part of the range of computational mining and models.



Source: D. Lauffenburger, MIT

Figure 2. Spectrum of computational mining/modeling methods.

The highly abstracted methods are most powerful when little prior knowledge exists concerning the key network components, and the most highly specified methods are most powerful when a deep and comprehensive amount of knowledge is available concerning not only the components, but also their connectivity and the mechanisms of their associated interactions. Intermediate to these two extremes are methods that enable determination of logical influences characterizing network component interactions—going a vital step beyond mere connectivity but not mandating intense knowledge of kinetic mechanisms. These methods include Bayesian network, Markov chain, decision tree, Boolean logic, and fuzzy logic models. At the more detailed end of the modeling spectrum is the development of numerical methods and software for ordinary differential equation (ODE) solution, differential-algebraic equation (DAE) solution, and sensitivity analysis of these types of systems. Discrete stochastic simulation—and hence multiscale simulation, since discrete stochastic simulation by itself often involves so much computational complexity that both algorithmic (multiscale algorithms) and high-performance computing must be brought to bear to speed up the simulation—is a necessary part of the computational arsenal for biochemical simulation.

Early advances in the field of modeling and analysis for systems biology have guided therapeutic interventions. Specific drug/disease combinations include heparin/anticoagulation (optimal control) and HAART/HIV (plasma PK targets) [34]. Targeting measurable quantities in a patient-tailored way (patient-specific medicine) is becoming more common as models and measurements coincide in diseases such as HIV [34] and diabetes [35].

General Comments

- The most promising opportunities are those problems that formulate in a manner most closely associated with "traditional" systems engineering problems: medical problems subject to high economic burden and having a suitable number of "easily accessible" measurements that characterize effect or from which treatment effect can be estimated.

- The maximum potential for (economic) impact of control in medicine and medical devices is probably in the area of poorly understood diseases having complex dynamic responses and sparse (in time or state dimension) measurements. Examples of this class include inflammation and highly prevalent cancers. Low customer expectations further motivate these applications for control-theoretic approaches.

- Measuring impact on a social scale provides a different perspective. Here impact can be made in those disease populations that are too small (for example, those with low-prevalence cancers) to economically justify involvement by a major drug company. Another socially motivated potential impact is the development and deployment of biomedical devices and medical treatments to the geographically or economically disadvantaged (for example, those living far from a major medical center in developed countries or patients in Africa). Again, low customer expectations further motivate these applications for control-theoretic approaches.

- The VAD application could employ extremal seeking methods in a periodic system. Adaptive imaging based on catheter tip position data combined with imaging technology is an opportunity. Another imaging challenge, this time in cancer, is automated image identification for cancer volume assessment.

- Model structure analysis and structure selection tools, used to quickly evaluate when the available measurements are adequately captured by the model structure chosen, are important to medical decision making. Prediction quality may depend on model accuracy and the ability to quickly identify a model that is lacking—and to simultaneously highlight the portions of the model in need of refinement—could provide both better healthcare decisions and rapid model improvement. As alluded to above, the continued development of parameter identification tools for data-sparse systems, as well as nonlinear identifiability tools to establish which model structures can be supported from a given data set, would assist in diseases where insufficient state or measurement information (either spatially or temporally) is a concern.

- Another need is for improved (white box) tools for modeling data from populations of individuals and individuals of a given population, and for making sure the population and individual models are consistent. With the levels of uncertainty involved and the nonlinear dynamics of populations, multiple statistical and parameter-estimation tools will need to be used in combination.

Challenges and Barriers to Translation

As noted in the introduction, the field of biological systems is relatively young in terms of practical applications (market products), and several challenges must be overcome in translating closed-loop technologies to practice. The sheer complexity of (non-engineered) biological (networked) systems is the overarching daunting challenge. More specific obstacles include:

- The translation of relevant clinical outcomes for patient health into corresponding metrics on the measured variables in the body remains a challenge for sensors and control design.

- In the case of ventricular assist devices, high-level physiologic control is a promising technology. How does one control the speed of the pump and in response to what sensors?

- Notably, the objective in many medically oriented problems is patient quality of life, a “soft” objective. Changing to quality-adjusted life years (QALYs) can provide a numeric metric, but this is only in the aggregate; it is also controversial because it may lead to some patients not being treated due to the insensitivity of their QALY score to a particular intervention or treatment.

- A critical theoretical challenge for controlled drug delivery is the handling of both intrapatient and interpatient variability. This problem is quite different from engineering systems where uncertainty may be present, but it is typically of fixed (for example, stationary) structure. In biology, the variability is profound, and the same subject can differ significantly from one day to the next, depending on such factors as stress and environment. In some specific situations, such as diabetes, the intrasubject variation in critical subject parameters (such as insulin sensitivity) far exceeds the interpatient variability.

- The advances made in biomedical devices with closed-loop control capabilities have been enabled by developments in sensing and actuation. Conversely, the lack of appropriate and safe measurement and actuation devices precludes many applications.

Barriers that have delayed the marketing of some control-enabled devices include:

- Regulatory approval (by the FDA in the U.S.) for the artificial pancreas (see below). These agencies have not handled feedback algorithms in the past, so they are adapting to specify requirements for regulatory approval. The control community could play a role here in designing protocols for “stress testing,” in other words, suitable disturbance scenarios to challenge the closed-loop designs.

- The barriers for cardiovascular devices are comparable to other aspects of FDA approval.

- The need for appropriate models and especially modeling paradigms for model-based control systems raises questions that do not have easy answers. How does one develop a reliable model for patients with widely varying physiological characteristics; how does one maintain such models; what model paradigm will facilitate model development for biomedical applications?

- Communication between systems engineers and clinicians is also a barrier. Each group speaks its own language, with associated jargon. Until representatives of the two groups develop a common language, often as a natural outcome of a close collaboration, engineering solutions may not be solving clinical problems in an optimal way (if at all).

From a regulatory standpoint, the focus should be on device (rather than drug) development, as the pathway to acceptance is generally faster. A further concern in control algorithm development is the burden of proof required for algorithm-based device approval (superiority vs. non-inferiority trials); the technical complexity in the algorithm (for example, closed loop, model-based, predictive, adaptive?), and the potential inability to a priori provide bounds on device performance for all individuals, may cause device rejection unless all possible failure modes are characterized and evaluated in significant detail. A secondary technical concern is the inherent variability or uncertainty encountered in a clinical patient population, which is typically greater than 100% (parametrically). A final complicating factor is economics. The price of a device is both market and development cost driven. Devices are often too expensive for the vast majority of patients; lack of insurance coverage may make it impossible to realize profitable sales volumes.

The regulatory approval process and the economics of the healthcare system are probably the greatest barriers and are also the least technical in nature.

Conclusions

All of the opportunities discussed in this section are effectively worthless in the medical arena if they cannot be translated to clinical practice. This fact simply highlights (1) the need to communicate more effectively the strengths and weaknesses of control tools and calculations with noncontrol experts, and (2) the requirement that interfaces for any or all of the aforementioned tools be constructed such that the tools can be deployed in a clinical environment by conventional healthcare providers such as nurses.

Finally, to underscore the importance of the promise of biological systems as a target domain for the controls community, we note that three of the National Academy of Engineering Grand Challenges [36] have direct relevance for control systems technology in medicine:

- Engineer better medicines. Engineers are developing new systems to use genetic information, sense small changes in the body, assess new drugs, and deliver vaccines.
- Advance health informatics. Stronger health information systems not only improve everyday medical visits, but they are essential to countering pandemics and biological or chemical attacks.
- Reverse-engineer the brain. For decades, some of engineering's best minds have focused their thinking skills on how to create thinking machines—computers capable of emulating human intelligence.

Selected recommendations for research in the control of biological systems:

- Success in the development of the artificial pancreas, and of other closed-loop biomedical devices, will be contingent on the development of robust, verifiable advanced control algorithms.
- Algorithms for controlled drug delivery are an exciting research opportunity; advances are needed to characterize and to accommodate the considerable intrapatient as well as interpatient variability that exists in disease (and healthy) populations.
- Biological control and diagnostic applications require modeling and system identification approaches that integrate structure determination, parameter estimation, and model verification—and human understandability of generated models is an important criterion.

References

[1] W.B. Cannon. *The Wisdom of the Body.* New York: W.W. Norton & Company, 1932.

[2] F.S. Grodins. *Control Theory and Biological Systems.* New York: Columbia University Press, 1963.

[3] G.W. Swan. "Optimal control applications in biomedical engineering – a survey," *Opt. Control Appls. Methods*, vol. 2, pp. 311-334, 1980.

[4] M.C.K. Khoo. *Physiological Control Systems: Analysis, Simulation, and Estimation*. New York: Wiley, 1999.

[5] F.J. Doyle III, L. Jovanovic, D.E. Seborg, R.S. Parker, B.W. Bequette, A.M. Jeffrey, X. Xia, I.K. Craig, and T. McAvoy, "A tutorial on biomedical process control," *J. Process Control*, vol. 17, pp. 571-630, 2007.

[6] J. Boston, M. Simaan, J. Antaki, and Y. Yu. "Control issues in rotary heart assist devices," in Proc. American Control Conference, 2000, pp. 3473-3477.

[7] L. A. Geddes. "Historical highlights in cardiac pacing," *Eng. Med. Biol. Mag.,* vol.2, pp. 12–18, 1990.

[8] B. Paden, J. Ghosh, and J. Antaki. "Control system architectures for mechanical cardiac assist devices," in Proc. American Control Conference, 2000, pp. 3478-3482.

[9] L. Baloa, D. Liu, and J. Boston. "Control of rotary assist devices," in Proc. American Control Conference, 2000, pp. 2982-2986.

[10] G. Giridharan and M. Skliar. "Nonlinear controller for ventricular assist devices," *Artificial Organs,* vol. 26, pp. 980-984, 2002.

[11] S.A.P. Haddad, R.P.M. Houben, and W.A. Serdijn. "The evolution of pacemakers." *Eng. Med. Biol. Mag.,* vol. 25, pp. 38–48, 2006.

[12] H. Schima, M. Vollkron, U. Jantsch, R. Crevenna, W. Roethy, R. Benkowski, G. Morello, M. Quittan, M. Hiesmayr, and G. Wieselthaler. "First clinical experience with an automatic control system for rotary blood pumps during ergometry and right-heart catheterization," *J. Heart Lung Transplant,* vol. 25 no. 2, pp. 167-73, 2006.

[13] B.W. Bequette. "A tutorial on biomedical process control. III. Modeling and control of drug infusion in critical care," *J. Process Control,* vol. 17, no. 7, pp. 582-586, 2007.

[14] S. Schraag. "Theoretical basis of target control anaesthesia: History, concept and clinical perspectives," *Best Practices & Research Clinical Anesthesiology,* vol. 15, no. 1, pp. 1-17, 2001.

[15] B.G. Goudra and J.E. Mandel. "Target-controlled infusions/patient-controlled sedation," *Techniques in Gastrointestinal Endoscopy,* vol. 11, pp. 181-187, 2009.

[16] T.D. Egan and S.L. Shafer. "Target-controlled infusions for intravenous anesthetics. Surfing USA not!" *Anesthesiology,* vol. 99, no. 5, pp. 1039-1041, 2003.

[17] P.J. Manberg, C.M. Vozella, and S.D. Kelley. "Regulatory challenges facing closed-loop anesthetic drug infusion devices," *Clinical Pharmacology & Therapeutics,* vol. 84, no. 1, pp. 166-169, 2008.

[18] Medical News Today. "New generation of insulin pump – Starbridge secures global diabetes contract [Online], 19 November 2005. Available at http://www.medicalnewstoday.com/articles/ 33847.php.

[19] E. Pfeiffer, C. Thum, and A.H. Clemens. “The artificial beta cell – a continuous control of blood sugar by external regulation of insulin infusion (glucose controlled insulin infusion system),” *Horm. Metab. Res.*, vol. 487, pp. 339-342, 1974.

[20] A. Clemens. “Feedback control dynamics for glucose controlled insulin infusion system.” *Med. Progr. Technol.*, vol. 6, pp. 91-98, 1979.

[21] A. Albisser. Devices for the control of diabetes mellitus, in *Proc. IEEE*, vol. 67, pp. 1308-1322, 1979.

[22] H. Broekhuyse, J. Nelson, B. Zinman, and A. Albisser. “Comparison of algorithms for the closed-loop control of blood glucose using the artificial beta cell,” *IEEE Trans. Biomed. Eng.*, vol. 28, pp. 678-687, 1981.

[23] G.M. Steil and M.F. Saad. “Automated insulin delivery for type 1 diabetes,” *Curr. Opin. Endocrinol. Diabetes*, vol. 13, pp. 205-211, 2006.

[24] C. Owens, H. Zisser, L. Jovanovic, B. Srinivasan, D. Bonvin, and F.J. Doyle III, “Run-to-run control of blood glucose concentrations for people with type 1 diabetes mellitus,” *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 996-1005, 2006.

[25] R. Parker, F. Doyle, and N. Peppas. “A model-based algorithm for blood glucose control in type I diabetic patients,” *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 148-157, 1999.

[26] R. Hovorka, L.J. Chassin, M.E. Wilinska, V. Canonico, J. Akwe, O.M. Federici, M. Massi-Benedetti, I. Hutzli, C. Zaugg, H. Kaufmann, M. Both, T. Vering, H. Schaller, L. Schaupp, M. Bodenlenz, and T. Pieber, “Closing the loop: The adicol experience,” *Diabetes Technology & Therapeutics,* vol. 6, pp. 307-318, 2004.

[27] G.M. Steil, K. Rebrin, C. Darwin, F. Hariri, and M.F. Saad. “Feasibility of automating insulin delivery for the treatment of type 1 diabetes,” *Diabetes,* vol. 55, pp. 3344-3350, 2006.

[28] C. Palerm, H. Zisser, W. Bevier, L. Jovanovic, and F.J. Doyle III. “Prandial insulin dosing using run-to-run control: Application of clinical data and medical expertise to define a suitable performance metric,” *Diabetes Care*, vol. 30, pp. 1131-1136, 2007.

[29] C. Ellingsen, E. Dassau, H. Zisser, B. Grosman, M.W. Percival, L. Jovanovic, and F.J. Doyle III. “Safety Constraints in an artificial β-cell: An implementation of model predictive control with insulin on board,” *J. Diab. Sci. Tech.,* vol. 3, pp. 536-544, 2009.

[30] F.H. El-Khatib, S.J. Russell, D.M. Nathan, R.G. Sutherlin, and E.R. Damiano. “A bihormonal closed-loop artificial pancreas for type 1 diabetes,” *Sci. Transl. Med.,* vol. 2, no. 27, pp. 1-12, 2010.

[31] R. Hovorka, J.M. Alle, D. Elleri, L.J. Chassin, J. Harris, D. Xing, C. Kollman, T. Hovorka, A.M. Larsen, M. Nodale, A. De Palma, M.E. Wilinska, C.L. Acerini, and D.B. Dunger. “Manual closed-loop insulin delivery in children and adolescents with type 1 diabetes: A phase 2 randomised crossover trial,” *Lancet,* Feb. 4, 2010.

[32] E. Dassau, C. Palerm, H. Zisser, B.A. Buckingham, L. Jovanovic, and F.J. Doyle III. “In silico evaluation platform for artificial pancreatic β-cell development – a dynamic simulator for closed-loop control with hardware in the loop,” *Diabetes Tech. Ther.*, vol. 11, pp. 187-194, 2009.

[33] H. Zisser, L. Robinson, W. Bevier, E. Dassau, C. Ellingsen, F.J. Doyle III, and L. Jovanovic. “Bolus calculator: A review of four smart insulin pumps.” *Diabetes Tech. Ther.,* vol. 10, pp. 441-444, 2008.

[34] C.V. Fletcher, R.C. Brundage, T. Fenton, C.G. Alvero, C. Powell, L.M. Mofenson, and S.A. Spector. “Pharmacokinetics and pharmacodynamics of efavirenz and nelfinavir in HIV-infected children participating in an area-under-the-curve controlled trial.,” *Clin. Pharmacol. Ther.,* vol. 83, no. 2, pp. 300-6, 2008.

[35] Plank et al. 2006 J. Plank, J. Blaha, J. Cordingley, M.E. Wilinska, L.J. Chassin, C. Morgan, S. Squire, M. Haluzik, J. Kremen, S. Svacina, W. Toller, A. Plasnik, M. Ellmerer, R. Hovorka, and T.R. Pieber. "Multicentric, randomized, controlled trial to evaluate blood glucose control by the model predictive control algorithm versus routine glucose management protocols in intensive care unit patients," *Diabetes Care,* vol. 29, pp. 271-6, 2006.

[36] National Academy of Engineering. Grand Challenges for Engineering [Online], 2008. Available at http://www.engineeringchallenges.org/.

Related Content

The Impact of Control Technology report also includes more than 40 flyers describing specific "success stories" and "grand challenges" in control engineering and science, covering a variety of application domains. The ones below are closely related to the topic of this section.

Grand Challenges

- Biophysical Networks – *F.J. Doyle III*
- Dynamics and Control for the Artificial Pancreas – *F.J. Doyle III*
- High-Performance Control with Slow Computing! – *R. Murray*
- Redesigning a Bacterium Control System – *R. Murray*

These flyers—and all other report content—are available at http://ieeecss.org/main/IoCT-report.

Control for Renewable Energy and Smart Grids

Eduardo F. Camacho, Tariq Samad, Mario Garcia-Sanz, and Ian Hiskens

Introduction

The use of renewable energy increased greatly just after the first big oil crisis in the late seventies. At that time, economic issues were the most important factors, hence interest in such processes decreased when oil prices fell. The current resurgence of interest in the use of renewable energy is driven by the need to reduce the high environmental impact of fossil-based energy systems. Harvesting energy on a large scale is undoubtedly one of the main challenges of our time. Future energy sustainability depends heavily on how the renewable energy problem is addressed in the next few decades.

Although in most power-generating systems, the main source of energy (the fuel) can be manipulated, this is not true for solar and wind energies. The main problems with these energy sources are cost and availability: wind and solar power are not always available where and when needed. Unlike conventional sources of electric power, these renewable sources are not "dispatchable"—the power output cannot be controlled. Daily and seasonal effects and limited predictability result in intermittent generation. Smart grids promise to facilitate the integration of renewable energy and will provide other benefits as well.

Industry must overcome a number of technical issues to deliver renewable energy in significant quantities. Control is one of the key enabling technologies for the deployment of renewable energy systems. Solar and wind power require effective use of advanced control techniques. In addition, smart grids cannot be achieved without extensive use of control technologies at all levels.

> Control is a key enabling technology for the deployment of renewable energy systems. Solar and wind power require advanced control techniques for high-performance and reliable operation.

This section of the report will concentrate on two forms of renewable energy—wind and solar—and on the role of smart grids in addressing the problems associated with the efficient and reliable delivery and use of electricity and with the integration of renewable sources. Solar and wind power plants exhibit changing dynamics, nonlinearities, and uncertainties—challenges that require advanced control strategies to solve effectively. The use of more efficient control strategies would not only increase the performance of these systems, but would increase the number of operational hours of solar and wind plants and thus reduce the cost per kilowatt-hour (KWh) produced.

Both wind and solar have tremendous potential for fulfilling the world's energy needs. In the case of wind, if conventional onshore wind turbines with 80-m towers were installed on 13% of the earth's surface, the estimated wind power that could be commercially viable is 72 terawatt (TW). That amounts to almost five times the global power consumption in all forms, which currently averages about 15 TW. With capacity that has tripled in the last five years, wind energy is the fastest growing energy source in the world. Using larger wind turbines to convert kinetic energy into electricity has significantly increased the average power output of a wind turbine unit; most major manufacturers have developed large turbines that produce 1.5 to 3.5 megawatts (MW) of electric power, even reaching 5 to 6 MW per

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

turbine in some cases. At the end of 2009, with 159.2 gigawatt (GW) of wind-powered generators worldwide, primarily grouped together to create small wind farms, the global collective capacity was 340 terawatt-hour (TWh) of energy annually, or 2% of global electric energy consumption. Several countries have achieved relatively high levels of wind power penetration: about 19% in Denmark, 14% in Spain and Portugal, and 7% in Germany and Ireland. Government subsidies have been a key factor in increasing wind power generation. These subsidies, in turn, have often been justified by the renewable portfolio standards (RPSs) that several countries have adopted and that require increasing the production of energy from renewable sources. In particular, RPSs generally obligate utilities to produce a specified fraction of their electricity from renewable energy. The European Union has a regionwide RPS of 20% by 2020; the United States of 20% by 2030, with different targets and years depending on the state (for example, 15% by 2025 in Arizona and 20% by 2020 in Colorado).

Although wind energy is a clean and renewable source of electric power, many challenges must be addressed. Wind turbines are complex machines, with large flexible structures working under turbulent and unpredictable environmental conditions, and are connected to a constantly varying electrical grid with changing voltages, frequency, power flow, and the like. Wind turbines have to adapt to those variations, so their efficiency and reliability depend heavily on the control strategy applied. As wind energy penetration in the grid increases, additional challenges are being revealed: response to grid disturbances, active power control and frequency regulation, reactive power control and voltage regulation, restoration of grid services after power outages, and wind prediction, for example.

Another abundant, sustainable source of energy is the sun. One of the greatest scientific and technological opportunities we face is developing efficient ways to collect, convert, store, and utilize solar energy at an affordable cost. The solar power reaching the earth's surface is about 86,000 TW. Covering 0.22% of our planet with solar collectors with an efficiency of 8% would be enough to satisfy the current global power consumption. Estimates are that an energy project utilizing concentrating solar power (CSP) technology deployed over an area of approximately 160 x 160 km in the Southwest U.S. could produce enough power for the entire U.S. consumption.

Solar-sourced electricity can be generated either directly using photovoltaic (PV) cells or indirectly by collecting and concentrating the solar power to produce steam, which is then used to drive a turbine to provide the electric power (CSP). We focus on CSP in this section, as control has greater relevance to it.

Concentrating solar thermal systems use optical devices (usually mirrors) and sun-tracking systems to concentrate a large area of sunlight onto a smaller receiving area. The concentrated solar energy is then used as a heat source for a conventional power plant. A wide range of concentrating technologies exists, the main ones being parabolic troughs, solar dishes, linear Fresnel reflectors, and solar power towers. The primary purpose of concentrating solar energy is to produce high temperatures and therefore high thermodynamic efficiencies.

Parabolic trough systems are the most commonly used CSP technology. A parabolic trough consists of a linear parabolic mirror that reflects and concentrates the received solar energy onto a tube (receiver) positioned along the focal line. The heat transfer fluid is pumped through the receiver tube and picks up the heat transferred through the receiver tube walls. The parabolic mirror follows the sun by tracking along a single axis. Linear Fresnel reflectors use various thin mirror strips to concentrate sunlight onto tubes containing heat transfer fluid. Higher concentration can be obtained, and the mirrors are cheaper than parabolic mirrors, but a more complex tracking mechanism is needed.

The main control problems with solar plants are related to sun tracking and control of the thermal variables. Although control of the sun-tracking mechanisms is typically done in an open-loop mode, control of the thermal variables is mainly done in closed loop. Solar plants exhibit changing dynamics, nonlinearities, and uncertainties, characteristics that result in detuned performance with classical PID control. Advanced control strategies that can cope with these issues are needed for better performance and for decreasing the cost per kilowatt-hour generated.

The uncertainty and intermittency of wind and solar generation are major complications that must be addressed before the full potential of these renewables can be reached. The *smart grid*—an evolution of electricity networks toward greater reliance on communications, computation, and control—promises a solution. The term gained prominence through the U.S. Energy Independence and Security Act (EISA) of 2007, the European Technology Platform for the Electricity Networks of the Future, and similar initiatives across numerous other countries. The U.S. Department of Energy has provided a concise description of the smart grid [1]:

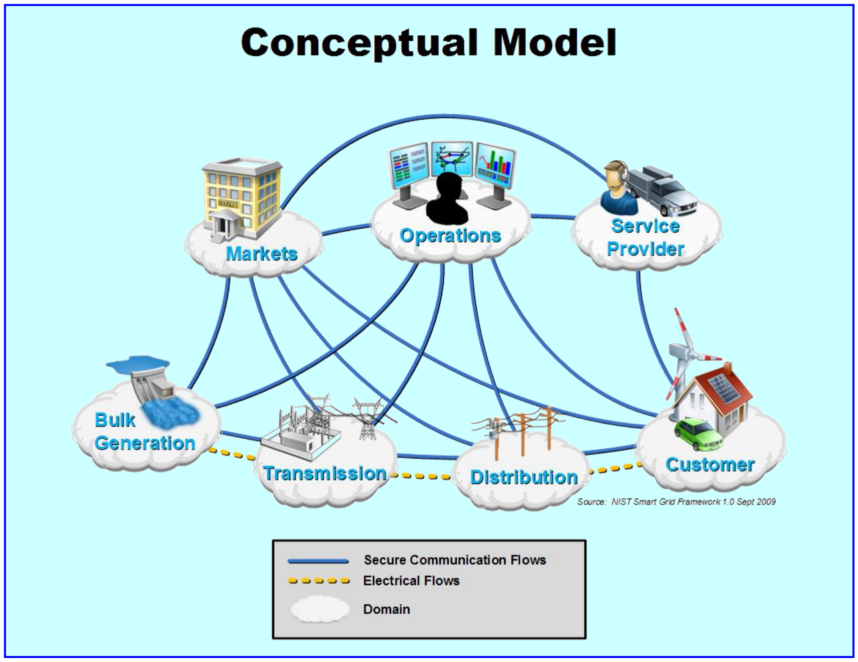
> The application of advanced digital technologies (i.e., microprocessor-based measurement and control, communications, computing, and information systems) are expected to greatly improve the reliability, security, interoperability, and efficiency of the electrical grid, while reducing environmental impacts and promoting economic growth. Achieving enhanced connectivity and interoperability will require innovation, ingenuity, and different applications, systems, and devices to operate seamlessly with one another, involving the combined use of open system architecture, as an integration platform, and commonly shared technical standards and protocols for communications and information systems. To realize Smart Grid capabilities, deployments must integrate a vast number of smart devices and systems.

The EU's SmartGrids technology platform summarizes the benefits of smart grids as follows. They:

- Better facilitate the connection and operation of generators of all sizes and technologies;
- Allow consumers to play a part in optimizing the operation of the system;
- Provide consumers with greater information and options for choice of supply;
- Significantly reduce the environmental impact of the whole electricity supply system;
- Maintain or even improve the existing high levels of system reliability, quality and security of supply;
- Maintain and improve the existing services efficiently;
- Foster market integration.

The broad spectrum of entities and stakeholders covered by the smart grid is evident from the conceptual model of Fig. 1. The smart grid further broadens the already highly distributed nature of power systems by extending control to the consumer level. The smart grid can be conceptualized as an extensive cyber-physical system that supports and significantly enhances controllability and responsiveness of highly distributed resources and assets within electric power systems.

> The smart grid can be conceptualized as an extensive cyber-physical system that supports and facilitates significantly enhanced controllability and responsiveness of highly distributed resources within electric power systems.



Source: NIST Smart Grid Framework 1.0, Sept. 2009

Figure 1. Depiction of the NIST smart grid conceptual model [2].

The term *smart grid* implies that the existing grid is dumb, which is far from true. The current grid structure reflects carefully considered trade-offs between cost and reliability. The responsiveness achievable through smart grid concepts will, however, play a vital role in achieving large-scale integration of new forms of generation and demand. Renewable generation will make an increasingly important contribution to electric energy production into the future. Integration of these highly variable, widely distributed resources will call for new approaches to power system operation and control. Likewise, new types of loads, such as plug-in electric vehicles and their associated vehicle-to-grid potential, will offer challenges and opportunities. Establishing a cyberinfrastructure that provides ubiquitous sensing and actuation capabilities will be vital to achieving the responsiveness needed for future grid operations. Sensing and actuation will be pointless, though, without appropriate controls.

Successful Applications of Control

Wind Energy

Charles F. Brush is widely credited with designing and erecting the world's first automatically operating wind turbine for electricity generation. The turbine, which was installed in Cleveland, Ohio, in 1887, operated for 20 years with a peak power production of 12 kW (Fig. 2). An automatic control system ensured that the turbine achieved effective action at 6.6 rpm (330 rpm at the dynamo) and that the dc voltage was kept between 70 and 90 volts. Another remarkable project in early wind energy research was the 1.25-MW wind turbine developed by Palmer Putnam [3] in the U.S. The giant wind turbine, which was 53 m (175 feet) in diameter, was installed in Vermont, Pennsylvania, around 1940 and featured two blades with a hydraulic pitch control system.

Modern wind-driven electricity generators began appearing during the late 1970s. At that time, the average power output of a wind turbine unit was about 50 kW with a blade length of 8 m. Since then,

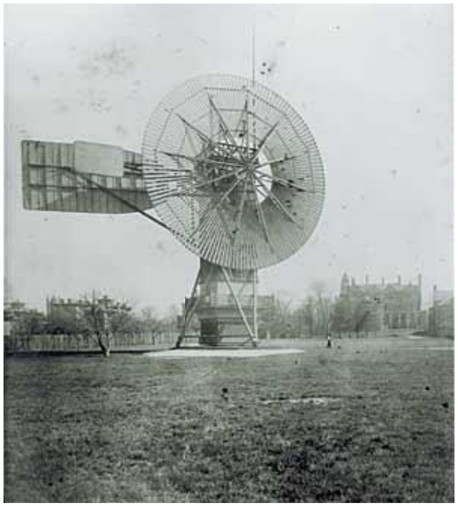


Figure 2. Charles F. Brush's wind turbine (1887, Cleveland, Ohio), the world's first *automatically operating* wind turbine for electricity generation.

the size of the machines has increased dramatically. Nowadays, the typical values for power output of the modern turbines deployed around the world are about 1.5 to 3.5 MW with blade lengths of more than 40 m for onshore and 60 m for offshore applications. Simultaneously, the cost per kilowatt has decreased significantly, and the efficiency, reliability, and availability of the machines have definitely improved.

New multidisciplinary computer design tools [4],[5], able to simulate, analyze, and redesign in a concurrent engineering way the aerodynamics, mechanics, and electrical and control systems under several conditions and external scenarios [6],[7],[8], have extended the capability to develop more complex and efficient wind turbines. In this new approach (Fig. 3), the control system designs, and the designers' understanding of the system's dynamics from the control standpoint, are playing a central role in new engineering achievements.

Far better than in the old days, when the design of any machine was carried out under a rigid and sequential strategy, starting from the pure aerodynamics and following with the mechanical, the electrical, and finally the control system design, the new tools have opened the door to a more central role for control engineers. The new philosophy brings a concurrent engineering approach, where all the engineering teams work simultaneously to achieve the optimum wind turbine design. This strategy allows the control engineers to interact with designers from the other fields from the very beginning, discussing and changing the aerodynamics, mechanics, and electrical systems to improve the dynamic behavior, efficiency, reliability, availability, and cost, and finally to design the most appropriate controllers for the machine.

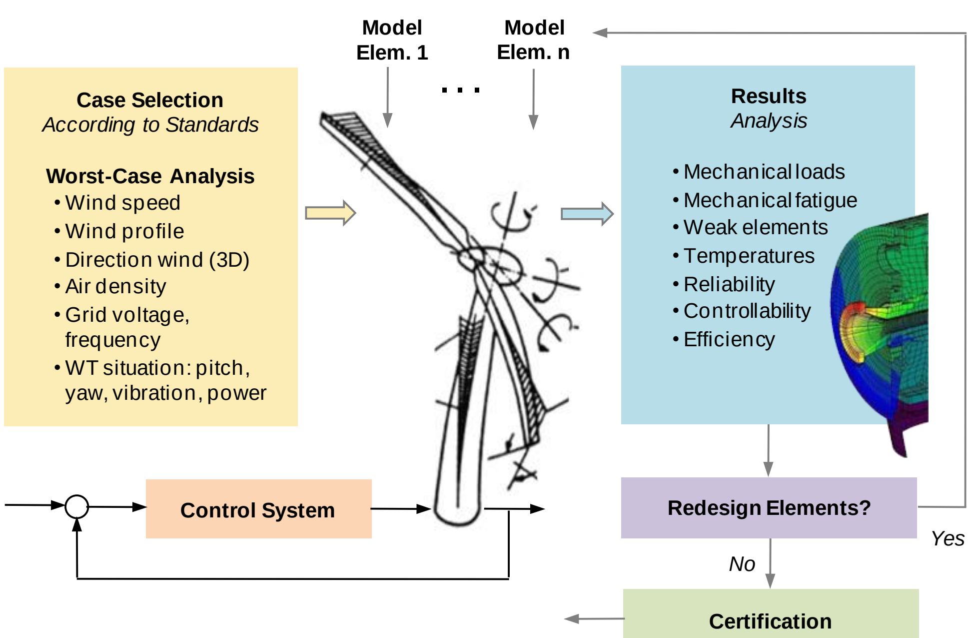


Figure 3. Multidisciplinary computer design tools for wind turbine design.

Nowadays, there are essentially two types of wind turbines: constant-speed and variable-speed machines. Until the late nineties, the constant-speed concept dominated the market. Today, it still represents a significant share of the operating wind turbines, but newer requirements have led to the emergence of variable-speed designs [5],[9],[10],[11].

Three main alternative strategies are used for regulating the amount of power captured by the rotor: passive stall control or fixed pitch, variable pitch control, and active stall control. So far, over the entire range of wind turbine sizes, no one of these strategies has taken the lead over the others. However, as machines get larger and power production increases, the trend is toward pitch control and active stall control [5],[9],[10],[11].

The configuration of a fixed-speed wind turbine is based on a gearbox and an asynchronous generator, which is usually a squirrel-cage induction generator to reduce costs. The gearbox links the wind turbine shaft with the rotor of a fixed-speed generator, providing the high rotational speed required by the generator. The generator produces electricity through a direct grid connection, and a set of capacitors is used to compensate reactive power. Due to lack of a frequency converter, the generator speed is dictated by the grid frequency. One disadvantage of fixed-speed operation is poor aerodynamic efficiency, particularly at partial-load operation. From the electrical system's standpoint, another disadvantage is that this type of operation has a detrimental effect on voltage because asynchronous generators demand reactive power from the grid.

Another alternative to the popular squirrel-cage asynchronous generator is the so-called slip control method, which adjusts the slip continuously. In this case, a wound rotor is connected to some variable resistors through slip rings. By changing the electrical resistance of the rotor, small changes in the rotational speed variation of about 10% above the synchronous speed can be compensated for without varying the generator output frequency.

Many options have been developed to achieve some degree of speed variation: (1) dual-speed generators with pole switching (the use of a lower speed in low wind conditions improves performance and reduces noise emissions); (2) variable-resistance asynchronous generators for a low range of variable speed; (3) doubly fed induction generators (DFIGs) for a moderate range of variable speed; and finally, (4) direct-drive multipole synchronous generator systems and (5) hybrid systems (combination of multipole generators with small gearboxes), both for a wide range of variable speed.

Especially dominant in new markets is the DFIG, also called the wound rotor induction generator. In this machine, the stator windings are directly connected to the grid, while a frequency converter interfaces between the standard wound rotor and the grid. The stator winding connection carries most of the power production, although the frequency converter may carry up to a third of the total power, depending on the operating mode. This configuration allows the machine to control the slip in the generator, and thus the rotor speed can vary moderately, achieving better aerodynamic efficiency. Furthermore, as the converter controls the rotor voltage magnitude and phase angle, partial control of active and reactive power is also possible.

Finally, another approach, which will probably dominate in offshore applications, is the multipole synchronous generator connected to the grid through a power electronic converter that handles the full power production. This concept, also called the direct-drive machine, takes advantage of the wide speed range allowed by the full-scale frequency converter. The generator can operate at any rotational speed, allowing operation to track the optimal speed for each wind condition. Among the main advantages of this approach are low maintenance costs and high reliability due to omission of the gearbox, improved aerodynamic efficiency, and the ability to assist grid voltage control.

A generic qualitative power curve for a variable-speed pitch-controlled wind turbine is shown in Fig. 4. Four zones and two areas are indicated in the figure [12]. The rated power P_r of the wind turbine (that is, the actual power supplied to the grid at wind speed greater than V_r) separates the graph into two main areas. Below rated power, the wind turbine produces only a fraction of its total design power, and therefore an optimization control strategy needs to be performed. Conversely, above rated power, a limitation control strategy is required.

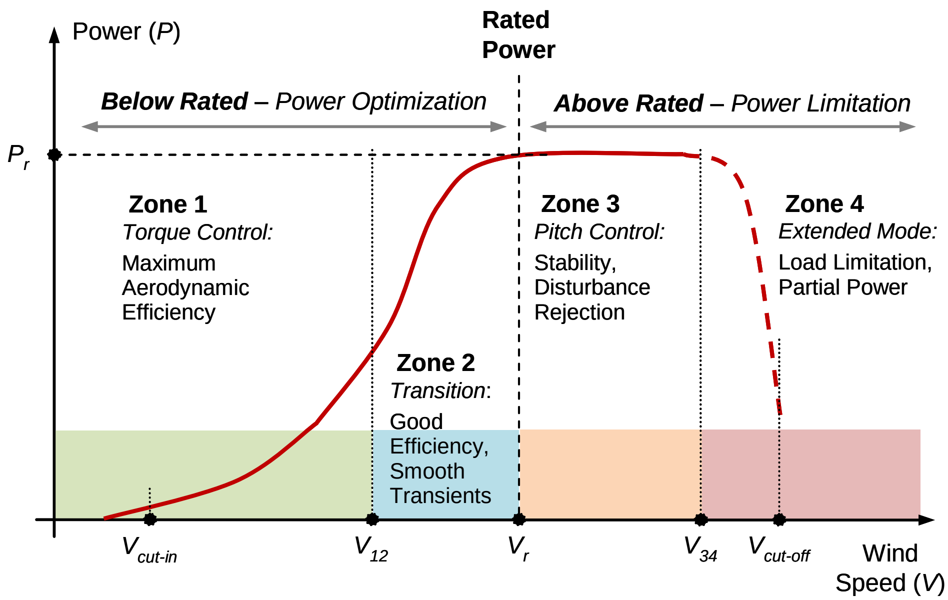


Figure 4. Power curve of a wind turbine and control zones.

For passive-stall-controlled wind turbines, in which the rotor blades are fixed to the hub at a specific angle, the generator reaction torque regulates rotor speed below rated operation to maximize energy capture. Above a specific wind speed, the geometry of the rotor induces stall. In this manner, the power delivered by the rotor is limited in high wind conditions thanks to a particular design of the blades that provokes loss of efficiency.

In pitch control, the power delivered by the rotor is regulated either by pitching the blades toward the wind to maximize energy capture or by pitching to feather to discard the excess power and ensure that the mechanical limitations are not exceeded. At rated operation, the aim is to maintain power and rotor speed at their rated value. To achieve this, the torque is held constant and the pitch is continually changed following the demands of a closed-loop rotor speed controller that optimizes energy capture and follows wind speed variations. In contrast, below rated operation there is no pitch control; the blade is set to a fine pitch position to yield higher power capture values while the generator torque itself regulates the rotor speed.

Active stall control is a combination of stall and pitch control. It offers the same regulation possibilities as the pitch-regulated turbine but uses the stall properties of the blades. Above rated operation, the control system pitches the blades to induce stall instead of feathering. In this technique, the blades are rotated only by small amounts and less frequently than for pitch control.

Solar Energy

A handful of thermal solar energy plants, most of them experimental, have been developed over the last two decades. The Solar One power tower [13], developed in Southern California in 1981, was in operation from 1982 to 1986. It used 1,818 mirrors, each 40 m², for a total area of 72,650 m². The plant was transformed into Solar Two by adding a second ring of larger (95 m²) heliostats and molten salts as a storage medium. This gave Solar Two the ability to produce 10 MW and helped with energy storage, not only during brief interruptions in sunlight due to clouds, but also to store sufficient energy for use at night. Solar Two was decommissioned in 1999 but proved it could produce power continuously around the clock.

The Solar Tower Power Plant SSPS was developed in 1980 in the Plataforma Solar de Almeria (PSA) on the edge of the Tabernas Desert in Spain (Fig. 5). The plant had 92 heliostats (40 m²) producing 2.7 MWth at the focal point of the 43-m-high tower where the heat was collected by liquid sodium. The PSA has a number of experimental plants such as the CESA-1 7-MWth central receiver system and the SSPS-OCS 1.2-MWth parabolic-trough collector system with associated thermal storage.

Figure 5. Plataforma Solar de Almeria (PSA).

The Solar Energy Generating Systems (SEGS) [14] begun in 1984 in the Mojave Desert in California uses parabolic-trough technology (Fig. 6). SEGS is composed of nine solar plants and is still the largest solar-energy-generating facility in the world with a 354-MW installed capacity. The plants have a total of 936,384 mirrors and cover more than 6.5 km^2. Lined up, the parabolic mirrors would extend more than 370 km.

The number of commercial solar power plants has been increasing in the last few years. New installations include the 10-MW (PS10) and the 20-MW (PS20) power tower (Fig. 7) plants; the 50-MW Solnova 1 and Solnova 3 trough plants designed, built, and operated by Abengoa Solar near Seville in Southern Spain; and the 50 MW Andasol 1 and Andasol 2 plants owned by ACS Group.

Photo credit: Alan Radecki

Figure 6. SEGS plants III-VII in California, U.S.A.

Solar power plant systems cannot be controlled with simple control strategies; they require advanced algorithms to compute the solar reflector positions as well as for self-calibration and prediction of the reflectors [15]. The sun vector needs to be computed, and for each heliostat, the normal vector is computed such that it divides the angle formed by the sun vector and the vector joining the center of the heliostat with the receiver. The current trend in solar concentrator tracking systems is to use open-loop controllers that compute the direction of the solar vector based on location and time. Nevertheless, error sources such as time of day, sun model, latitude and longitude of the site, heliostat position in the field, and control interval increase the complexity of the control system. Structural and mechanical sources of error, mainly due to tolerances (joints, encoder) and incorrect mirror facet alignment (optical errors), further add to the approximations in calculating solar position and other variables.

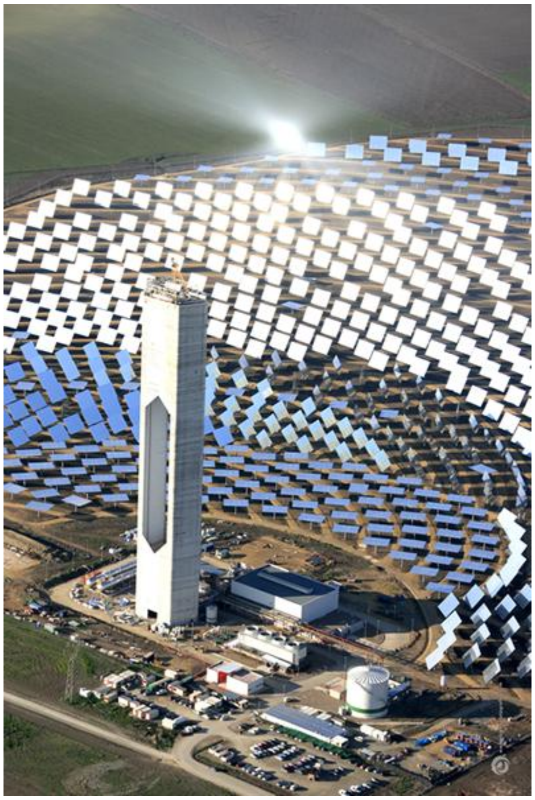


Figure 7. Abengoa Solar PS 20 power tower (Sevilla, Spain).

Heuristic control algorithms and CCD cameras have been used [16] to cope with some of these errors. The sunbeam centroid position errors are used to calibrate heliostat tracking parameters. The system can also be used during operation, as an individual heliostat can be deviated from its spot to correct its offset in real time.

To avoid deterioration due to excessive thermal gradients in central volumetric receivers, multi-aiming strategies are used [17] to obtain an appropriate flux distribution. Individual heliostats are deliberately aimed at different aiming points in such a way that more uniform irradiance is obtained in the central receiver.

Parabolic trough systems concentrate sunlight onto a receiver pipe located along the focal line of a trough collector. A heat transfer fluid, typically synthetic oil, is heated as it flows along the receiver pipe. For maximum efficiency, a constant supply of hot oil is required at some prespecified temperature, despite variations in the ambient temperature, inlet temperature, and direct solar radiation. Over the last 25 years, considerable research has been devoted to improving the efficiency of solar thermal power plants with distributed collectors in terms of control and optimization. Activities performed by control groups related to this field cover modeling, identification and simulation, classical proportional-integral-derivative control (PID), feedforward control (FF), model-based predictive control (MPC), adaptive control (AC), gain-scheduled control (GS), cascade control (CC), internal model control (IMC), time delay compensation (TDC), optimal control (LQG), nonlinear control (NC), robust control (RC), fuzzy logic control (FLC), and neural network control (NNC) . Most of this work is summarized in [12]. The control of steam-generating parabolic trough systems is a more challenging problem [18].

Smart Grids

Power systems are fundamentally reliant on control, communications, and computation for ensuring stable, reliable, efficient operations. Generators rely on governors and automatic voltage regulators (AVRs) to counter the effects of disturbances that continually buffet power systems, and many would quickly lose synchronism without the damping provided by power system stabilizers (PSSs). Flexible AC transmission system (FACTS) devices, such as static var compensators (SVCs) and high-voltage DC (HVDC) schemes, rely on feedback control to enhance system stability. At a higher level, energy management systems (EMSs) use supervisory control and data acquisition (SCADA) to collect data from expansive power systems and sophisticated analysis tools to establish secure, economic operating conditions. Automatic generation control (AGC) is a distributed closed-loop control scheme of continental proportions that optimally reschedules generator power setpoints to maintain frequency and tie-line flows at their specified values.

Historically, distribution systems have had a minimal role in power system operation and control. Many distribution utilities have employed demand management schemes that switch loads such as water heaters and air conditioner to reduce load during peak conditions or emergency situations. The controllability offered by such schemes has been rather limited, however. This lack of involvement of distribution is largely a consequence of the technical difficulties involved in communicating (with sufficient bandwidth) with consumers. Smart grids promise cost-effective technology that overcomes these limitations, allowing consumers to respond to power system conditions and hence actively participate in system operations.

Smart grid concepts encompass a wide range of technologies and applications. We describe a few below that are currently in practice with the caveat that, at this early stage in the development of smart grids, the role of control, especially advanced control, is limited:

- Advanced metering infrastructure (AMI) is a vision for two-way meter/utility communication. Two fundamental elements of AMI have been implemented. First, automatic meter reading (AMR) systems provide an initial step toward lowering the costs of data gathering through use of real-time metering information. They also facilitate remote disconnection/reconnection of consumers, load control, detection of and response to outages, energy theft responsiveness, and monitoring of power quality and consumption. Second, meter data management (MDM) provides a single point of integration for the full range of meter data. It enables leveraging of that data to automate business processes in real time and sharing of the data with key business and operational applications to improve efficiency and support decision making across the enterprise.

- Distribution management system (DMS) software mathematically models the electric distribution network and predicts the impact of outages, transmission, generation, voltage/frequency variation, and more. It helps reduce capital investment by showing how to better utilize existing assets, by enabling peak shaving via demand response (DR), and by improving network reliability. It also facilitates consumer choice by helping identify rate options best suited to each consumer and supports the business case for renewable generation solutions (distributed generation) and for electric vehicles and charging station management.

- Geographic information system (GIS) technology is specifically designed for the utility industry to model, design, and manage their critical infrastructure. By integrating utility data and geographical maps, GIS provides a graphical view of the infrastructure that supports cost reduction through simplified planning and analysis and reduced operational response times.

- Outage management systems (OMSs) speed outage resolution so power is restored more rapidly and outage costs are contained. They eliminate the cost of manual reporting, analyze historical outage data to identify improvements and avoid future outages, and address regulatory and consumer demand for better responsiveness.

- Intelligent electronics devices (IEDs) are advanced, application-enabled devices installed in the field that process, compute, and transmit pertinent information to a higher level. IEDs can collect data from both the network and consumers' facilities (behind the meter) and allow network reconfiguration either locally or on command from the control center.

- Wide-area measurement systems (WAMS) provide accurate, synchronized measurements from across large-scale power grids. They have been implemented in numerous power systems around the world, following initial developments within the Western Electricity Coordinating Council (WECC) through the early 1990s [19]. WAMS consist of phasor measurement units (PMUs) that provide precise, time-stamped data, together with phasor data concentrators that aggregate the data and perform event recording. WAMS data plays a vital role in post-disturbance analysis, validation of system dynamic models, FACTS control verification, and wide-area protection schemes. Future implementation of wide-area control schemes are expected to build on WAMS.

- Energy management systems (EMSs) at customer premises can control consumption, onsite generation and storage, and potentially electric vehicle charging. EMSs are in use today in large industrial and commercial facilities and will likely be broadly adopted with the rollout of smart grids. Facility energy management can be seen as a large-scale optimization problem: Given current and (possibly uncertain) future information on pricing, consumption preferences, distributed generation prospects, and other factors, how should devices and systems be used optimally?

Smart grid implementations are occurring rapidly, with numerous projects under way around the world. Fortum's "intelligent management system of electric consumption" uses advanced metering devices to gather customer's consumption data and metering management systems to store and analyze this information. Vattenfall's "automatic household electricity consumption metering system" is another example of a European project that is focused on remote measurement of consumers. Also, projects such as Elektra's "distribution management system" improve quality of service by implementing next-generation devices to manage and control information (SCADA), DMS to plan and optimize distribution system operations, and ArcFM/Responder to improve outage response times.

Market Sizes and Investment [15], [20], [21]

Wind Energy

With many thousands of wind turbines in operation, the total worldwide installed capacity is currently about 160 GW. According to the World Wind Energy Association, the net growth rate is expected to be more than 21% per year. The top five countries, the United States, Germany, Spain, China, and India, currently share about 73% of the world capacity.

The cost of electricity from utility-scale wind farms has dropped by more than 80% over the last 20 years, reaching values of about $2.2 and $4.6 million per megawatt for onshore and offshore applications, respectively, in 2010. According to the U.S. Department of Energy, the capital cost of onshore applications can be further reduced to about 10% of current cost over the next two decades. In addition, several countries have adopted special programs to subsidize and promote wind energy. Among the most successful ones are the feed-in-tariff (FiT) programs and the production tax credit (PTC) programs.

The FiT programs have been adopted by more than 60 countries and states all over the world, including some of the top-producing countries: Germany, Spain, Canada, and Denmark. They typically include: (1) guaranteed grid access for the wind farm, (2) long-term contracts to sell the electricity produced by the wind turbines, and (3) purchase prices for distributed renewable generation that are substantially higher than the retail price of electricity (and will gradually be reduced toward grid parity).

A production tax credit program has been adopted in the United States. This federal incentive provides a credit of a varying number of cents per kilowatt-hour (currently 2.1 cents). Since its establishment in 1992, the PTC has had an "on-again/off-again" status, which has contributed to boom-bust cycles of the wind energy industry in the U.S.

More wind power was installed in the EU in 2008 than any other electricity-generating technology [22]. In leading the EU power sector for the first time, wind accounted for 36%, or 8,484 MW, of new capacity based on investments of €11 billion in the EU alone. By comparison, the gas sector created 6,932 MW (29%) of new capacity, new solar photovoltaic installation capacity was 4,200 MW (18%), new capacity from oil was 2,495 MW (10%), from coal, 762 MW (3%), and from hydro, 473 MW (2%). The 65 GW of EU wind energy capacity installed by the end of 2008 will avoid the emission of 108 million tons (Mt) of CO_2 annually—equivalent to taking 55 million cars off the road and equaling 24% of the EU-27's Kyoto obligation.

The EU wind energy sector directly employed approximately 108,600 people in 2007 [23]. Including indirect employment, the wind energy sector employs 154,000 in the EU. On average, 12,047 new direct wind energy jobs have been created per year in the five-year period 2002-2007.

Solar Energy

Solar photovoltaic generation installed capacity has grown about 40% since 2002. Thermal power plants are growing rapidly, with more than 2 GW under construction and some 14 GW announced through 2014. Spain is the epicenter of solar thermal power development with 22 projects under development for 1,081 MW capacity [24]. In the United States, 5,600 MW of solar thermal power projects have been announced. Currently (as of July, 2010), 679 MW of CSP capacity are installed worldwide. The U.S. is the market leader in terms of installed capacity with 63% market share, followed by Spain with 32%. These two markets will continue to be crucial for the development of the industry into the next decade, with Spain accounting for the largest share of projects under construction with almost 89%. Solar generation is taking off in emerging regions as well; both China and India have announced plans for large-scale solar plants.

On July 3, 2010, U.S. President Obama announced that "the Department of Energy is awarding nearly $2 billion in conditional commitments to two solar companies. The first is Abengoa Solar, a company that has agreed to build one of the largest solar plants in the world right here in the United States. Once completed, this plant will be the first large-scale solar plant in the U.S and it will generate enough clean, renewable energy to power 70,000 homes. The second company is Abound Solar Manufacturing, which will manufacture advanced solar panels at two new plants. When fully operational, these plants will produce millions of state-of-the-art solar panels each year" [25]. The Solar Energy Technologies Program (SETP, or Solar Program) launched by the U.S. Department of Energy works to develop cost-competitive solar energy systems for America. More than $170 million is spent each year in research and development (R&D) on both photovoltaics and concentrating solar power. The greatest R&D challenges are reducing costs, improving system performance, and finding new ways to generate and store energy captured from the sun [26].

In terms of the technology employed, the market is dominated by parabolic trough technology, which accounts for 88% of operating plants and 97.5% of projects under construction.

The China Renewable Energy Scale-up Programme (CRESP) recently released a report on solar power generation economic incentive policies. The report suggested measures such as taxation and financial preference, discounted loans, and direct financial subsides and included information on preferential price policies and management, increasing technical research and development investment, strengthening R&D capacity, establishing technical standards, management regulations, and an authentication system. The Chinese National Development and Reform Commission’s 11th 5-year plan (2006-2010) includes 200 MW of commercial CSP plants [24]. China is currently the market leader in the PV manufacturing industry. A licensing agreement to build at least 2 GW of solar thermal power plants in China over the next 10 years was recently announced [6]. The deal represents the country’s first major move into concentrating solar thermal power. The Chinese government also recently announced aggressive plans to increase the country’s renewable power generation capacity to 15% by 2020 [27].

India's “New Solar Mission” [28] is the most ambitious solar energy development plan in the world. Its goal is for the country to be generating 20 GW of energy from sunlight by 2022. Going by International Energy Agency forecasts, this will make India the producer of almost three-quarters of the world's total solar energy output. The “New Solar Mission” has set forward a three-stage approach to hitting the 2022 target. The first stage will comprise 1,100 MW of grid-connected power and up to 200 MW of nongrid capacity by 2013.

Smart Grid

The smart grid’s technology market is expected to see 20% annual growth, going from $70 billion in 2009 to about $171 billion by 2014, according to market reports by Specialist in Business Information (SBI). In 2010 alone, the U.S. and China will spend more than $7 billion on smart grid technology and implementation, according to the research and consulting firm Zpryme. Due to these and many other initiatives, the smart grid communication market is expected to have opportunities of $16 to $20 billion per year, and transmission and distribution infrastructures will see investment of $41 billion through 2015.

The European Electricity Grid Initiative (EEGI) is one of many European projects focused on smart grid research and implementation. One of the EEGI’s main goals is to achieve the 20-20-20 climate package challenge: a 20% cut in emissions of greenhouse gases by 2020 (compared with 1990 levels), a 20% increase in renewable energy use by 2020, and a 20% cut in energy consumption by 2020. The total budget for this program is estimated at €2 billion ($2.54 billion). U.S. initiatives include the “Grid 2030 Vision,” which consists of achieving three major elements: a national electricity backbone, regional interconnections, and local distribution. To achieve this vision, the U.S. government plans on investing more than $38 billion to create the first “smart grid with continental dimensions.”

Application Challenges/Opportunities for Research [5], [20], [29]

Wind Energy

The enormous and unique worldwide possibilities for large-scale wind energy development over the next few decades depend greatly on how critical technology challenges are addressed. New ideas and control engineering solutions are needed to open virgin global markets. Among others, we emphasize the seven technology challenges (TCs) listed in Table 1.

Table 1. Wind Energy Challenges

TC.1	Cost reduction for a zero-incentive situation
TC.2	Efficiency maximization
TC.3	Mechanical load attenuation
TC.4	Large-scale grid integration and penetration
TC.5	Extreme weather conditions
TC.6	Offshore wind turbines
TC.7	Airborne wind energy systems

- **TC.1.** Although the cost of utility-scale wind farms has dropped by more than 80% over the last 20 years, most wind energy systems, including all offshore applications, still need significant government support to be feasible. However, that subsidy cannot be sustained long term at large scale. Thus, the long-term economic sustainability of wind energy imperatively requires improving the wind energy business model so that costs are similar to conventional power generation. This important objective will be achieved by (1) the development of new control systems, materials, blades, electromechanics, and power systems for the wind turbine, and (2) automatic low-cost blade and tower manufacturing systems for mass production.

- **TC.2.** Efficiency maximization implies generating more energy over the low-to-medium operating wind spectrum. Research opportunities for efficiency maximization include: (1) smart blades with advanced airfoils, new sensors and actuators, and specific control systems; (2) new rotor configurations; (3) variable-diameter rotors, which could significantly increase the efficiency of the turbine by presenting a large area to capture more energy in low winds and a reduced area to protect the system in high winds; and (4) turbines with taller towers to capture more energy in regions with high wind shear. In all cases, advanced control strategies to damp out tower motion by using blade pitch and generator torque control are critical.

- **TC.3.** Large multi-megawatt machines need very large rotor diameters. To allow the rotor to grow larger and capture more energy, new active and independent pitch control and torque control systems must be developed to reduce towertop motion, power fluctuations, asymmetric rotor loads, mechanical fatigue, and individual blade loads, achieving higher reliability and lower maintenance [5],[29],[30]. These developments will also help improve gearbox reliability.

- **TC.4.** In a large-scale wind energy scenario, the wind farms will have to support the grid by providing (1) fault ride-through capability; (2) voltage regulation and reactive power control; (3) primary frequency control; (4) oscillation damping; (5) low harmonics content; and by (6) avoiding power flickers and (7) carrying a share of power control capability for the grid. There is no generally accepted "maximum" level of wind penetration. The limit for a particular grid will depend on existing generating plants, wind turbine technology, wind turbine control systems, grid demand management, pricing mechanisms, grid capacity and topology, storage type and availability, and wind resource reliability and diversity [31]-[33].

- **TC.5.** Extreme cold and humid weather conditions can stop the wind turbines from working during winter months due to ice formation on the blades in quantities that would degrade the turbine performance and cause blade imbalance. By integrating ice protection systems in the blades and managing them with an appropriate control system, the wind turbines will produce a greater amount of power during winter, opening new markets at northern latitudes and many offshore locations such as the freshwater Great Lakes in the U.S. and Canada.

- **TC.6.** Offshore wind power is a promising technology with enormous energy potential. With fewer logistic constraints than onshore applications, over the next few years offshore turbines will reach a typical size of 5 to 8 MW and a rotor diameter of more than 150 m, adopting tip speeds slightly higher than those of onshore turbines. The offshore foundation system depends on the water depth. Most of the projects installed so far have been in water less than 22 m deep, with a demonstration project in Scotland at a depth of 45 m. Shallow-water technology currently uses monopiles for about 20-m depths. Very deep water applications, with floating foundations, still need reliable solutions, including advanced control systems to deal with wind, ocean waves, tides, ice formation, and water currents simultaneously. In addition, research opportunities for offshore applications include: (1) new ideas to reduce the cost from the current 20 cents/kWh to 7-9 cents/kWh by 2030, according to U.S. Department of Energy goals; (2) remote, intelligent turbine condition monitoring and self-diagnostic systems; (3) dedicated deployment vessels; (4) analytical models to characterize wind, ocean currents, tides, ice, and ocean waves; (5) high reliability; (6) predictive maintenance techniques; and (7) grid technologies for electricity transmission back to shore.

- **TC.7.** An airborne wind energy system is a wind turbine that is supported in the air without a tower. Two technologies have been proposed: ground generator systems and aloft generator systems. In both cases, the wind turbines have the advantages of an almost constant and high wind speed and a low-cost structure without the expense of tower construction. Advanced multivariable robust control strategies for attitude and position control of the flying structure and reliable control algorithms to govern the system under bad weather conditions, such as lightning or thunderstorms, are critical. No commercial airborne wind turbines are in regular operation yet.

Solar Energy

One of the 21st Century's Grand Challenges for Engineering identified by the U.S. National Academy of Engineering is to make solar energy economical: “Overcoming the barriers to widespread solar power generation will require engineering innovations in several arenas—for capturing the sun’s energy, converting it to useful forms and storing it for use when the sun itself is obscured” [34].

Solar energy can be made more economical by reducing investment and operating costs and by increasing solar plant performance. The solar field represents the largest share of the cost of any CSP plant. Depending on the technology, this cost could vary from about 43% for tower and Fresnel technology to almost 60% for parabolic trough and dish Stirling CSP plants. The most significant cost reductions are likely to come from innovations in solar field design, which could bring down the levelized cost of energy (LCOE) by 15% to 28%, depending on the technology.

Advanced control can help reduce operating costs and increase solar plant performance. The main control challenges are:

- Optimal robust control techniques able to maintain the operating temperature as close to optimum as possible despite disturbances such as changes in solar irradiance level (caused by clouds), mirror reflectivity, and other operating conditions.

- Optimal and hybrid control algorithms that determine optimal operating points and modes and take into account the production commitments, expected solar radiation, state of energy storage, and electricity tariffs.

- Modes and methods for forecasting solar radiation using heterogeneous information (cameras, satellites, weather forecasts).

- Algorithms to estimate main process variables and parameters from heterogeneous and distributed measurements (oil temperature and solar radiation at different parts of the field, mirror reflectivity, thermal losses).

- Automatic mirror cleaning devices. The main factor degrading the optical performance of concentrating mirrors is accumulation of dirt on the mirror surface. Cleaning mirrors represents a considerable expense in manpower and water, usually a scarce resource where solar plants are located. Automatic devices need to be developed that minimize the use of water and degradation of the reflective surface.

- Heliostat self-calibration mechanisms. Heliostats need to be retuned periodically because of errors in the sun model, latitude and longitude of the site, heliostat position in the field, mechanical errors, optical errors, and the like. Heliostat recalibration may represent an important cost in manpower and time when done manually. Methods are needed for fast, automatic, online recalibration of heliostats.

- Fault detection and isolation in solar power plants. Algorithms are needed to detect and isolate faults and malfunctions in power plants, such as detection of hot spots, receivers with broken glass covers or vacuum losses, and heliostat faults.

Smart Grids

A significant challenge associated with smart grids is the integration of renewable generation. Traditionally, power systems have addressed the uncertainty of load demand by controlling supply. With renewable energy sources, however, uncertainty and intermittency on the supply side must also be managed. Demand response and load control—direct and indirect mechanisms to adjust consumption—are required. Direct load control—load adjustments made directly by the utility—must be nondisruptive in the sense that consumers are unaware of the control actions. Indirect demand response, such as providing price signals or other incentives for consumers to modify their loads, is already being practiced in commercial and industrial facilities, and some pilot projects are under way for homes as well. Modeling, optimization, and control issues are crucial. For example, instability may result in both the market side and the grid side if real-time pricing is implemented without adequate understanding of control principles. With slower time scales for price adjustment (a more likely scenario), instability will not be a primary concern, but the cost of suboptimal performance may be considerable. For example, slow price adjustment limits the ability for demand to track variations in renewable generation output, increasing the reliance on storage and nonrenewable sources for power balance.

Also on the consumer side, the integration of storage, distributed generation, and plug-in (possibly hybrid) electric vehicles all present both opportunities and challenges. Any local storage or generation can, at least in principle, help with managing varying grid supply. But each component has characteristics that must be considered and incorporated in the control scheme. Plug-in vehicles, when broadly deployed, are especially notable in that they represent a large load (charging rates for individual vehicles may be higher than typical peak load in a home), and consumers will expect full (or at least commute-sufficient) state of charge by morning. Some neighborhood or higher level control will likely be necessary to regulate the overnight load. Given that wind generation is typically at maximum overnight, such controls will play a vital role in achieving optimal use of wind resources.

As noted above, price signals are already being communicated to users by utilities or service providers with media ranging from advanced metering infrastructure (AMI) to the Internet. Here, too, control-relevant issues arise, and on both the supply and demand sides. Thus, a utility needs to generate control signals (a simple example is time-of-use prices, which impose different consumption costs at different times of the day according to a fixed and broadcast schedule) that, based on models of expected consumer behavior, will maximize the utility's objective—incorporating profitability, renewable energy use, stability/loadability requirements, and other criteria. Conversely, consumers must determine how to schedule their load and, where available, how and when to operate distributed generation and storage resources to best satisfy their objectives. Furthermore, large consumers and utilities will sometimes negotiate together for load profiles and prices, thereby combining two already large optimization problems into a multi-objective problem.

Another promising focus for the controls community related to the smart grid is power electronics, which is playing an increasingly important role in grid connection of loads and generation. Devices that use power electronics for grid connection include plug-in electric vehicles, variable-frequency drives, and many of the newer forms of renewable generation. Power electronic interfaces tend to decouple device behavior from grid disturbances. This decoupling can have a detrimental effect on the response of the grid frequency and can accentuate voltage collapse. Power electronic interfaces can be controlled in ways that alleviate these undesirable effects, within the bounds of physical capabilities. The required controls are location-specific and also vary with system conditions.

Complexities abound across the transmission and distribution infrastructure, with inherent interactions between continuous dynamics and discrete events. Power systems should therefore be modeled as large-scale hybrid dynamical systems, where the continuous dynamics are best represented by differential-algebraic models. State dimension is frequently in the tens of thousands. Smart grids imply incorporating cyberinfrastructure into this physical model, and doing so in a way that is computationally feasible yet preserves the dominant characteristics of the cyber-physical interactions. Furthermore, smart grids will add large numbers of devices that actively participate in systemwide control actions. Modeling each individual device is infeasible, yet their consolidated response must be accurately represented. The overall modeling problem is multiscale in terms of both time and model fidelity.

Finally, on the architectural front, smart grids will require new distributed control structures to fully exploit the new, and widely distributed, sensors and actuators. It is infeasible for a centralized controller to address every controllable load individually, yet actions taken by local controllers must be consistent with global performance objectives.

Conclusions

Most national energy policies worldwide aim at ensuring an energy portfolio that supports a cleaner environment and stronger economy and that strengthens national security by providing a stable, diverse, domestic energy supply. Clean energy is a global and urgent imperative. Renewable generation, especially from wind and solar, and smart grid concepts are critical technologies needed to address global warming and related issues. The key challenge is to reduce the cost of renewable energies to affordable levels. Control and related technologies will be essential for solving these complex problems.

Acknowledgments

The authors would like to acknowledge Thomas Bak, Paul Houpt, Dragan Obradovic, Tomás Sanchez, Desirée Ortiz Marín, Peter Terwiesch, Kishan Baheti, Pravin Varaiya, Eyad Abed, Jakob Stoustrup, and Jorge Pereira for their contributions to this section.

Selected recommendations for research in the control of renewable generation and smart grids:

- For concentrated solar power plants, integrated control systems are needed that incorporate advanced estimation and forecasting, heliostat self-calibration, and hybrid/robust closed-loop control.
- Novel high-altitude systems promise tremendous improvement in wind power generation—but the associated, complex modeling and control challenges must first be addressed.
- Control is critical for realizing visions for smart grids—in particular, distributed decentralized control system architectures encompassing end-to-end communication and power flows are needed.

References

[1] U.S. Department of Energy. Office of Electricity Delivery and Energy Reliability, Recovery Act Financial Assistance Funding Opportunity Announcement, Smart Grid Investment Grant Program, DE-FOA-0000058, June 25, 2009.

[2] Office of the National Coordinator for Smart Grid Interoperability. NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 1.0, U.S. National Institute of Standards and Technology Special Publication 1108 [Online], 2009. Available at http://www.nist.gov/public_affairs/releases/upload/smartgrid_interoperability_final.pdf.

[3] P.C. Putnam. *Power from the Wind.* New York: Van Nostrand Reinhold, 1948.

[4] Garrad Hassan. GH Bladed software, 2010.

[5] M. Garcia-Sanz and C.H. Houpis. *Wind Energy Systems: Control Engineering Design.* Boca Raton, FL: Taylor & Francis, 2011.

[6] Germanischer Lloyd. Wind turbine standards and certification, 2010.

[7] IEC (International Electrotechnical Commission). IEC-61400, Wind turbine standards, 2007.

[8] NREL (National Renewable Energy Laboratory). NWTC Design Codes (FAST), An Aeroelastic Design Code for Horizontal Axis Wind Turbines, 2010.

[9] AWEA (American Wind Energy Association). Wind Energy Basics [Online], 2007. Available at http://www.awea.org/newsroom/ pdf/Wind_Energy_Basics.pdf.

[10] T. Burton, D. Sharpe, N. Jenkins, and E. Bossanyi. *Wind Energy Handbook.* London: Wiley, 2001.

[11] E. Hau. *Wind Turbines. Fundamentals, Technologies, Application, Economics* (2nd ed.). Berlin: Springer, 2006.

[12] E.F. Camacho, F. Rubio, M. Berenguel, and L. Valenzuela. "A survey on control schemes for distributed solar collector fields (part 1 and 2)," *Solar Energy,* vol. 81, pp. 1240-1272, 2007.

[13] P. DeLaquil, B. Kelly, and R. Lessley. "Solar One Conversion Project," *Solar Energy Materials,* vol. 24, no. 1-4, pp. 151-161, December 2, 1991.

[14] D. Kearney. "Solar Electric Generating Stations (SEGS)," *IEEE Power Engineering Review*, vol. 9, no. 8, pp. 4–8, 1989. doi:10.1109/MPER.1989.4310850.

[15] E.F. Camacho, M. Berenguel, F.R. Rubio, and D. Martinez. *Control of Solar Energy Systems.* Berlin: Springer Verlag, 2010.

[16] M. Berenguel, F.R. Rubio, A. Valverde, M.R. Arahal, E.F. Camacho, and M. López. "An artificial vision-based control system for automatic heliostat position in offset correction in a central receiver solar power plant," *Solar Energy,* vol. 76, no. 5, pp. 523-653, June 2004. ISSN 0038-092X.

[17] F.J. García-Martín, M. Berenguel, A. Valverde, and E.F. Camacho. "Heuristic knowledge-based heliostat field control for the optimization on the temperature distribution in a volumetric receiver," *Solar Energy,* vol. 66, no. 5, pp. 355-369, August 1999. ISSN 0038-092X.

[18] L. Valenzuela, E. Zarza, M. Berenguel, and E.F. Camacho. "Direct steam generation in solar boilers," *IEEE Control Systems Magazine,* vol. 24, no. 2, pp. 15-29, 2004.

[19] J.Y. Cai, Z. Huang, J. Hauer, and K. Martin. "Current status and experience of WAMS implementation in North America," in Proc. 2005 IEEE/PES Transmission and Distribution Conference & Exhibition: Asia and Pacific, Dalian, China, 2005.

[20] U.S. Department of Energy. 20% Wind Energy by 2030. Increasing Wind Energy's Contribution to U.S. Electricity Supply, DOE/GO-102008-2567, July 2008.

[21] European Wind Energy Association. Wind Energy - The Facts. Earthscan, 2009.

[22] European Wind Energy Association. Winning with European Wind, EWEA 2008 Annual Report, June 2009.

[23] European Wind Energy Association. Wind at work: Wind Energy and Job Creation in the EU, January 2009.

[24] SolarPaces. Concentrating Solar Power Gobal Outlook 09 [Online], 2009. Available at http://www.solarpaces.org/Library.

[25] White House. Weekly address [Online], July 3, 2010. Available at, http://www.whitehouse.gov/blog/ 2010/07/03/weekly-address-a-solar-recovery#.

[26] U.S. Department of Energy. Solar Energy Technologies Program - 2009 Annual Report, 2009.

[27] Businesswire. [Online], 2010. Available at http://www.businesswire.com/portal/site/home/permalink/ ?ndmViewId =news_view&newsId=20100108006041&newsLang=en.

[28] J. Nehru. National Solar Mission Towards Building SOLAR INDIA [Online], 2009. Available at http://mnre.gov.in/pdf/mission-document-JNNSM.pdf.

[29] M. Garcia-Sanz. "Wind turbines: New challenges and advanced control solutions," *International Journal of Robust and Nonlinear Control,* vol. 19, no. 1, pp. 1-116, 10 January 2009. London: Wiley, ISSN: 1049-8923.

[30] E.A. Bossanyi. "Individual blade pitch control for load reduction," *Wind Energy,* vol. 6, no. 2, pp. 119–128, 2003.

[31] T. Ackermann. *Wind Power in Power Systems.* London: Wiley, 2005.

[32] S. Heier. *Grid Integration of Wind Energy Conversion Systems* (2nd ed.). London: Wiley, 2006.

[33] IEEE. "Working with wind: Integrating wind into the power system," *IEEE Power and Energy Magazine*, vol. 3, no. 6, November–December 2005.

[34] National Academy of Engineering. Grand Challenges for Engineering [Online], 2008. Available at http://www.engineeringchallenges.org/.

Related Content

The Impact of Control Technology report also includes more than 40 flyers describing specific "success stories" and "grand challenges" in control engineering and science, covering a variety of application domains. The ones below are closely related to the topic of this section.

Grand Challenges

- Control for Energy-Efficient Buildings – *P. Stluka and W. Foslien*
- Control for Grid Responsiveness – *I. Hiskens*
- Control for Smart Grids – *T. Samad*
- Control for Wind Power – *L. Fagiano and M. Milanese*

These flyers—and all other report content—are available at http://ieeecss.org/main/IoCT-report.

Cross-Application Perspectives: Tools and Platforms for Control Systems

Ken Butts and Andras Varga

Introduction

Control analysis and design methods are based on rigorous theoretic and systematic foundations. Regardless of application, the concepts of stability, controllability, observability, performance, and robustness are used to analyze the system or device to be controlled and design the control laws that satisfy the given system requirements. Due to the engineering nature of the discipline, control designers have enjoyed a rich history of simultaneous development of theoretical concepts and supporting numerical algorithms. These numerical algorithms are the basis for control-domain-specific computer-aided engineering toolkits that provide ready access to best practices, rigor, and scalability. Although the number of users for this tooling is relatively small, control engineering's fundamental, application-agnostic approach yields a vibrant and sustainable tool market.

Control-oriented software platforms represent integrated software environments whose tool chains cover most aspects of typical industrial control design workflows. One such workflow is the well-known system engineering "V" diagram of Fig. 1, where the requirements are processed through design and implementation with a focus on step-by-step process deliverables and milestones.

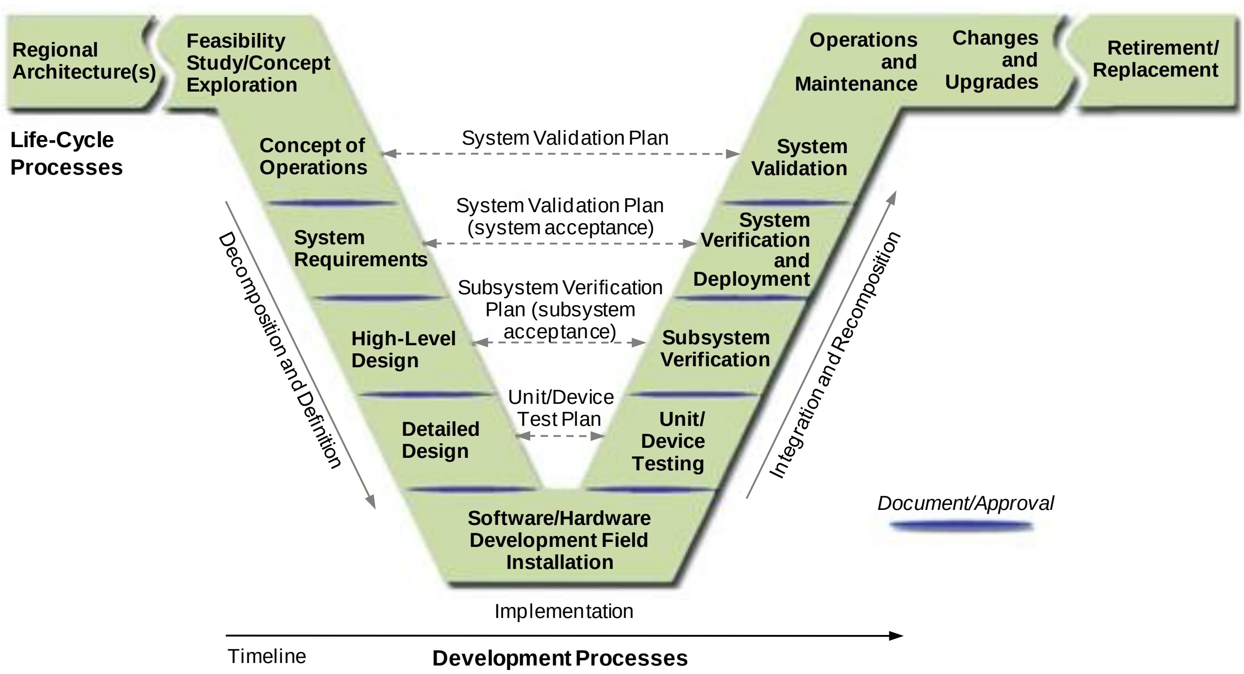
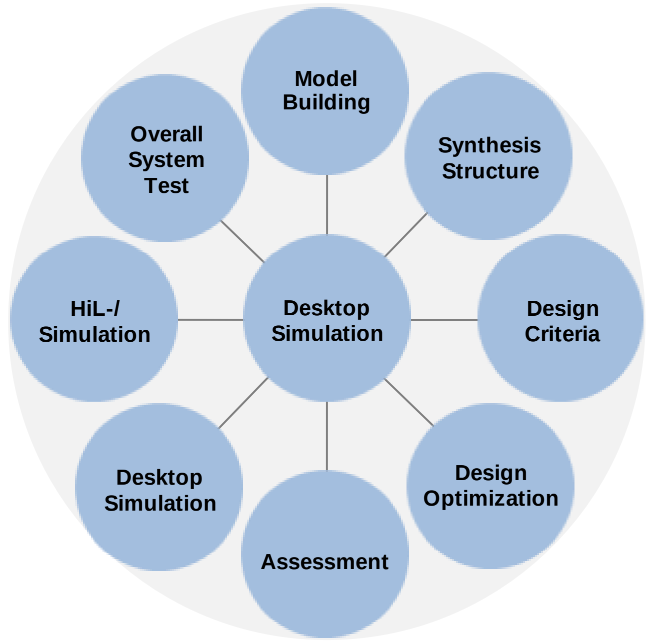


Figure 1. Systems engineering V diagram [1].

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Alternatively, a control engineering task view is shown in Fig. 2. From this perspective, one can see how an effective software platform supports requirements-centered control engineering iteration, where results are validated against the requirements and the requirements are occasionally updated to satisfy design tradeoffs. Typical commercial platforms provide tool chains with rich functionality for standard tasks such as control-oriented plant model building, control system design, or simulation-based assessment while simultaneously providing support to related aspects such as rapid prototyping, software/ hardware-in-the-loop verification, or code generation. The widespread use of these platforms in industry and academia confirms the needs for a broad spectrum of tools supporting all aspects of control systems development. Several success stories provide evidence that without such platforms, many advanced control applications, especially in the aerospace, automotive, and process industry areas, would not be possible.



Source: DLR

Figure 2. Model-based control system design.

Control systems are becoming ubiquitous, more critical, more heterogeneous, and more distributed. These systems must be developed reliably, predictably, and productively. Consider the statement by the U.S. President's Council of Advisors on Science and Technology: "It is difficult to overstate the contribution of networking and information technology (NIT) to America's security, economy, and quality of life. This contribution is the consequence of rapid advances in an array of technologies, some now ubiquitous, such as Internet search engines and wireless devices. Other technologies, such as simulation software and embedded systems, are essential to the effective performance of sectors that include national security, energy, health care, manufacturing, and transportation. The cumulative effect of these technologies on life in the United States and around the world has been profound and beneficial" [2]. Control engineering tools and platforms are used to aid modeling and analysis, address large-scale systems, reduce time-to-market, and mitigate the limits of ad hoc and manual design processes.

Control engineering tools and platforms are used to aid modeling and analysis, reduce time-to-market, and overcome the limits of ad hoc and manual design processes.

Specific Industrial Needs

Although the fundamentals of control engineering are common across application domains, some variations in approach can be identified. For example, model-building tools for the aerospace and automotive industries must primarily support physically oriented modeling, whereas for the process industries, due to the higher complexity of plant dynamics, model-building tools must use a data-driven plant modeling approach. The widely used block-diagram-based (also called *causal*) model-building approach has well-known intrinsic limitations, and therefore, in some domains (such as aerospace and automotive), object-oriented (also sometimes called *acausal*) modeling techniques are gaining greater acceptance, with obvious benefits of more physical insight and better reusability of model components.

Controller tuning philosophies also vary markedly across the aerospace, automotive, and process control industries. The control designer is solely responsible for parameter tuning in the aerospace industry, whereas in the automotive industry, downstream product specialists typically set the tuning parameters to meet vehicle performance targets. Auto-tuning algorithms are often used in the process industries to allow plant control engineers to perform control law tuning on site.

The control designer is solely responsible for parameter tuning in the aerospace industry, whereas in the automotive industry, downstream product specialists typically set the tuning parameters to meet vehicle performance targets. Auto-tuning algorithms are often used in the process industries to allow plant control engineers to perform control law tuning on site.

Assessment needs also vary significantly among activity domains. Although simulation-based assessment (also in conjunction with Monte Carlo techniques) is widely used across many industries, the assessment of safety-critical control applications (such as flight control) requires rigorous verification of the robustness of control laws in the presence of uncertainties. Since the verification effort can represent more than 50% of the total development cost of the control system, new optimization-based approaches (such as worst-case search) in conjunction with parallel computation techniques can contribute greatly to cost reduction.

Automatic code generation is widely used for rapid prototyping, especially in the automotive and aerospace industries. In particular, the requirements for code generation in the aerospace industry are more stringent because certifiable code imposes the use of certified code-generation tools as well, which are usually not part of common control design platforms. The gap in the tool chain may even necessitate manual recoding of control laws. No single platform can presently address the entire development workflow in aerospace.

Furthermore, functional safety standards (see [3] for a draft version) establishing development requirements for automotive safety-critical systems will result in demands for qualified processes and tooling. It is imperative that the automotive control design community properly manage the adoption of such standards for systems, hardware, and software development.

To make the general workflow situation slightly more concrete, we list a representative tool suite for automotive control system development:

- Requirements management: IBM Telelogic DOORS®.
- Plant modeling:
  - Acausal: Modelica Association Modelica®;
  - Causal: MathWorks Simulink®;
  - Empirical: MathWorks Model-Based Calibration Toolbox™;
  - Parameter identification: MathWorks System Identification Toolbox™.

- Control design:
  - Analysis and synthesis: MathWorks Control System Toolbox™;
  - Algorithm specification: MathWorks Simulink®;
  - Automatic code generation: dSPACE TargetLink.
- Verification and validation:
  - Model-in-the-Loop, Rapid Controller Prototyping, Software-in-the-Loop, Processor-in-the-Loop, Hardware-in-the-Loop: dSPACE;
  - Automatic Test Generation: Reactive Systems Reactis®.
- Calibration:
  - Design of experiments: MathWorks Model-Based Calibration Toolbox™;
  - Test automation: A&D Technology ORION;
  - Data acquisition, visualization, and analysis: ETAS INCA.

Opportunities for Research and Education

Given the dynamic nature of the development of control theory, gaps naturally exist between theory, design methodologies, and supporting tools. To reduce these gaps, many opportunities exist for improving tooling, including the following:

- Support for cyberphysical systems and networked embedded control: The rapid advancements in control, computing, communication, and the physical sciences enable system designs that are beyond our ability to analyze and verify. Theoretical system research in the cyberphysical systems and networked embedded control domains is expected to lead to systematic design methodologies and thence to the development of appropriate tools.
- Architectural analysis and design: Control systems must provide cross-cutting qualities such as function, safety, reliability, security, and energy efficiency. Architecture-based annotation and abstraction techniques allow system designers to model, assess, and confirm these qualities for large-scale, component-based systems [4]. The deployment of new instructional modules would accelerate the adoption of this emerging system engineering discipline, and research on new system verification methods that use these annotations and abstractions will greatly enhance system designers' capabilities and capacities.
- Productivity increase: A significant increase in the efficiency of control system development can be expected in different industries by developing and using adequate tools. Some examples are: control-oriented physical plant modeling (all domains); efficient certification/qualification tools (aerospace/automotive); certifiable/qualifiable autocode from control design tools (aerospace/automotive); verification and validation tools allowing automatic test generation at the system level or for requirements coverage (all domains); system identification tools for use in the field (process industries). With the advent of cheap multicore/cluster computing architectures, the use of parallel computational techniques will be an important facilitator of productivity increases.

- Numerical algorithms: System-theoretic algorithms will continue to form the core of computer-aided control systems design (CACSD). Numerical algorithms and accompanying software tools developed for new control domains often represent enabling technologies for the applicability of advanced control techniques. From an educational perspective, more emphasis needs to be placed on numerical algorithms in control engineering curricula.

Additional Workshop Participants

The authors would like to thank the following workshop participants for their contributions to this section: Maryam Khanbaghi (Corning), Alexander Knoll (EADS), Massimo Maroni (Alenia Aermacchi), Johann Bals (DLR), Dragan Obradovic (Siemens), Thomas Bak (Aalborg University), Luigi Glielmo (Università del Sannio in Benevento), and Anuradha Annaswamy (MIT) .

References

[1] Federal Highway Administration, U.S. Department of Transportation. “Systems Engineering for Intelligent Transportation Systems” [Online], Publication No. FHWA-HOP-07-069 EDL-14340, January 2007, Available at http://ops.fhwa.dot.gov/publications/seitsguide/section3.htm.

[2] President’s Council of Advisors on Science and Technology. “Leadership Under Challenge: Information Technology R&D in a Competitive World: An Assessment of the Federal Networking and Information Technology R&D Program” [Online], August 2007. Available at http://www.nitrd.gov/pcast/reports/PCAST-NIT-FINAL.pdf.

[3] International Standards Organization. ISO/DIS 26262-(1 to 10), ISO Standards, ICS, 43 Road vehicles engineering, 43.040.10 [Online], 2009, Under development, Draft available at http://www.iso.org/iso/home.htm.

[4] Society of Automotive Engineers. SAE Architecture Analysis & Design Language (AADL), SAE Standard AS-5506 [Online], January 20, 2009, Available at www.aadl.info.

Cross-Application Perspectives: Application and Market Requirements

Greg Stewart and Tariq Samad

Introduction

Is deep controls expertise sufficient to make an impact on industry and society? Control technology has, after all, had a transformational effect on several application domains and is seen as a crucial enabler for dramatic advances in several others. A common core—the fundamentals of control science and engineering—underlies these past successes and future prospects.

The success of control in practice is contingent on gaining an understanding of the intricacies and idiosyncrasies of specific domains.

Yet the answer to the question is an emphatic no. Success in practice requires considerably more than generic controls expertise. Exploiting the intellectual richness of the field is contingent on gaining a deep understanding of the intricacies and idiosyncrasies of specific domains.

Different application areas differ in ways that are often underappreciated. These points of difference are a mélange of technical and nontechnical factors. A short list includes industry supply chains, hardware and software constraints, engineering organizational structures, sensor and actuator quality and availability, the prevalence of legacy versus new systems, first-principles understanding, educational level of staff, the availability of operational data, and regulatory requirements.

In this section we address the issues and requirements involved in realizing practical, successful industry deployments of new control technology. We contrast selected domains with regard to application and market requirements and discuss aspects of industry applications that are critical to understand in attempting to achieve impact with advanced control.

The Role of Context in Control Engineering Innovation

By definition, *innovation* involves changing an existing situation. As control engineering is fundamentally about the integration of many elements—plant, sensors, actuators, computing, algorithms—it is essential that control engineering researchers fully appreciate all aspects of the environment they wish to improve. This includes understanding the current control design process and performance criteria, then evaluating the changes that are incurred with

As control engineering is fundamentally about the integration of many elements—plant, sensors, actuators, computing platform, algorithms—it is essential that control engineering researchers fully appreciate all aspects of the environment they wish to improve.

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

the proposed innovation. Csikszentmihalyi uses an old Italian expression: “Impara l'arte, e mettila da parte” (learn the craft, and then set it aside). What tasks or expenses will the innovation simplify or eliminate? What new tasks will be introduced as a result of the innovation? Will the innovation bring a net benefit (usually measured in money) to the industrial application?

To answer these questions, one must not consider an innovation in isolation, but instead must evaluate the overall benefit of the new system created by integrating the innovation into the previous system.[1]

When developing an advanced control innovation, one should bear in mind which portions of the current control design process will need to be changed or replaced. To cite a few examples:

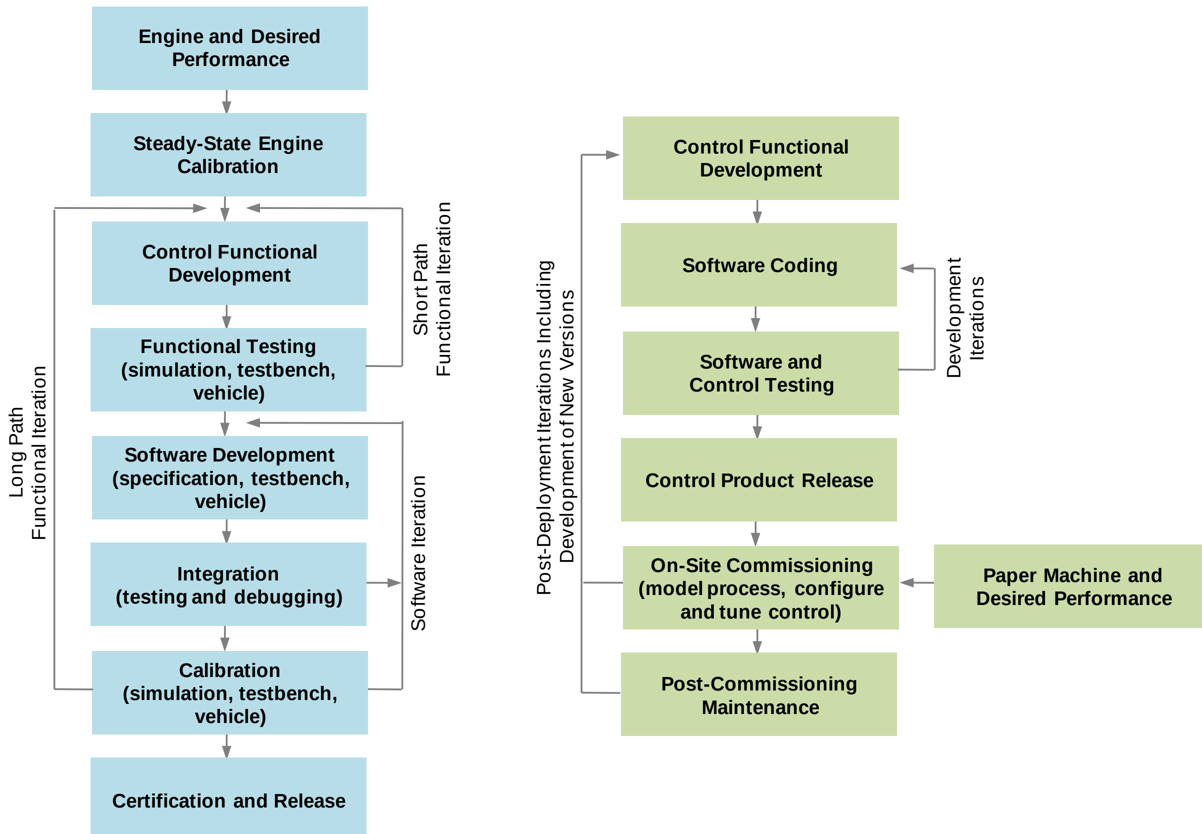
1. Inventing a new PID controller tuning technique may affect only the person responsible for tuning that loop.
2. Introducing a new H_∞ or nonlinear controller would have the impact of 1 above and further require that the real-time control software be changed, a technique for obtaining plant models be in place, and often an industrial-quality (intuitive and error-free) tuning tool that enables nonexperts to tune the advanced controller be available.
3. The introduction of a computationally intensive technique such as model predictive control (standard, not explicit MPC) would have the impact of 1 and 2 and may require an upgrade of the hardware platform to host the algorithm.

To surface some of the key ideas, we can contrast various control design processes. Fig. 1 illustrates the high-level workflow for the development of (1) heavy-duty engine control, and (2) papermaking control. In both cases, many of the familiar design steps are present, but a stark contrast exists in the position of the plant itself. In engine control, a production or prototype engine is available at the start of the control development process, and the control development proceeds with the use of engine measurements and experiments until the tailored and tuned control strategy is released along with the engine production fleet. Conversely, in process control, each plant is often custom designed, and thus each plant is usually very different from every other plant. Advanced control tools are typically developed with the facility to accommodate these plant-to-plant differences by virtue of including model identification software and the ability to straightforwardly configure the controller structure (number of setpoints, actuators, constraints) at the time of commissioning the control. Furthermore, once an engine control is released to market, relatively few opportunities exist for modifying the controller. In process control, the expectation is that during the “post-commissioning maintenance” phase, the models will be re-identified and the control retuned on a frequent, sometimes even weekly, basis.

Generally, an innovation must consider which portion(s) of the current system it will change or replace. The changed system must be “complete” in the sense that the user must be able to perform his/her tasks from beginning to end. Two common examples of incomplete innovations can be cited: (1) a control innovation whose tuning requires Ph.D.-level control expertise where such does not exist at the application, and (2) a complex advanced control algorithm whose memory and processor requirements are too large for the target hardware platform. These classes of innovations cannot be adopted on their own but instead require additional work to become industrially viable.

[1] Here the term *system* is understood in its broad sense to mean an overall situation that may include a design process or accepted method for performing a task or set of tasks.

Furthermore, before performing the work and incurring the expense required to adopt an innovation, an organization will weigh the potential value the innovation is expected to bring. A successful innovation will bring more value than it costs, where these criteria are considered along the usual dimensions that include equipment costs, development time, performance, training, and personnel costs.



(a) Engine Control Development Process (b) Papermaking Control Development Process

Figure 1. Two example industrial control development processes. Many of the activities are included in both situations, but the ordering of the steps is quite different. In particular, the point at which plant-specific information is available is quite different (early for engine control, late for process control) and thus influences the structure of the respective control development processes.

The Elements of Control in Practice

Generally, the impact of an innovation must be evaluated on a case-by-case basis for each potential industry application. However, although each application has its unique characteristics, we find that the control-relevant facets of engineering problems follow some general categories:

- Plant,
- Sensors and actuators,
- Hardware platform,
- Software structure and process,

- Controller tuning (including model identification), and
- Certification.

A common theme in control engineering is evident when a change in one aspect of the application environment enables changes (whether intended or not) in other aspects. For example, both papermaking control and thickness control in steel cold tandem mills initially relied on a proprietary software platform that made it challenging to introduce advanced control. Once open software platforms were introduced, it became much easier to introduce advanced control at the software application level, and both industries now employ robust control and multivariable control in many applications.

Plant

The challenges presented to the control engineer by the plant are generally well known. Some of the leading considerations include the degree of nonlinearity; the complexity of the dynamics; the magnitude of model uncertainty; the constraints on input, output, and states; and the condition number of multivariable plants.

Despite its maturity as a discipline, control engineering is often a technology that is considered only after the plant has been designed. The design of a plant such that it can be effectively controlled is still rare in many applications.

Emerging needs:

- Co-design of plant, sensors, actuators, and control for desired closed-loop performance.
- Control-oriented modeling in terms of physical-based parameters. This would enable a common language between plant designers and control engineers.

Sensors and Actuators

Sensors and actuators are the "handles" by which a control algorithm accesses a plant. Both classes of instrumentation will have requirements in terms of cost, range, bandwidth, and reliability. When considering actuators, it is especially important to understand the role of typical nonlinearities in the control loop—backlash can often be accommodated by detuning the control algorithm, whereas stiction may not.

The performance of sensors is particularly important as feedback control is designed to translate the sensor information into the operation of the plant itself. Sensor accuracy, bias, and cross-sensitivities to their anticipated environment must be considered by the control engineer during the design.

Emerging needs:

- Smart sensors with onboard observers.
- Integration of hardware sensors with inferential sensing for redundancy.
- Networks of wireless sensors.

Computational Platform

The parameters of the intended computational platform are a key consideration when developing a control algorithm. Processor speed, memory, sampling time, architecture, and redundancy all play a role in determining the feasibility of implementation of the algorithm. In the automotive industry, the processor speeds may be in the range of 40 to 56 MHz and 2 to 4 MB of flash memory may be available for control to be executed within milliseconds. On the other hand, modern equipment in the process industries may have a 2.83-GHz processor and 3 GB of memory to execute control actions in seconds or minutes. Very different control approaches may be considered in each case.

The design of control for embedded processors may require the consideration of additional computational aspects such as numerical accuracy in fixed- or floating-point applications.

Emerging needs:

- Hardware-specific algorithm design (for example, designing control algorithms that are robust to fixed-point implementation).
- Control-specific hardware design.

Software Development Process

Since modern control is typically implemented as algorithms in a software application, the importance of the software development process is central. Typically this process follows the phases of proof of concept, application prototyping, testing, software specification, software coding for target, software testing, and finally performance testing. In some industries, the software development process is identified as a key bottleneck in reducing time to market. This is one of the areas where modern control engineering could be expected to make a contribution.

In many applications, a control engineering innovation must accommodate the requirement to develop code to an industry standard. In particular, applications in industries such as aerospace and automotive are finding validation and verification tasks accounting for around half the cost of overall product development.

The issue of interfacing a new control strategy to legacy software is very important. In industrial situations, advanced control is often introduced into an existing software environment, and thus it becomes crucial to define the scope of the control innovation early and evaluate its impacts on the overall software environment. For example, replacing several SISO control loops with a MIMO controller is not always straightforward. In legacy systems, the SISO loops may exist in disparate portions of the overall software environment and can be expected to be connected to legacy diagnostics functionality, which must be maintained in the new control development.

Emerging needs:

- Verifiable control design methods.

Controller Tuning (Including Model Identification)

Given that a control strategy's success or failure can be determined by how its tuning parameters are set, it is surprising how much more attention the research community typically devotes to the development of the core control algorithm, often leaving the setting of the tuning parameters to the end users' discretion. In many industrial situations, the personnel responsible for controller tuning may have little or no advanced control training yet are responsible for delivering acceptable closed-loop performance. This simple fact goes a long way toward explaining the persistence of PID control and simple tuning rules in industrial practice.

Emerging needs:

- Techniques that guarantee closed-loop performance while requiring “industry-realistic” control knowledge.
- Computationally efficient tuning algorithms.
- Systematic and reliable modeling and tuning.

Acknowledgments

This section is the result of a breakout session held at the International Workshop on the Impact of Control: Past, Present, and Future, October 18–20, 2009, in Berchtesgaden, Germany. The authors gratefully acknowledge the contributions of Christian Philippe, Francesco Cuzzola, Paul Houpt, Clas Jacobson, Keith Glover, L.K. Mestha, Boris Lohmann, Manfred Morari, Paolo Coeli, Bob Yeh, Thomas Mannchen, and Ilya Kolmanovsky. Thanks also to Johan Backstrom of Honeywell for providing additional insight into the history of industrial control in the process industries.

Part 2

SUCCESS STORIES IN CONTROL

High-resolution versions of the flyers in this section are available at www.ieeecss.org/main/IoCT-report.

Auto-tuners for PID Controllers

Despite all the progress in advanced control, the PID remains the most popular controller. Any stable system can be controlled with an integrating controller; performance can be increased by adding proportional and derivative action. There is ample evidence that many manually tuned PID controllers do not work well. Automatic tuning has improved performance and simplified use.

PID controllers come in different shapes: as stand-alone components, as elements of distributed control systems, or embedded in instruments and systems.

PID control is used everywhere—in cellular phones, vehicles, process control, heating, ventilation, air conditioning, machine tools, and motor drives. Many PID controllers are found in cars, for example, in engine, cruise, and traction control. PID control is also embedded in instruments like atomic force microscopes and adaptive optics. Because of their widespread use, it is difficult to precisely estimate the number of control loops that are installed each year, but an educated guess is that it is in the billions.

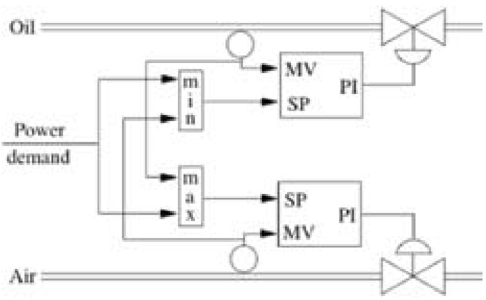


Figure 1

The PID controller is based on very simple ideas. As illustrated in the idealized formula below, the controller output is a combination of three terms:

- The proportional term reacts to current errors.
- Past errors are accounted for by the integral term.
- The derivative term anticipates future errors by linear extrapolation of the error.

$$u_{PID}(t) = k_p e(t) + k_i \int e(t)dt + k_d \frac{d}{dt} e(t)$$

A remarkable property of a controller with integral action is that it gives the correct steady state, if a steady state exists, even for nonlinear processes.

Predicting a noisy signal by linear extrapolation is difficult; it is also difficult to find values of derivative gain k_d that give a robust system. Most PID controllers are in fact used as PI controllers.

A Real PID Controller

PID control is much more than what is captured by the simple idealized formula. To get a functioning controller, one must consider filtering of the measured signal, protection for integral windup, as well as bumpless mode and parameter changes.

Complex System

The PID controller is a simple system. Well-developed architectures exist for building complex systems from the bottom up by combining PID controllers with linear and nonlinear elements such as cascade, mid-range, selector control, and gain scheduling. Figure 1 shows a system for controlling a burner that guarantees there will always be excess air.

Automatic Tuning

Traditionally, PID controllers were tuned manually using simple rules that date back to Ziegler and Nichols in the 1940s. The rules were based on process experiments. The step response method is based on measurement of the open-loop step response. The frequency response method is based on a closed loop experiment where the system is brought to the stability boundary under proportional control. Unfortunately, the traditional rules gave systems with poor performance.

Automatic tuning has increased the use of derivative action. It has even been said: "This controller must have automatic tuning because it uses derivative action."

Automatic tuning can be done in many ways. In rule-based methods that mimic an experienced instrument engineer, features of the closed-loop response are calculated and controller parameters are adjusted based on empirical rules. Other methods are based on estimation of low-order process models, typically first-order dynamics with time delays. The controller parameters are then determined by a variety of control design methods.

Relay auto-tuning is another widely used approach that has proven to be robust and that brings attractive theoretical properties as well.

Contributors: Karl Johan Åström and Tore Hägglund, Lund University, Sweden

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

PID auto-tuners are in widespread use, especially in the process and manufacturing industries. All major instrumentation and control suppliers offer auto-tuning as a feature in their products. Auto-tuning software is also commercially available for PC, SCADA, and DCS platforms and in the simulation programs Simulink and LabView (Figure 2).

Figure 2: PID auto-tuners

Relay Auto-tuning

In relay auto-tuning, the process is first brought to oscillation by replacing the PID controller with a relay function (Figure 3). The controller parameters are then determined from the period and the amplitude of the oscillation. An interesting feature of relay auto-tuning is that it automatically generates signals that are customized for modeling critical aspects of the process. The relay can also be applied to a closed-loop system.

For typical process control applications, the relay auto-tuners can be designed so that tuning can be executed simply by pushing a button; there is no need to set any parameters. The auto-tuner can also be used to generate gain schedules automatically.

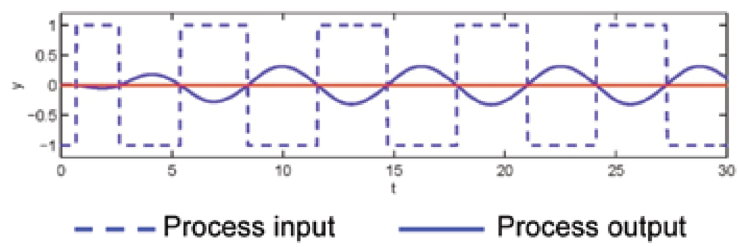
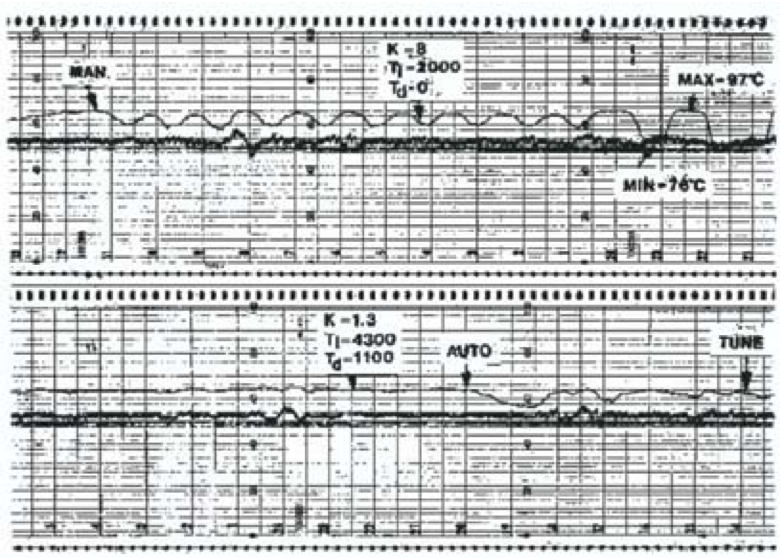


Figure 3

Relay auto-tuning of a temperature control loop on a distillation column.

The data are from a recorder where time runs from right to left. A PI controller produced oscillations as seen in the top plot. The PID controller was switched to manual at 11:15. The oscillation stops but the process drifts. An auto-tuner was installed and tuning was initiated at time 14:00 by pushing the tuning button; no further manual interaction was involved. Tuning is completed at time 20:00 and the controller switches to automatic with good control performance. The auto-tuner introduced derivative action with prediction time 1100s.



For further information: K.J. Åström and T. Hägglund, Advanced PID Control, ISA, Research Triangle Park, NC, 2004; T.L. Blevins, et al., Advanced Control Unleashed: Plant Performance Management for Optimum Benefit, ISA, Research Triangle Park, NC, 2003

Advanced Tension Control in Steel Rolling Mills

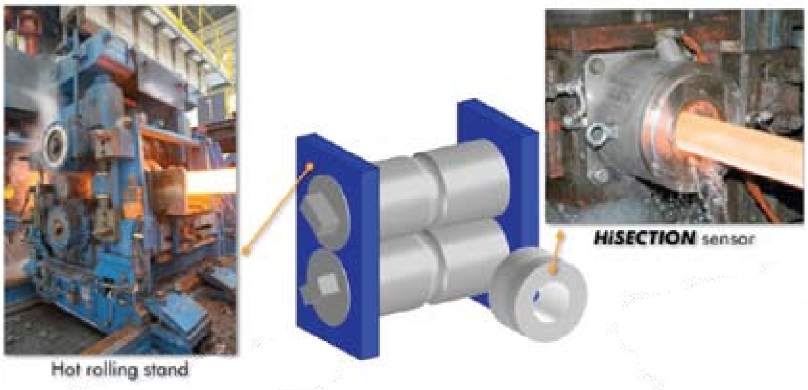
A rolling mill made of a sequence of rolling stands yielding the desired final steel bar section as the bar moves through the mill

Ever-increasing competition in the global steel market has led to the need for significant cost savings in terms of increased production, more stringent tolerances on final product dimensions, and less shop floor utilization. Tension control systems for rolling mills have been a specific target of development because of their cost and impact on product quality.

Traditional tension control systems for steel bars/billets (so-called "long products") involve using loopers between stands to avoid bar tension. Loopers deviate the hot steel trajectory, creating a "buffer" of material between stands to compensate for unanticipated speed fluctuations.

Loopers take valuable space and may cause cobbles; that is, sudden blocking of the hot steel flow leading to loss of production, safety issues, and possibly equipment damage.

HiTension is an innovative and effective architecture for accurately controlling the tension of the steel bar between the stands based on HiSection eddy-current section sensors. Accurate interstand section measurement paves the way to tension control, thus avoiding the need for loopers in rolling mills. Furthermore, improved section tolerances and increased yield are attained.



Schematic view of a rolling stand equipped with HiSection sensor

Enabling Sensing Technologies

- Key to bar/billet tension control is the availability of accurate interstand section measurements.
- Usually only the final downstream section is measured. With eddy-current section sensors, each intermediate section can be measured online.
- Section measurements on each stand allow for identification of an accurate mathematical model of the rolling process.
- This way, the bar/billet section fluctuations can be actively corrected by a model-based multiple-input, multiple-output advanced control system.

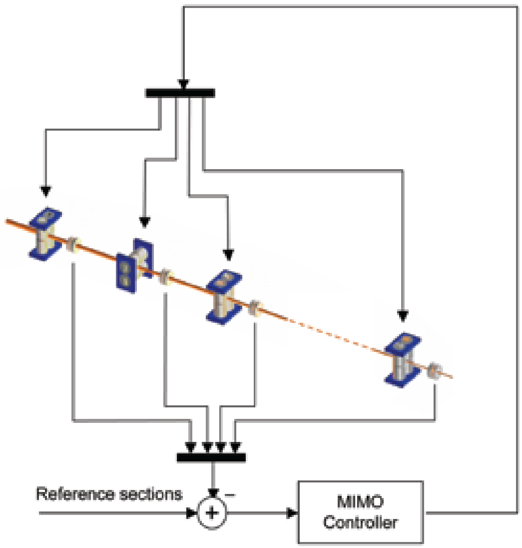
Inventions and Innovations

- First looperless multivariable feedback controller for steel bar/billet tension in hot rolling mills worldwide
- Direct feedback of interstand section measurements with eddy-current section sensors

HiTension and HiSection are products of Danieli Automation S.p.A.

Contributors: Thomas Parisini, Imperial College London, U.K. & University of Trieste, Italy; Lorenzo Ciani and Riccardo M. G. Ferrari, Danieli Automation, Italy

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.



Control Architecture

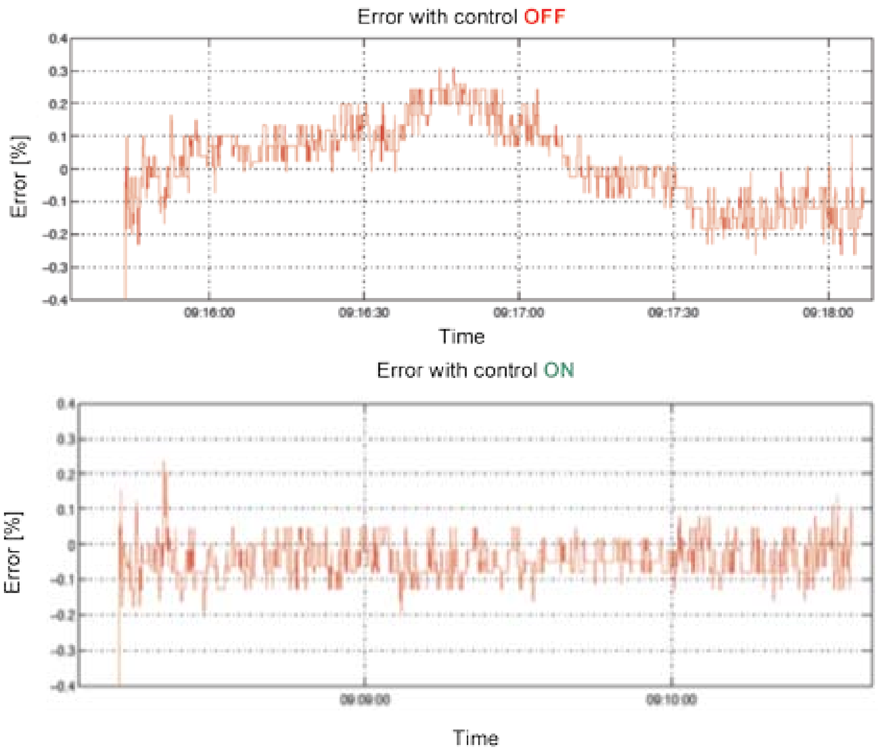
- Each interstand section is measured in real time and compared to reference value. The error variables are fed to the controller that regulates the stand speed.
- Controller parameters are tuned online at the beginning of production based on the identified process model.

Left: Section control architecture for a bar/billet rolling mill. The interstand section is fed by the HiSection sensors to a multivariable controller acting on the stand rotation speeds.

Operational Results

- The HiTension architecture is now running during regular production at a steel plant in northeastern Italy.
- Use of the tension control leads to
  - Tighter tolerances on final section: <0.5% of nominal value;
  - 50% reduction of section error variance on the whole bar/billet.

Right: Comparison of the error on the diameter of a bar, at a fixed point, without (upper) and with (lower) HiTension control. The controller cancels out the low-frequency error component that is present when the controller is turned off.



Benefits

The HiTension bar/billet rolling mill control system has resulted in several very significant benefits:

- Major improvement in rolling quality
- More stable rolling conditions
- Major reduction in cobbles frequency
- Increased productivity through reduction of out-of-tolerance production
- Solution toward looperless rolling mills for long steel products

Advanced Energy Solutions for Power Plants

Fuel costs, energy conversion efficiency, and environmental impacts of fossil-fueled plants have become priorities in both developed and developing countries. Advanced Energy Solutions (AES), a product of Honeywell Process Solutions, is an advanced process control product that significantly improves power plant efficiency and reduces plant emissions.

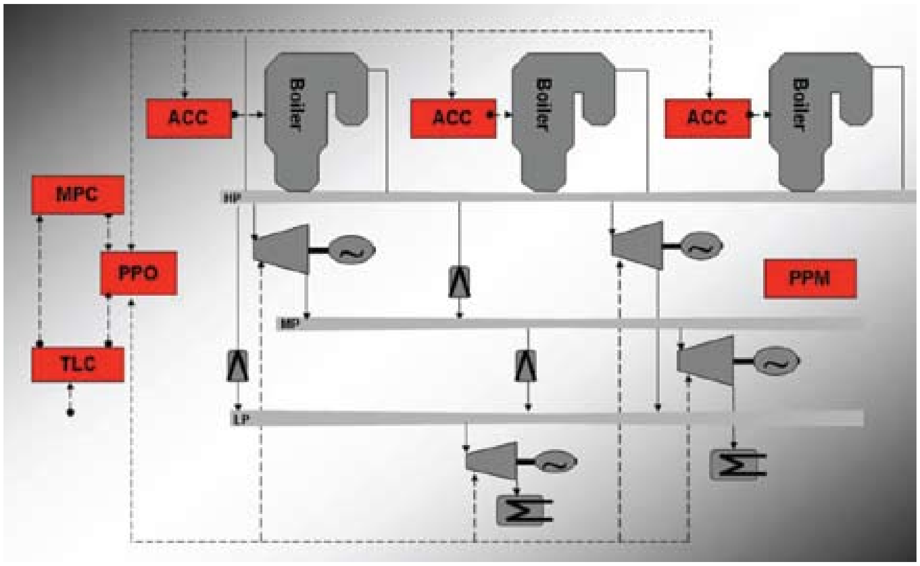
AES provides combustion control in boilers; coordinates multiple boilers, turbines, and heat recovery systems for optimal operation of entire power plants; and provides dynamic balancing of power production to demand.

The AES solution is particularly effective for fossil-fueled power plants and has been applied on several coal-fired boiler applications in Europe, Africa, and Asia.

Successful Applications Worldwide

AES and its component technologies have been implemented in the following plants:

- Co-generation plant Otrokovice, Czech Republic
- ECG Kladno, Czech Republic
- Samsung Fine Chemicals, Korea
- Nam JeJu power plant, Korea
- Sinopec JinShan power plant, China
- SASOL steam plant, Secunda, South Africa



Solution Overview

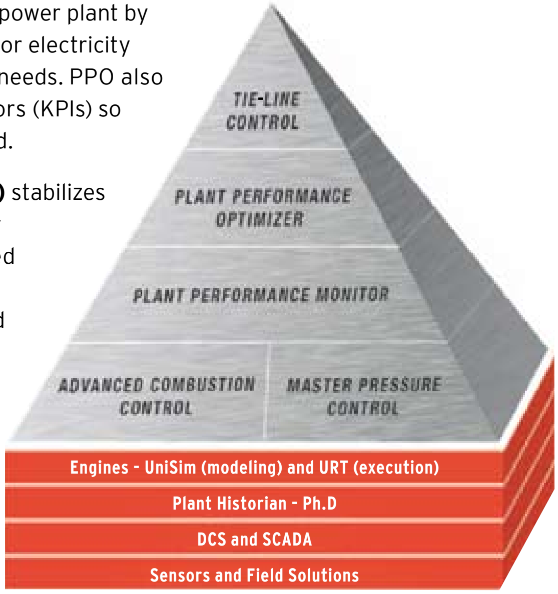
AES is a software-based product that can be implemented as a hierarchical application layer on baseline distributed control systems (DCSs). Several modules are available:

Advanced Combustion Controller (ACC) optimizes air distribution and tightly coordinates control of fuel and air ratio for advanced control of the combustion process.

Plant Performance Optimizer (PPO) increases the efficiency and reliability of the power plant by optimizing the utilization of steam for electricity generation and process or heating needs. PPO also analyzes key performance indicators (KPIs) so business objectives can be achieved.

Master Pressure Controller (MPC) stabilizes steam pressure and prevents boiler and turbine outages using advanced predictive control algorithms. It continuously balances produced and consumed steam and increases asset life by minimizing wear.

Tie-Line Controller (TLC) is a power quota planning and real-time execution toolkit for management of energy supply and demand.



Contributor: Vladimir Havlena, Honeywell, Czech Republic

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

"As the first company in the world to apply advanced control application technology to CFB units, Sinopec significantly enhanced the effectiveness and control performance of the distributed control system at the CFB boiler level and for the entire plant. Even more impressive, all improvements were achieved by implementing software rather than executing a major hardware refurbishment at the plant. We have also to date achieved an estimated $1 million of savings on the supply of energy to our refinery."

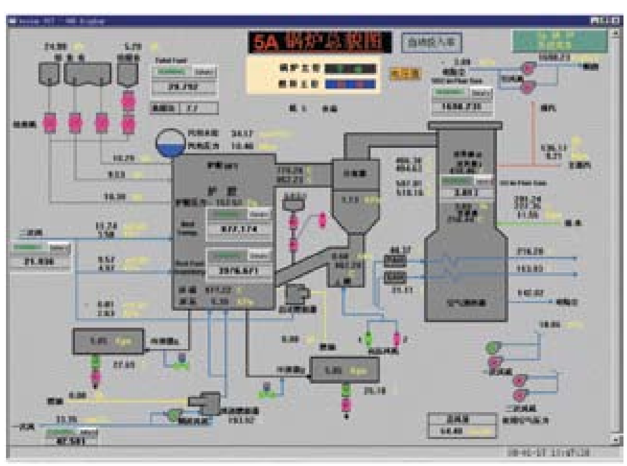
— Zhao Weijie, Chief Engineer, Sinopec Shanghai Petrochemical Company (2008)

Inventions and Innovations

Advanced Energy Solutions incorporates innovative concepts to improve energy efficiency, reduce emissions, and improve the economic operation of industrial and utility fossil-fueled power plants:

- Dynamic coordination of the air-fuel ratio (AFR) in the boiler reduces the AFR variation and enables combustion optimization. An extension of linear model predictive control technology for ratio control was developed.
- Turbulence during combustion results in emissions being highly stochastic. Deterministic optimization methods were unable to provide satisfactory performance. AES's "cautious optimization" strategy takes uncertainty into account.
- One of the key challenges for coal-fired power plants is the variability in the BTU content of the coal. With advanced estimation and inferential sensing technology, leaking air variation and coal quality variation are identified and combustion parameters are optimized online.
- The solution has been extended for circulating fluidized bed (CFB) boilers. CFB boiler dynamics depend significantly on the accumulated char in the bed. An inferential bed fuel inventory (BFI) sensor was developed to estimate the accumulated char level and adapt the model used for predictive control accordingly.
- Another innovation is the plantwide optimization of boilers, turbines, and heat recovery systems to improve the end-to-end efficiency of a power plant.

Most Innovative Power Technology of the Year Award from *Asian Power* magazine, 2008



For the application of AES to Sinopec's Shanghai Petrochemical Company Principal Power Plant in Shanghai, Honeywell received the 2008 Most Innovative Power Technology of the Year Award from Asian Power, the leading publication for energy professionals in Asia.

For further information: http://www.tinyurl.com/HPS-AES

Performance Monitoring for Mineral Processing

The chemical and metallurgical process industries face stiff challenges in the form of increasing energy costs, increasingly stringent environmental regulations, and global competition. Although advanced control is widely recognized as essential to meeting these challenges, implementation is hindered by more complex, larger-scale circuit configurations, the tendency toward plantwide integration, and in some cases an increased lack of trained personnel. In these environments, where process operations are highly automated, algorithms to detect and classify abnormal trends in process measurements are critically important. Advanced algorithms and measurement control systems have been designed and implemented for process performance monitoring and operational performance management, yielding substantial benefits in operating installations.

Monitoring and Control with Computer Vision

Most of the developments were based on the application of computer vision systems in areas where no devices were previously in place to measure key variables in reaction systems:

- Customized computer vision algorithms are used to estimate the proportion of fines in coal feed systems accurately enough to allow online control (Figure 1). Excessive fines in the feed may adversely affect gas flow through the reactor burden, leading to substantial losses associated with suboptimal operation. Previous methods for online analysis of particulate feeds were either nonexistent or based on inefficient automated sieve samples taken from belts.
- Multivariate feature extraction from platinum froth images could be used as a basis for system identification and advanced control of froth flotation systems. With the aid of advanced process models, these features can be used to reliably estimate prevailing process conditions, which is not otherwise feasible (Figure 2).

Figure 1: An image from a computer vision system on a coal conveyor is automatically processed to estimate the proportion of fines in the coal.

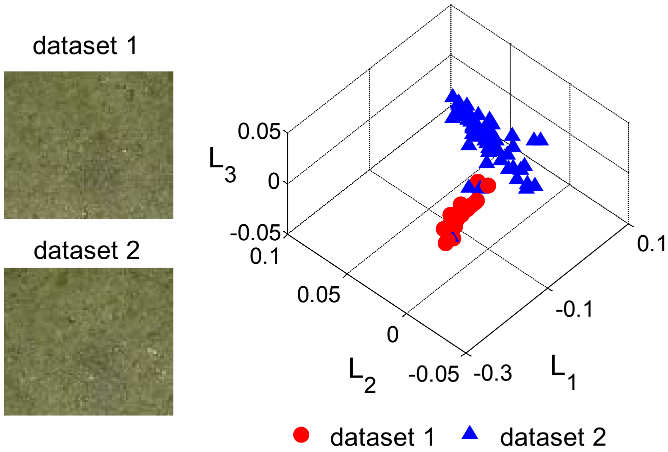


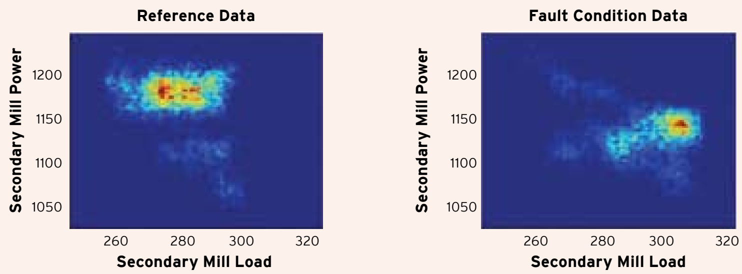
Figure 2: Change in process conditions is detected by an online computer vision system by projection of froth image features to a control chart.

These techniques have been developed by the Anglo Platinum Centre for Process Monitoring and their associates at the University of Stellenbosch in South Africa to enable their industrial partners to realize the competitive advantages afforded by advanced control and monitoring systems.

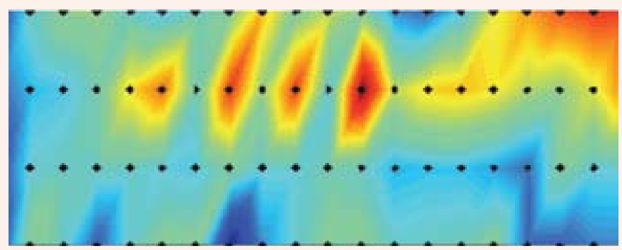
Contributor: Chris Aldrich, University of Stellenbosch, South Africa

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Advanced process monitoring technology can expedite safe operation closer to production limits, eliminate major process upsets, and reduce minor, often unreported, abnormal situations. Capacity increases of 3-8% have been estimated for a range of process industries with monitoring and related technologies (www.asmconsortium.net); this amounts to $200—500 million per year for the South African platinum industry alone.



Monitoring of an autogenous mill on an industrial concentrator plant



Prognostic monitoring of wall temperatures for a metallurgical furnace

Process Fault Detection and Identification

The Centre and its associates have developed state-of-the-art algorithms (Figure 3) for process fault diagnosis that are currently used plantwide by one of the premier mineral processing companies in South Africa. The following benefits have resulted from use of the algorithms:

- The turnaround time associated with attending to plant faults has been reduced from several weeks or even months to less than three days.
- The frequency of large events leading to process circuit and equipment downtime–and associated losses in revenue–has been substantially reduced.
- Plant alarms and the cost of alarm management have been reduced considerably.
- The development of large-scale process monitoring systems has enabled early detection of thermal runaway in metallurgical furnaces.
- The detection and identification of different operational states in autogenous and semi-autogenous mills can at times be realized with an accuracy of up to 80%–providing a basis for better control and potentially large reductions in energy usage by mineral processing plants.

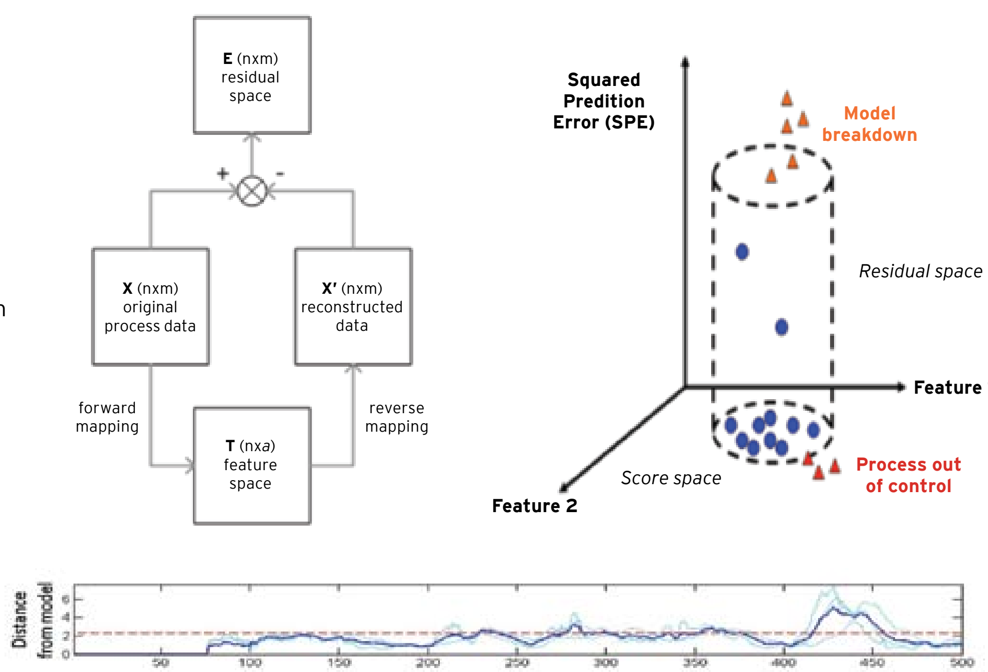


Figure 3: Monitoring the model prediction error in an appropriately constructed feature space enables advance warning of process upsets.

For further information: The Anglo Platinum Centre for Process Monitoring, www.sun.ac.za/cpm

Advanced Control for the Cement Industry

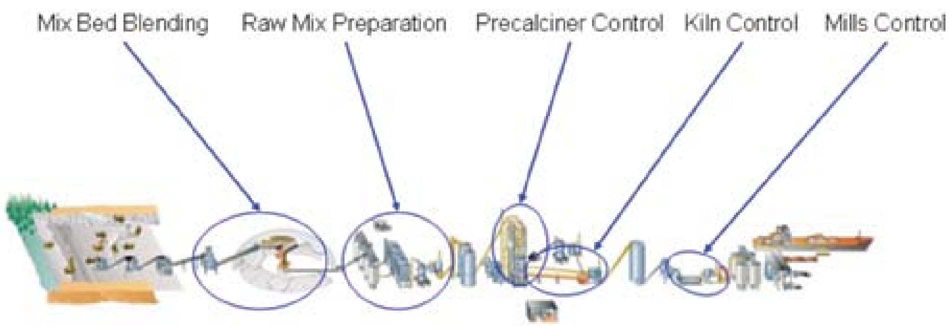
The cement industry of the 21st century is confronted with disparate goals that at first glance seem to conflict. For example, there is enormous pressure to increase profit and margins, while at the same time there is considerable public interest in the sustainable and environmentally friendly use of natural resources. In other words, plant operators find themselves in a situation where they need to react fast and optimally to continuously changing conditions while still meeting various and probably conflicting objectives. Thus, there is a need for tools that bring the plants to their optimal economic performance allowed by the technological, environmental, and contractual constraints. From a technological standpoint, these tools are related to mathematical programming: optimization subject to constraints. The cpmPlus Expert Optimizer (EO) was developed to address these challenges, in particular for cement plants.

Solution Overview

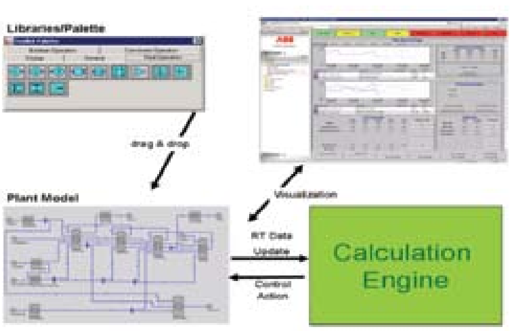
Over many years, a variety of strategies for control and optimization of key industrial processes have been developed and implemented in EO, with particular focus on control and optimization in the cement industry:

- Raw materials blending
- Vertical mills for raw meal grinding
- Calciners and rotary kilns
- Vertical and ball mills for cement grinding

The technology has been deployed in cement plants worldwide. Most installations have been made in blending, kiln, and grinding operations. More than 45 blending systems, 195 rotary kilns, and 90 ball mills have been commissioned by the ABB team in recent years.



cpmPlus Expert Optimizer Applications Scope in the Cement Industry



$$
\begin{aligned}
x(t+1) &= Ax(t)+B_1u(t)+B_2\delta(t)+B_3z(t)\\
y(t) &= Cx(t)+D_1u(t)+D_2\delta(t)+D_3z(t)\\
E_2\delta(t)+E_3z(t) &\le E_1u(t)+E_4x(t)+E_5
\end{aligned}
$$

Energy Efficiency and CO2 Reduction

The cpmPlus Expert Optimizer is a generic platform for development of advanced process control solutions at ABB. It is primarily designed for closed-loop control, optimization, and scheduling of industrial processes, although it can also be used for open-loop decision support applications. When this platform is used, the solution of the problems described above can be attacked with techniques such as model predictive control (MPC) in its mixed logical dynamical (MLD) systems formulation that includes Boolean variables and logical constraints.

For ease of use, the technology has been embedded in a graphical modeling toolkit that allows maximal flexibility during model and cost function design while hiding the mathematical complexity from the user.

Contributor: Eduardo Gallestey, ABB, Switzerland

Global Fuels Award, 2008

In 2008, the cpmPlus Expert Optimizer received the "Global Fuels Award for most innovative technology leading to electrical energy savings." The award was granted by the Global Fuels 2008 conference in London.

Selected Success Stories

Switzerland: Material blending at Untervaz

The Untervaz plant wanted to reduce raw mix quality variability, reduce the associated material costs, and increase the useful lifetime of the quarry. This would also allow the plant to have better process parameters in the kiln, getting closer to clinker quality targets, increasing production, and reducing the risk of process disruptions. In March 2007, ABB extended Untervaz's Expert Optimizer to include ABB's Raw Mix Preparation (RMP) solution. The technologies used are MPC and MLD systems. The benefits achieved by the installation are that raw mix quality variability has been reduced by 20% and kiln process variability has also been reduced. New daily clinker production records have been achieved in the time since RMP has been online.

Germany: Precalciner with alternative fuels at Lägerdorf

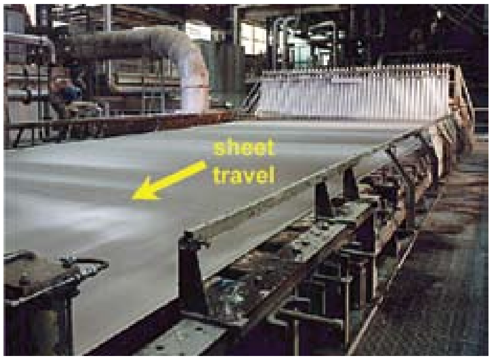
The Lägerdorf plant wanted to increase alternative fuels utilization, get closer to the optimal calcination conditions, and reduce the risk of process disruption. In August 2006, ABB successfully installed Expert Optimizer, encompassing a Precalciner Temperature (PCT) control solution, on the calciner at Lägerdorf. The technologies used are MPC and MLD. The installation achieved a dramatic increase in the use of alternative fuels. Furthermore, it was possible to reduce temperature variability, bring the precalciner average temperature toward optimal values, and reduce the risk of cyclone blockages.

Italy: Cement grinding at Guidonia

Buzzi Unicem wanted a solution for its Guidonia plant that would increase the productivity of its cement grinding system, consisting of three mills. ABB installed Expert Optimizer on the mills at the Guidonia plant between December 2006 and January 2007. The EO team overcame the challenges at the Guidonia plant by applying the MPC approach together with a tailor-made parameter adaptation and process supervision procedure. The benefits are better grinding process parameters and operation closer to process constraints. The specific energy consumption was reduced by as much as 5%.

For further information: www.abb.com/cpm

Cross-Direction Control of Paper Machines



Cross-Direction versus Machine Direction

Traditionally, paper machine control was limited to the "machine direction"—reducing the variation in the paper linearly along the direction of travel, without coordinating control "across" the sheet. Cross-direction (CD) control is a considerably more challenging problem—its solution revolutionized paper making!

A paper machine is a technological marvel! Think of it as a 100-meter-long, 10-meter-wide wire screen that can move at faster than 100 km/hour. At one end of the machine, pulp stock is extruded onto the wire screen; this stock is composed of about 99.5% water and 0.5% fibers. Over a 100-meter-long machine, the paper sheet travels a path that may cover more than 200 meters.

Ultimately paper is produced; the moisture content at the dry end is about 5-8% water and 92-95% wood fibers. With new "cross-direction" control technology, the paper can be produced to a thickness ("caliper") uniformity of within a few microns over entire production reels, each of which can contain a 40 km length of the paper sheet.

Scanning sensors measure sheet properties such as weight, moisture, and caliper in a zig-zag path on the moving paper sheet, and up to 300 individually controlled actuators in each of up to seven actuator beams are used for pulp stock metering, steam heat drying, water shower re-wetting, and induction heating.

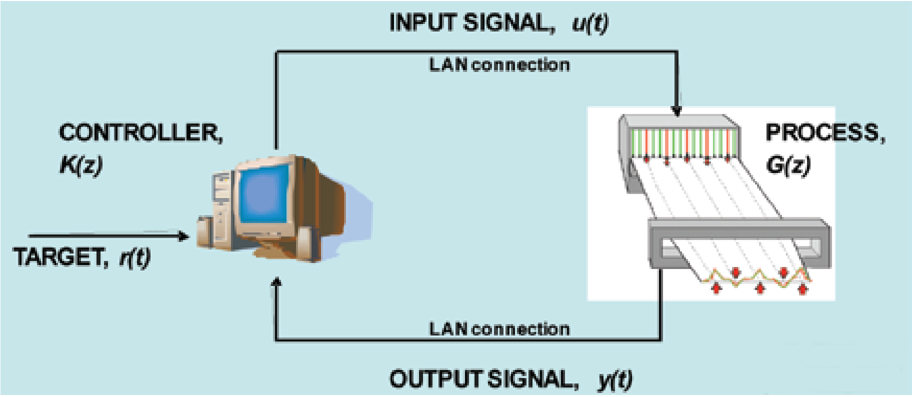
Problem Characteristics

The response of the paper to the cross-directional actuators has both a dynamic component and a spatially distributed component.

Analysis of the response in frequency domains reveals that the dynamic response is small at fast temporal frequencies and the spatial response is small at short spatial wavelengths (corresponding to the existence of small singular values).

A common closed-loop instability occurs at slow temporal frequencies and short spatial wavelengths—a result of combining aggressive control action (an integrator!) with low plant gain.

Honeywell's cross-direction paper machine control products, IntelliMap and PerformanceCD, are deployed on more than 300 paper machines. Benefits include up to 50% higher performance and about 80% reduction in control tuning time.



A long-standing industrial control problem solved by robust control design.

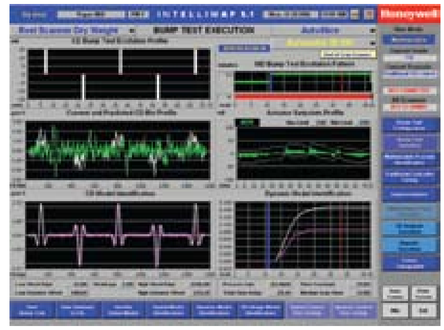
Contributor: Gregory Stewart, Honeywell, Canada

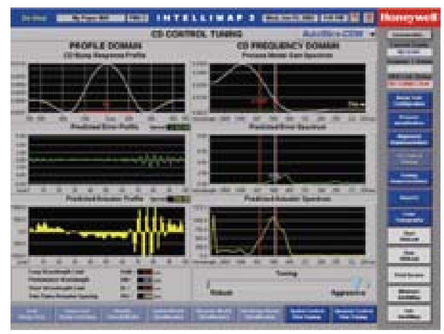
Operators and engineers in paper manufacturing plants are not experts in control theory. Providing tools that can be used by paper industry personnel has been instrumental to the success of this CD control innovation.

Design and Tuning Tools

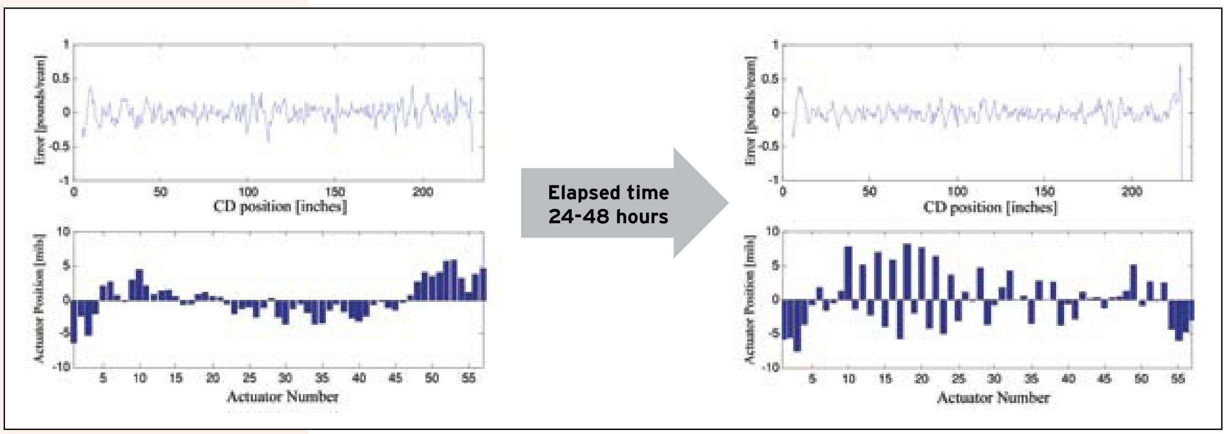
The tools provided for the success of this CD control innovation include:

- A test and identification tool that automatically makes changes to the actuators, collects the resulting sensor data, and analyzes the data to develop a large-scale, spatially distributed model.
- An automated controller tuning tool that takes the developed model and determines optimal tuning parameters for the cross-directional controller. The controller is designed using a technique called two-dimensional robust loop shaping.





The Previous State of the Art: *Prior to the introduction of advanced control, these systems would often develop a slow closed-loop instability, sometimes taking as long as 24 to 48 hours to appear after controller tuning.*



Awards

- IEEE Control Systems Technology Award from the IEEE Control Systems Society for "innovative application of modern identification and control methods to the papermaking process" awarded to D. Gorinevsky and G. Stewart in 2001.
- *IEEE Transactions on Control Systems Technology* Outstanding Paper Award given to G. Stewart, D. Gorinevsky and G. Dumont, "Feedback Controller Design for a Spatially Distributed System: The Paper Machine Problem," vol. 11, no. 5, September 2003.

For further information: Go to www.honeywell.com and search "paper quality control system"

Ethylene Plantwide Control and Optimization

Ethylene is the largest-volume industrial bulk commodity in the world. The majority of ethylene is used in the production of ethylene oxide, ethylene dichloride, ethylbenzene, and a variety of homo- and co-polymers (plastics ranging from plastic food wrap to impact-absorbing dashboards inside cars).

Ethylene plants are complex, large-scale, flexible factories that can process a wide variety of feedstocks, ranging from gases (such as ethane, propane, and LPG), to naphthas, to distillates and gas oils. Main products are polymer-grade ethylene and propylene. Operational objectives include yield improvement, production maximization, and energy intensity reduction.

Honeywell's advanced control and optimization technology has been applied to ethylene plants worldwide with substantial economic benefits, including millions of dollars from increased production annually and additional benefits from energy savings.

Process and Operating Characteristics

Universal:

- No product blending
- Stringent product quality requirements
- Slow dynamics from gate to gate
- Gradual furnace and converter coking
- Frequent furnace decoking and switching
- Converter decoking

Site-specific:

- Feed quality variations
- Product demand changes
- Sensitivity to ambient conditions
- Periodic switching (for example, dryers)

Operating Flexibilities and Solution Goals

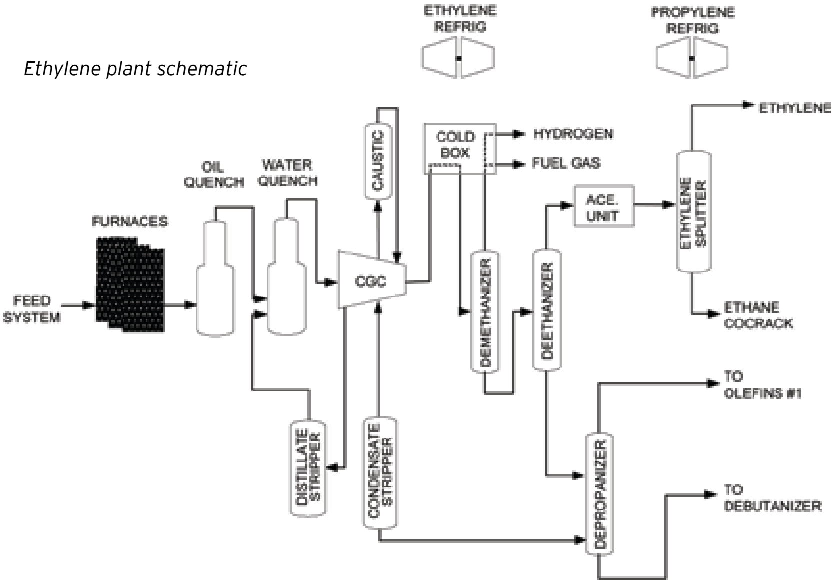
The main operating degrees of freedom for ethylene plantwide control and optimization include feed selection, furnace feed rates, cracking severity, dilution steam, cracked gas compressor and refrigeration compressor suction pressures, typical column variables (reflux, reboiler, and pressure), and converter temperature and H2 ratio. Advanced control and optimization goals include:

- Stabilizing operation
- Minimizing product quality giveaway
- Maximizing selectivity and yield
- Minimizing converter over-hydrogenation
- Minimizing ethylene loss to methane and ethane recycle

Combined Control-Optimization Solution

Unlike optimization approaches based on steady-state models, the solution featured here relies on dynamic models of model-predictive controllers (MPCs). There is no need to wait for the plant to reach steady state, and economic optimization is augmented into a standard MPC control formulation, known as range control. Nonlinearity of the plant is accounted for with successive linear dynamic models. The use of nonlinear dynamic models is in development and has been demonstrated experimentally.

Ethylene plant schematic

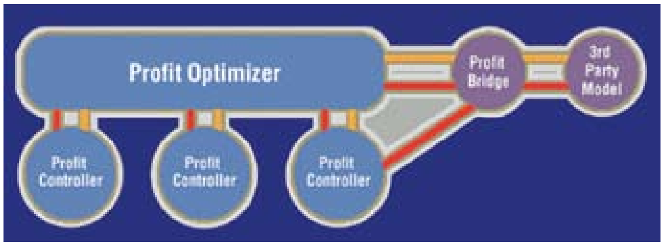


Contributors: Joseph Lu and Ravi Nath, Honeywell, USA

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Overview of an Ethylene Plantwide Control and Optimization Project

A typical advanced control and optimization solution for ethylene plants comprises a global optimizer (Profit Optimizer) that coordinates 15 to 30 model predictive controllers (Profit Controllers) for separation and quench towers, converters, and a fuel-gas system. MPC controllers execute every 30 to 60 seconds and the global optimizer every minute. Technip's SPYRO nonlinear model is used to update the furnace yield gains every 3 to 5 minutes.

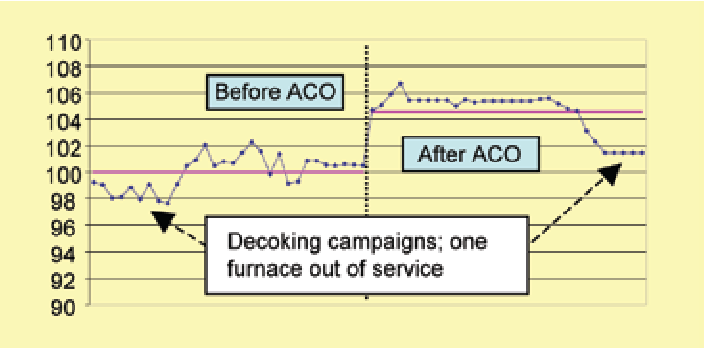


Honeywell's Profit Suite optimization and control products also include Profit Bridge for integrating third-party, nonlinear steady-state models; Profit Stepper for model identification (including closed-loop identification that allows models to be developed while the plant operates); and an advanced single-loop controller, Profit Loop.

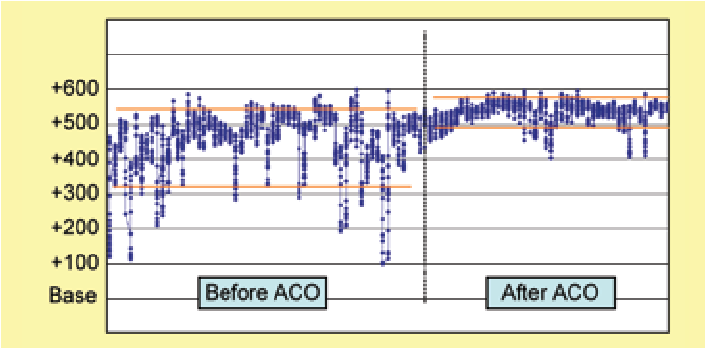
The technology is based heavily on dynamic models. Steady-state nonlinear models are used selectively for calculating critical gains. All Profit Controllers operate off linear dynamical models (usually developed with the Profit Stepper application). The base Profit Optimizer model is automatically aggregated from the Profit Controller models; Bridge Models and source/clone structures are added to define interactions among the controllers.

Once a plant has been commissioned, validated, and put into operation (a process that takes 9 to 12 months), little maintenance is typically required. Clients either dedicate a half-time control engineer to monitor and perform minor services or depend on quarterly visits by Honeywell staff. (In contrast, traditional real-time optimization solutions typically require a full-time modeling person and a half-time control engineer to maintain the solution.)

Normalized olefin production rate before and after advanced control and optimization (ACO):
Production is increased and product variability reduced.



Ethane content (ppm) in product ethylene:
The plantwide optimization and control solution gives tighter control and also reduces quality giveaway; the product purity specification is met without incurring the expense of further reducing ethane content.



Awards:

American Automatic Control Council, 2010, Control Engineering Practice Award for "Innovation in advanced control and optimization with sustained impact on the process industries."

Control Engineering Magazine, 1999, Editor's Choice Award for RMPCT (now Profit Controller).

Broad Process Industry Impact!

The sequentially linear, dynamical model predictive control and optimization solution showcased here has been applied since 1995 to more than 10 industries, such as refining, petrochemical, oil and gas, coal gasification, LNG and LPG, pulp and paper, polymer, and aluminum.

Production increases valued at $1.5–$3 million annually are typical for ethylene plants. Energy savings are an additional and significant benefit.

For further information: Go to www.honeywell.com and search "Advanced Control and Optimization."

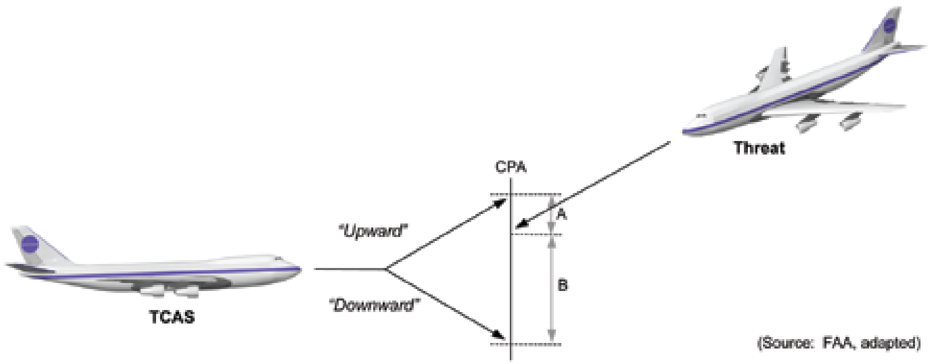
Automated Collision Avoidance Systems

CTAS provides automation tools for air traffic controllers to use in planning and controlling arrival traffic. It includes methods for achieving acceptable aircraft sequencing and separation as well as improved airport capacity. This photo shows CTAS' Traffic Management Advisor (TMA) in operation in the Denver airport facility. FAA operates TMA at all of its centers and says it has fundamentally changed National Airspace System (NAS) operations.

Control theory to date has achieved tremendous success in the analysis and synthesis of automatic guidance systems in aerospace. Aircraft and spacecraft autopilots are elegant schemes that employ a hierarchical structure. For example, pilots can fly today's commercial jets by programming into their onboard flight management system a set of waypoints describing the desired flight path over time. These waypoints are automatically translated into a sequence of guidance commands for the aircraft, and these in turn into the actuator commands for the aircraft throttle and control surfaces. In recent years, this success of guiding single aircraft has been extended to the relative control of two or more vehicles in the design of airborne collision avoidance systems and tactical separation assurance tools in air traffic control (ATC). Although many of these systems have been developed as prototypes within the research community, several have been tested and are now operational in NASA's Center TRACON Automation System (CTAS).

Collision avoidance and separation assurance tools can be classified into three groups according to the time horizons over which they operate. The Traffic Alert and Collision Avoidance System (TCAS) operates over a time horizon of less than a minute and is called an immediate collision avoidance scheme. New automated methods for both ground-based and airborne collision avoidance, known as midterm collision avoidance schemes, are being designed for a time horizon of a few minutes. Tactical air traffic control schemes provide separation to aircraft and generally operate over a longer time horizon of about 30 minutes. These are denoted as separation assurance schemes. Separation assurance is an air traffic control responsibility for aircraft to maintain a separation of 5 nautical miles (lateral) and 1000 feet (vertical).

In the TCAS protocol, when an intruder aircraft is declared to be a "Threat" to the "TCAS" aircraft, the Resolution Advisory (direction and rate) is designed, based on range and altitude tracks, to give the most separation at the closest point of approach (CPA).



Today, TCAS is installed on all commercial aircraft with at least 30 passenger seats operating in the U.S. It receives and displays bearing and relative altitude information about all other aircraft within a 40-mile radius and provides alerts and its Resolution Advisory with respect to the aircraft that poses the greatest potential threat.

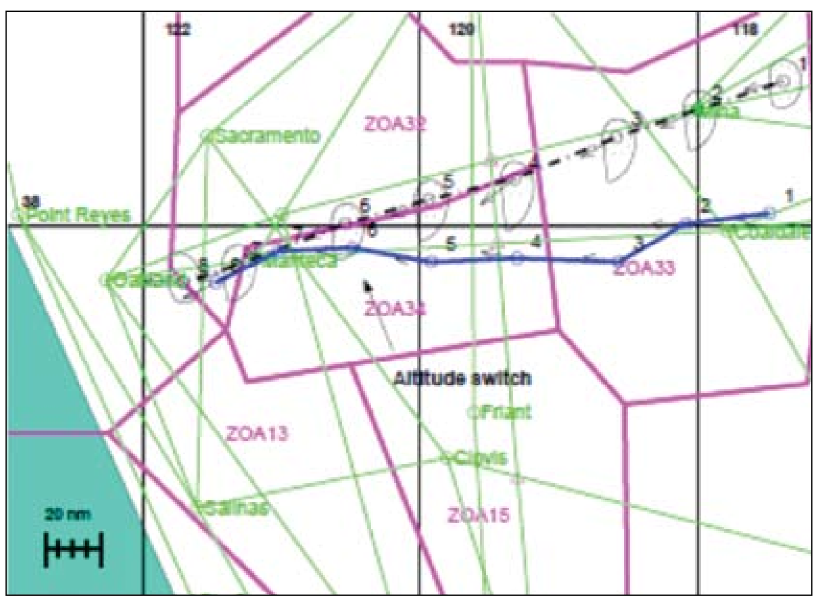
Contributors: Claire Tomlin, University of California, Berkeley, and Heinz Erzberger, NASA, USA

For the ground-based system, automated tools such as the midterm collision avoidance scheme called Tactical Separation Assisted Flight Environment (TSAFE) are being developed to provide conflict alert and resolution advisories to the air traffic controller or to the pilot directly via data link. Separation assurance functionality, operating over a time horizon of about 30 minutes, consists of a tactical ATC function that handles the coordination of traffic in a local area (for example, the maneuvers for efficiently spacing and sequencing traffic to a metering fix or to a runway). Fast optimization schemes for routing aircraft in the presence of metering and capacity constraints have also been developed with the goals of providing advisories or automated functionality to tactical-level control.

The horizontal resolution method used in TSAFE generates a set of maneuvers to ensure achieving the specified minimum separation between aircraft. These maneuvers consist of a turn to a specified heading followed by straight-line flight. TSAFE also uses vertical maneuvers when required. The solution is generated analytically and is thus computationally efficient and suitable for real-time implementations. The resolutions could be implemented by the air traffic controller or could be uplinked directly to the aircraft using existing data link technologies.

The figure on the left shows a highlight of a predicted conflict between two aircraft, with the red lines indicating the flight paths that lead to a predicted loss of separation in 9 minutes from the current positions of the two aircraft in conflict. The resolution trajectory is generated automatically and is shown in yellow for one of the two conflict aircraft. The small white, blue and green diamonds show the locations of neighboring traffic that was accounted for in the generation of the resolution trajectory.

Methods for analyzing the safety of collision avoidance systems have also been developed. One such method uses reachable set technology to determine the unsafe configurations of one aircraft with respect to another. As an aircraft approaches the boundary of the unsafe region, corrective action must be applied. This control action is computed automatically as part of the reachable set calculation. The figure on the left shows two aircraft arriving at Oakland airport. At the position labeled 6, both aircraft are inside the reachable set, indicating an unsafe configuration (loss of separation in 3 minutes). In the actual scenario, the controller performed an altitude change to resolve the conflict.



For further information: http://www.aviationsystemsdivision.arc.nasa.gov/index.shtml; Introduction to TCAS II, Version 7, U.S. Dept. of Transportation, FAA, November 2000; Erzberger and Heere, DOI:10.1234/09544100JAERO546; Mitchell, Bayen, and Tomlin, DOI:10.1109/TAC.2005.851439.

Digital Fly-by-Wire Technology

Digital fly-by-wire (DFBW) is one of many success stories where technology developed under the U.S. space program has proven beneficial in other areas. Based in part on a recommendation from Neil Armstrong, who was directly familiar with the Apollo Guidance Computer through his historic lunar landing, NASA's Dryden Flight Research Center chose to work with Draper Laboratory to adapt the concept for aircraft, beginning with experimentation on a U.S. Navy F-8 Crusader in 1972.

From top to bottom: Space Shuttle, Airbus A320, B-2 Stealth Bomber, Boeing 777, Dassault Falcon 7X, and Joint Strike Fighter X35. Sources: NASA, Airbus, Boeing, and Dassault

Draper developed DFBW as an extension of its work on the Apollo Guidance Computer. The concept uses a highly reliable computer and electronic flight control system, rather than mechanical or hydraulic-based systems, to stabilize and maneuver a vehicle. The computer is able to execute far more frequent adjustments than a human pilot, thus helping maintain stability while offering increased maneuverability.

The 15-year DFBW technology research program also demonstrated adaptive control laws, sensor analytical redundancy techniques, and new methods for flight testing digital systems remotely.

Real-World Applications

The F-8 digital fly-by-wire program served as the springboard for DFBW technology to be used in both military and civilian aircraft. Today, commercial launch service providers and satellite manufacturers also routinely use the technology in their vehicles and spacecraft. Below is a partial list of aircraft and spacecraft with DFBW technology:

- Space Shuttle
- Launchers: Ariane, Vega, Titan, Delta, Proton
- Airbus A320 (first airliner with DFBW controls)
- Boeing 777 and 787
- Jet fighters: F-18/22, Dassault Rafale, Eurofighter, Joint Strike Fighter X35
- Stealth Bomber: F-117, B-2
- Dassault Falcon 7X (first business jet with DFBW controls)
- Rotorcraft: V-22 Osprey, RAH-66 Comanche, AH-64 Apache, NH-90, Sikorsky S-92
- Several unmanned aerial vehicles (UAVs)

Apollo computer interface box used in the F-8C digital fly-by-wire program. Source: NASA

NASA used an F-8C for its digital fly-by-wire program, the first DFBW aircraft to operate without a mechanical backup system. This photo shows the Apollo hardware jammed into the F-8C. The computer is partially visible in the avionics bay. Source: NASA

Contributor: Christian Philippe, European Space Agency, The Netherlands

Space Technology Hall of Fame 2010

NASA's Dryden Flight Research Center, Draper Laboratory, The Boeing Company, and Airbus were inducted into the Space Technology Hall of Fame in 2010 for the development of digital fly-by-wire technology that makes modern aircraft easier and safer to operate.

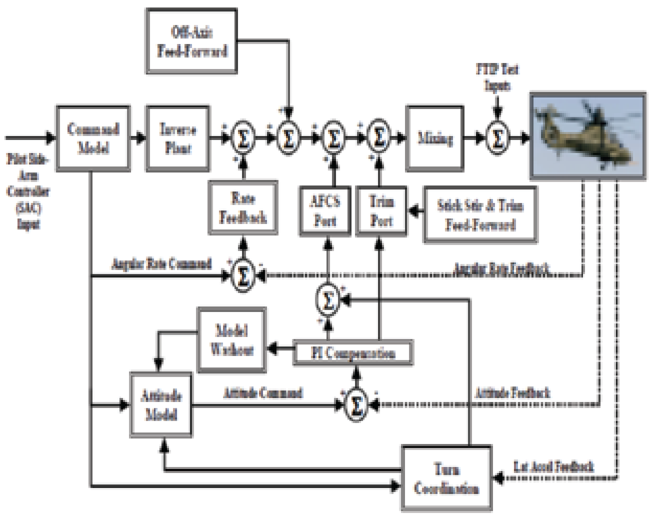
Left: RAH-66 Comanche demonstrating full flight envelope capability. Right: X-31 demonstrating high angle of attack maneuver. Sources: Boeing/Sikorsky and NASA

Major benefits of DFBW flight controls include:

- Overall cost reduction
- Overall airframe weight reduction
- Increased safety and reliability
- Fuel efficiency
- Reduced CO_2/NOx emissions
- Improved flying (or handling) qualities
- Improved passenger comfort
- Reduced pilot workload
- Ease of assembly and maintenance
- Improved survivability
- Improved mission performance

The following features, enabled by DFBW, are currently implemented onboard fighter aircraft:

- Reconfigurable flight control system allowing mission continuation or safe recovery following system failures or battle damage
- Flight envelope protection such as bank angle protection, turn compensation, stall and overspeed protection, pitch control and stability augmentation, and thrust asymmetry compensation
- Online system identification for verification of the aerodynamic effects on aircraft flexible modes



RAH-66 Comanche multimode control law architecture

For further information: J.E. Tomayko, Computers Take Flight: A History of NASA's Pioneering Digital Fly-By-Wire Project, The NASA History Series, NASA SP-4224, National Aeronautics and Space Administration, 2000; NASA Dryden Technology Facts – Digital Fly By Wire, http://mynasa1.nasa.gov/centers/dryden/about/Organizations/Technology/Facts/TF-2001-02-DFRC.html.

Nonlinear Multivariable Flight Control

F-35

Under contracts from NASA and the Air Force Research Laboratory, Honeywell and Lockheed Martin developed and documented a novel approach to flight control design. This multivariable flight control methodology based on nonlinear dynamic inversion was applied to the experimental X-35 military aircraft and the X-38 Crew Return Vehicle program and is now in production for the Lockheed Martin F-35 aircraft.

This work contributed to the advancement of advanced control technology and resulted in the control law being recognized as a design option for flight control development.

Key Innovations

The nonlinear dynamic inversion control approach is a systematic generalized approach for flight control. Using general aircraft nonlinear equations of motion and onboard aerodynamic, mass properties, and engine models specific to the vehicle, a relationship between control effectors and desired aircraft motion is formulated. A control combination is designed that provides a predictable response to a commanded trajectory. Control loops shape the response as desired and provide robustness to modeling errors. Once the control law is designed, it can be used on a similar class of vehicle with only an update to the vehicle-specific onboard models.

Specific innovations include:

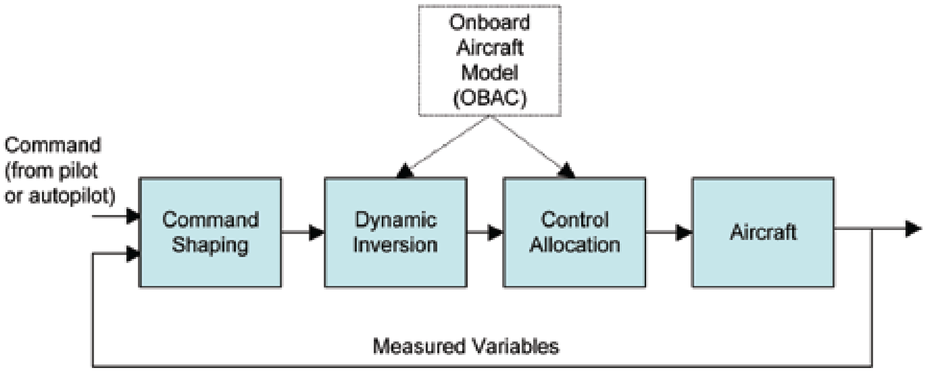
- The dynamic inversion control law
- A control allocation procedure
- An onboard aircraft model (OBAC)

Previous Practice

Nonlinear dynamic inversion is the first systematic approach to nonlinear flight control. Prior to this development, the control law was typically designed from a set of linear plant models and implemented with a gain-scheduled linear controller. The performance capabilities of the aircraft were not fully realized, and the manually intensive development process was time consuming.

Onboard Aircraft Model (OBAC)

The NASA/Honeywell/Lockheed Martin flight control approach includes the first use of an aircraft model in the control law. This model is used to derive coefficients for dynamic inversion and control allocation computations. Changes in vehicle structure during design often only require changing the OBAC model for the controller.



Contributors: John Bosworth, NASA, and Dale Enns, Honeywell, USA

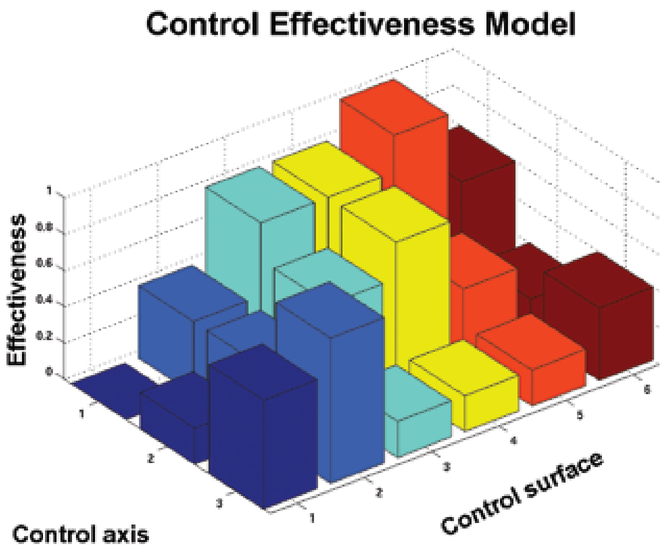
From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

F-18

X-38

Nonlinear Dynamic Inversion

Dynamic inversion "inverts" the aircraft model to identify what roll, pitch, and yaw moments will give the desired aircraft trajectory. Sophisticated numerical algorithms are used to ensure rapid computation and to deal with actuation rate and deflection limits. The inversion is done in real time during flight.



Control Allocation

Modern aircraft have redundant actuation capabilities: the same airframe response can be achieved with different combinations of actuators. The problem of determining which actuators to use, and to what extent, at a given instant is referred to as control allocation. The nonlinear dynamic inversion methodology computes an optimal control allocation taking into account saturation constraints on actuators. The actuators of interest for the applications developed are the control surfaces (ailerons, elevators, rudder) and thrust vectoring (directing the engine thrust). A control effectiveness model is used in the computation (see figure above).

Program History

The work described in this success story began in the mid-1980s as a theoretical development for a high angle-of-attack aircraft. A Honeywell nonlinear dynamic inversion design was selected as the controller for the F-18 research vehicle and implemented in a full hardware-in-the-loop piloted simulation.

In the late 1990s, NASA began the X-38 Space Station Crew Return Vehicle program. The nonlinear dynamic inversion controller was proposed for this program and allowed control updates as the vehicle structural design changed by simply updating the OBAC model.

These and other foundational projects led to the collaboration between Honeywell and Lockheed Martin and implementation of nonlinear dynamic inversion on the X-35 prototype and eventually the production F-35 vehicle, the latest state-of-the-art military aircraft. The controller has provided consistent, predictable control through the transition from conventional flight to hover and has also enabled a 4X to 8X reduction in nonrecurring engineering development cost.

For further information: Honeywell and Lockheed Martin, Multivariable Control Design Guidelines, Final Report, WL-TR-96-3099, Wright Patterson AFB, OH, U.S.A., 1996.

Robust Adaptive Control for the Joint Direct Attack Munition

Control theory has been the enabling technology in achieving man's dominance over flight. Early experimental aircraft were difficult to control, had limited flight envelopes and flight times, and the pilots had to exercise control over the aircraft's trajectory using mechanical systems. In these early aircraft, the pilots had to adapt to changing environmental conditions and/or failures of any aircraft components. Over several decades, propulsion systems matured, our understanding of flight dynamics and aerodynamics grew, and computers and digital fly-by-wire systems were developed, all of which have helped bring automation to flight control.

With recent advances in control theory, particularly in the area of robust and adaptive control, fully automatic flight is now possible even for high-performance air systems. Among the first application successes of this new technology has been its technical transition to guided munitions, in particular, the Joint Direct Attack Munition (JDAM) system.

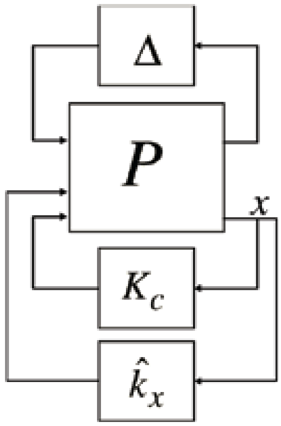
Robust Adaptive Control

Two techniques have been developed by control engineers and scientists to accommodate uncertainty in our knowledge of the system we are trying to control. The two techniques are complementary and have been combined to create robust adaptive controllers.

Robust control: Based on a mathematical model of the uncertainty, a formal design procedure is used to develop closed-loop controllers that will provide optimized performance and ensure stability over the range of uncertainty.

Adaptive control: Instead of developing a fixed controller over a space of model uncertainty, adaptive control adjusts the controller online based on detections of plant deviations from a reference model. Adaptive control augments and further extends the performance and robustness of the flight control system.

Shown at right is a control engineer's block-diagram representation of robust adaptive control. The nominal plant model P of the system under control (such as a missile) is subject to uncertainties Δ. The baseline flight controller K_c, designed using robust control techniques, is augmented with an adaptive controller. The state vector x is the input to both the baseline and adaptive controllers. The combination provides robust stability and performance over a substantially enhanced space of modeling uncertainties and can accommodate changes in the system under control.



JDAM

The Joint Direct Attack Munition is a guidance kit that converts unguided bombs into all-weather "smart" munitions. JDAM-equipped bombs are guided by an integrated inertial guidance system coupled to a Global Positioning System (GPS) receiver, giving them a published range of up to 15 nautical miles (28 km). The guidance system was developed jointly by the United States Air Force and the United States Navy. The JDAM was meant to improve upon laser-guided bomb and imaging infrared technology, which can be hindered by potentially bad ground and weather conditions.

Contributors: Kevin A. Wise and Eugene Lavretsky, Boeing, USA

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Improving Weapon Control and Effectiveness

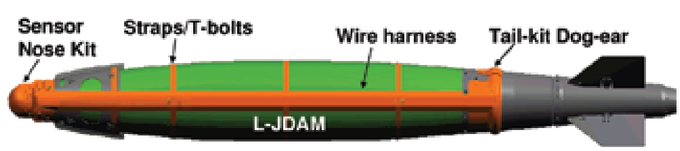
The U.S. Air Force Office of Scientific Research (AFOSR) sponsored researchers at The Boeing Company to develop and transition new robust adaptive control algorithms for application in the Joint Direct Attack Munition. The first transition has been to the 500-lb MK-82 JDAM.

Affordability and weapon accuracy (including collateral damage minimization) are among the primary objectives for the JDAM. The new robust adaptive control algorithms provide accurate control of the weapon, accommodating warhead aerodynamic uncertainties and off-nominal mass properties. Without control modifications, these uncertainties can significantly degrade weapon accuracy.

Photo sequence showing JDAM test against mobile ground targets

Other Applications

Other air systems are also prime candidates for robust adaptive control technology. Recently, another variation of the JDAM system has been developed and transitioned into production: a new dual-mode laser-guided JDAM system (L-JDAM) for detecting and prosecuting laser-designated targets (moving or fixed). For the L-JDAM development, the adaptive controller augments the baseline flight control system and only engages if the weapon begins to deviate from nominal behavior. This augmentation approach allowed The Boeing Company to develop and test the new laser variant without expensive wind tunnel testing, reducing development costs and schedule. The hardware modifications to create the L-JDAM weapon included the addition of a sensor nose kit (the sensor fit into the existing fuse well), wire harnesses, straps with barrel nuts, and symmetric tail kit dog ears where the sensor wire harness enters into the tail kit.



Laser guided MK - 82 scores direct hit against a moving target during tests at Eglin AFB

Affordable hit-to-kill accuracy minimizes collateral damage

Control of the Flexible Ares I-X Launch Vehicle

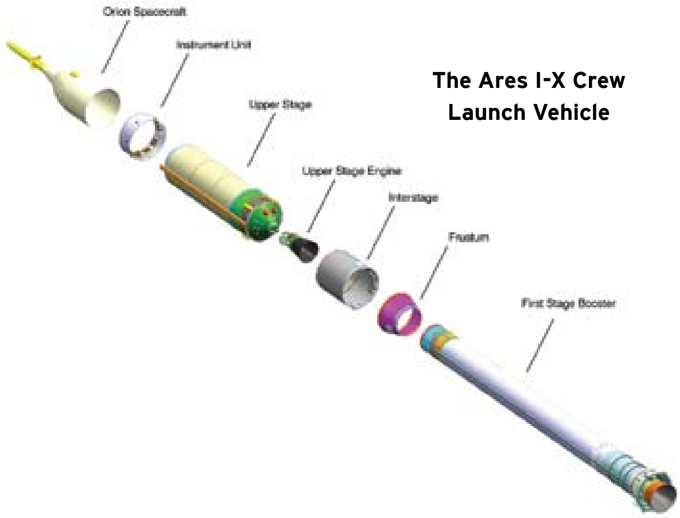
Ares I-X Flight Test Launch, October 28, 2009. Photo courtesy of NASA.

The Ares I-X flight test launch was the first flight test of an experimental launch vehicle as part of the NASA Constellation Program. The Ares I Launch Vehicle is planned as the crew launch vehicle replacement for the Space Shuttle, which is scheduled for retirement in 2011.

The Ares I-X configuration resembles the Saturn V vehicle but differs in ways that are significant from a flight control perspective. The first stage is a single, recoverable solid rocket booster derived from the Shuttle program. The first and upper stages are separated by a frustum and an interstage that houses the roll control system and avionics. A single liquid propellant engine powers the upper stage, and the upper stage reaction control system is located on the aft end of that stage. A redundant inertial navigation unit (RINU) and flight computers used for guidance, navigation, and control are located in the instrument unit at the top of the upper stage.

Challenges of Flexible Launch Vehicle Control

The ascent flight control system (AFCS) design for a flexible launch vehicle such as the Ares I-X is challenging due to the wide range of dynamic interactions between the vehicle and its environment, as well as varying mass properties, aerodynamic loads. and propulsion system characteristics that must be accommodated to maintain adequate margins on stability and performance. Launch vehicles are typically aerodynamically unstable due to the center of pressure being located above the center of mass. Ares I-X had an atypically large negative static margin due to the mass distribution in the first stage and the larger diameter upper stage. The low-frequency unstable aerodynamics were readily compensated by the relatively high-bandwidth first stage thrust vectoring. Due to the separation of control effectors (thrust vectoring) and flight control sensors (typically in the upper stage), control of the first bending mode was non-minimum phase, as is typical of flexible launch vehicles. These challenges are compounded by uncertainties in aerodynamics, ascent wind profiles, and the variability of vehicle mass properties and structural dynamics as propellant is consumed during flight. Lessons learned on the Ares I-X will lead to better design practices for the next generation of human-rated and heavy lift launch vehicles.



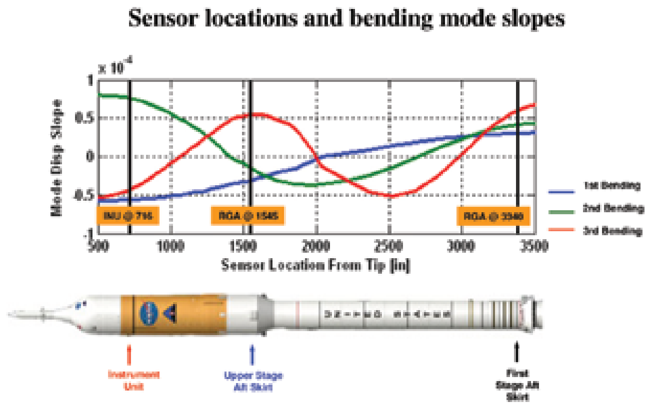
The Ares I-X Crew Launch Vehicle

Contributor: Mark Whorton, Teledyne Brown Engineering, USA

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

AFCS Design Process

The Ares I-X AFCS design approach begins with PID control designs for rigid-body performance in pitch and yaw and a phase plane control design for roll control. Multiple rate sensors are located along the structure to allow for blending of the sensed rotational rate. The need to increase robustness to force and torque disturbances such as those caused by wind and thrust vector misalignment led to the development of an anti-drift channel option for the autopilot. Unique to the Ares I-X flight test was the introduction of "parameter identification" maneuvers during ascent flight. Flight test instrumentation measured the dynamic response of the vehicle to these programmed torque commands, and post-flight data analysis was conducted to validate vehicle parameters obtained from test and analysis.

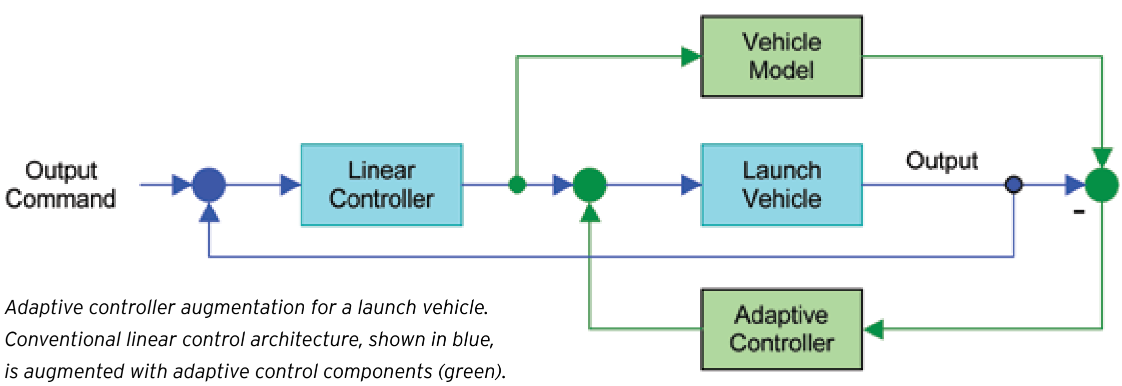


Adaptive Control: A Promising Future Trend in Launch Vehicle Control

Design and analysis of the Ares I-X AFCS indicates that classical control is sufficient to meet stability and performance requirements. Yet adaptive control concepts used in conjunction with classical approaches afford the opportunity to improve performance with increased robustness and crew safety.

Recent trends in adaptive control augmentation of existing controllers leverage the wealth of engineering heritage and experience in the development of classical control of launch vehicles while allowing for adaptation to recover and enhance stability and performance in the event of off-nominal vehicle response (see figure below). Having shown promise in missile and aircraft flight tests, these methods are showing preliminary benefits in design and analysis for the next generation of flexible launch vehicles as well.

Implementation of adaptive control for future launch vehicles will require technical and cultural transitions whereby new suites of tools for analysis and proof of stability and performance will gain confidence with program managers. Early progress is being made through theoretical developments that bridge the gap between classical and adaptive control. These developments are demonstrating analogs to traditional gain and phase margins using Monte Carlo-based gain-margin assessment and metrics such as time-delay margins. As these new "acceptance paradigms" mature and gain validity through practice, the next generations of aerospace vehicles will be safer and more capable than is possible with today's technology.



Adaptive controller augmentation for a launch vehicle. Conventional linear control architecture, shown in blue, is augmented with adaptive control components (green).

For further information: C. Hall, et al., Ares I flight control system overview, AIAA-2008-6287; M. Whorton, C. Hall, and S. Cook, Ascent flight control and structural interaction for the Ares-I crew launch vehicle, AIAA-2007-1780.

H-infinity Control for Telecommunication Satellites

Artist's rendering of Eutelsat W2A satellite in orbit, based on the Spacebus 4000 C4 platform, with deployed solar arrays and 12-m-diameter antenna. Source: Thales Alenia Space

In the early 1990s, European space industries initiated research in robust control, specifically H-infinity (H_∞), through close collaboration with external control laboratories. The research was motivated not only by the desire to gain experience in this new method, but also to evaluate its potential benefits, performance improvements, and development costs when compared to traditional (proportional-integral-derivative and linear-quadratic-Gaussian) controllers commonly used in the 1980s. The increasing performance and dynamic complexity of future space applications were also a source of motivation for the research program on robust control techniques.

The transfer of robust control techniques from research laboratories to industrial space applications covered not only the technique itself, but also the process-oriented engineering tools and methodologies required for modeling, design, and analysis of robust H_∞ controllers.

Telecommunication Satellite Control System Design: Challenges and Needs

Geostationary telecommunication satellite platforms typically consist of a central body and large (deployable) antennae together with low-damping flexible solar arrays that are rotating with respect to the Earth-pointing central body at a rate of one rotation per day. During orbit inclination correction maneuvers, the satellite is submitted to thruster-induced disturbance torques that require some few tens of nanometers control authority to limit the attitude depointing below 0.1 deg. Because of the low damping (typically 10^{-3}) and shifting frequency modes with high resonant peaks of the large rotating solar arrays (see Figure 1), a stiff filtering controller is required. Using classical control design techniques, the design problem is solved in an ad hoc fashion requiring skilled engineers to initiate the lengthy iterative design procedures, tune the convergence control parameters, and balance the multi-objective performance index.

The limited capability of the classical design procedures to adapt to other space control problems prompted the need to develop automated control design techniques, including systematic procedures to rapidly adapt to changes in dynamic models, to rapidly optimize performance under constraints of parameter uncertainties, and to address "flexible structure control" formulations in the frequency domain. From an industrial perspective, there was also a need to improve European system integrators' competitiveness within the global space market by reducing the overall telecommunication satellite development time. H_∞ techniques enabled fulfillment of these requirements.

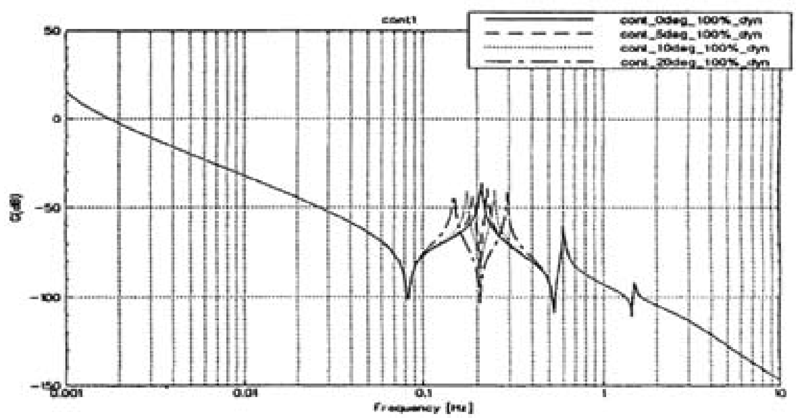


Figure 1: Structural flexible modes (above) of Astrium communication satellite platform Eurostar 2000+ (left). The shifting of frequency modes corresponds to different angular positions of the solar arrays. Source: EADS

Contributor: Christian Philippe, European Space Agency, The Netherlands

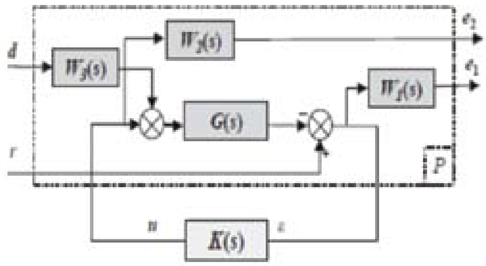


Figure 2: Standard four-block H_∞ scheme.
Source: Thales Alenia Space

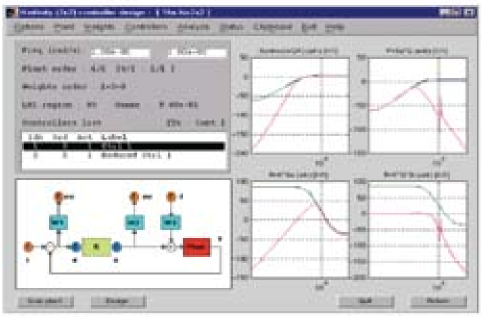


Figure 3: Ready-to-use engineering software tool based on MATLAB from Mathworks.
Source: Thales Alenia Space

Standard Four-Block H_∞ Scheme

The standard four-block H_∞ problem corresponds to the scheme of Figure 2, where e_1 and e_2 are signals to be controlled with respect to reference input r and disturbance input d. The closed-loop control objectives are attained through an appropriate tuning of the weighting filters $W_i(s)$ chosen to shape the four transfer functions from r and d inputs to ε (control error) and u (command) outputs. The weighting filters are tuned to ensure good reference tracking and disturbance rejection as well as to meet the desired control bandwidth and rolloff attenuation of the flexible structural modes. The H_∞ problem consists of finding the controller K(s) that fulfills the four main control design objectives:

1. Guaranteeing stability margins
2. Filtering flexible structural modes from solar arrays and/or antennas and propellant sloshing modes
3. Minimizing the attitude degradation in the presence of external and internal disturbances
4. Minimizing sensor noise transmission and fuel consumption

The standard four-block H_∞ scheme benefits from attractive numerical and physical properties and analytically guarantees the stability margin and robustness. Although the scheme offers an all-in-one design procedure, it must be used in association with order-reduction techniques to obtain the final controller. The all-in-one approach prevents the designer from having to do repeated iterations between preprocessing, optimal design, and post-processing, as experienced in the classical control design procedure. The all-in-one control design procedure has been implemented in a ready-to-use engineering software tool based on MATLAB from Mathworks and in dedicated control toolboxes (Figure 3).

Benefits

The development of H_∞ controllers for European telecommunication satellite platforms such as Eurostar 3000 and Spacebus 4000 has helped reduce the duration of the orbit inclination correction maneuver by 50%, equivalent to a propellant mass savings of about 20% when compared to the classical control design technique based on proportional-integral-derivative control combined with specific filters in the flexible modes frequency range.

Ariane 5 launcher (top) and SILEX (bottom). Source: EADS Astrium

Other Real-World Applications

H_∞ controllers have also been developed, implemented, and successfully flown on the Ariane 5 Evolution launcher, the Automated Transfer Vehicle (ATV), and earth observation, scientific, and exploration satellites, as well as pointing, acquisition, and tracking (PAT) systems. The benefits of applying H_∞ control techniques for Ariane 5 and the first European optical communication terminal in orbit (SILEX) are summarized below:

Ariane 5 Evolution Launcher
- Issue: structural bending and sloshing modes
- H_∞ controller synthesis (atmospheric phase)
- Benefit: thrust vector control actuation effort minimized

Optical Laser Terminal (PAT)
- Issue: performance limitations of traditional controller
- H_∞ controller synthesis (tracking mode at 20 Hz)
- Benefit: pointing stability performance of 0.1 µrad

For further information: B. Frapard and C. Champetier, H_∞ techniques: From research to industrial applications, Proc. 3rd ESA International Conference, Noordwijk, Netherlands, November 26-29, 1996; G. Pignie, Ariane 5 and Ariane 5 evolution GN&C overview, 34th COSPAR Scientific Assembly, The Second World Space Congress, Houston, TX, October 10-19, 2002; C. Charbonnel, H and LMI attitude control design: Towards performances and robustness enhancements, Acta Astronautica, vol. 54, pp. 307–314, 2004.

Control for Formula One!

A ballscrew inerter (flywheel removed) made at Cambridge University, Department of Engineering, in 2003, designed by N.E. Houghton

In August 2008, the deployment of a novel mechanical control device in Formula One racing was announced. Developed at the University of Cambridge by Malcolm Smith and colleagues, the device, called an "inerter," was deployed by the McLaren team in 2005 in Barcelona.

What Is an Inerter?

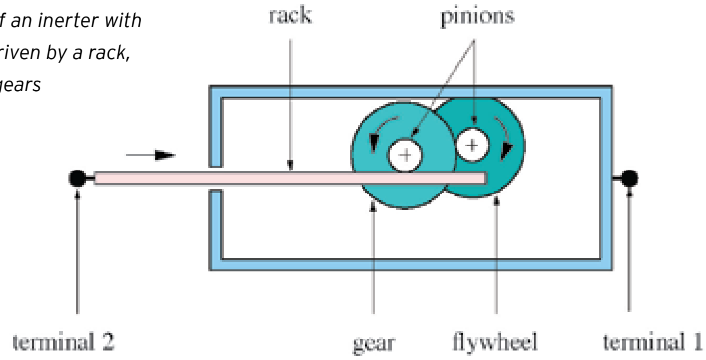
The standard analogy between mechanical and electrical networks relates force to current and velocity to voltage. The following correspondences exist between standard modeling elements:

- spring ↔ inductor
- damper ↔ resistor
- mass ↔ capacitor

The correspondence is perfect for the spring and damper, but the mass element is analogous to a grounded electrical capacitor and not to a general two-terminal capacitor. Without a two-terminal capacitor equivalent, mechanical systems are unable to provide the same flexibility in dynamic response that electrical systems can. The two-terminal electrical model suite above can be used to produce any "passive impedance" device.

The inerter overcomes this limitation of mechanical systems; this two-terminal element has the property that the applied force at the terminals is proportional to the relative acceleration between them.

Schematic of an inerter with a flywheel driven by a rack, pinion, and gears



Kimi Raikkonen crosses the finish line to take victory for McLaren in Barcelona 2005 in the first car to race the inerter. Photo courtesy of LAT Photographic

The First Application: Vehicle Suspensions

Malcolm Smith's group at Cambridge University, in attempting to build high-performance mechanical impedances for car suspensions, realized that the lack of a true capacitor equivalent was a fundamental limitation.

After several fruitless efforts to prove that such a device could not exist, they realized it could be built, and in a relatively simple manner. They ultimately developed several prototypes of the device they called the inerter.

Contributor: Malcolm Smith, Cambridge University, UK

From the Laboratory to the Racetrack

Analyses of inerter-based suspensions indicated a potential performance advantage for vehicle suspensions that might be large enough to interest a Formula One team. Cambridge University filed a patent on the device and then approached McLaren in confidence. McLaren signed an agreement with the University for exclusive rights in Formula One for a limited period.

After a rapid development process, the inerter was raced for the first time at the 2005 Spanish Grand Prix by Kimi Raikkonen, who achieved a victory for McLaren.

Stolen Secrets . . . and the Truth Ultimately Comes Out

During development, McLaren invented a decoy name for the inerter (the "J-damper") to keep the technology secret from its competitors for as long as possible. The "J" has no meaning and is just a ruse, and of course the device is not a damper. The idea behind the decoy name was to make it difficult for personnel who might leave McLaren to join another Formula One team to transfer information about the device and in particular to make a connection with the technical literature on the inerter, which Malcolm Smith and his group were continuing to publish.

This strategy succeeded in spectacular fashion during the 2007 Formula One "spy scandal," when a drawing of the McLaren J-damper came into the hands of the Renault engineering team. The FIA World Motor Sport Council considered this matter at a hearing in December 2007. According to the Council finding, "[a drawing of McLaren's so-called J-damper] was used by Renault to try to have the system that they thought McLaren was using declared illegal. This failed because Renault had certain fundamental misunderstandings about the operation of the J-damper system." A full transcript of the decision is available on the FIA website: http://www.fia.com/mediacentre/Press_Releases/FIA_Sport/2007/December/071207-01.html.

Neither the World Motor Sport Council nor McLaren made public what the J-damper was. Thereafter, speculation increased on Internet sites and blogs about the function and purpose of the device. Finally, the truth was discovered by Craig Scarborough, a motor sport correspondent from *Autosport* magazine. *Autosport* ran an article on May 29, 2008, which revealed the Cambridge connection and that the J-damper was an inerter.

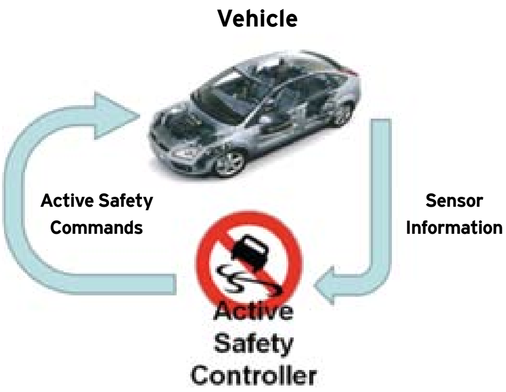
More Applications Anticipated

With the truth out, and McLaren's exclusivity expired, Cambridge University decided to enter a license agreement with Penske Racing Shocks, enabling Penske to supply inerters to any team in Formula One. A more widespread use of inerters in Formula One is now anticipated. The Cambridge University research group is working with partners to develop other applications of the inerter.

Kimi Raikkonen leading the field in the McLaren-Mercedes MP4-20 at the Spanish Grand Prix, May 8, 2005, Circuit de Catalunya, Barcelona, Spain. Photo courtesy of LAT Photographic

For further information: M.C. Smith, Synthesis of mechanical networks: The inerter, IEEE Transactions on Automatic Control, vol. 47, no. 10, October 2002; http://www.admin.cam.ac.uk/news/dp/2008081906; http://www.eng.cam.ac.uk/news/stories/2008/McLaren

Active Safety Control for Automobiles



The rapid evolution of technology over the last 20 years has made automobiles much safer than ever before. Active safety is a relatively young branch of the automobile industry whose primary goal is avoiding accidents and at the same time facilitating better vehicle controllability and stability, especially in emergency situations.

The driver + vehicle + environment form a closed-loop system, with the driver providing control actions by manipulating three primary actuators: the steering wheel and the brake and accelerator pedals. In certain cases, as a result of environmental or vehicle conditions, or the driver's actions, the car may end up in an unsafe state, with the driver's ability to control the vehicle curtailed. Active systems correct such situations by automatically applying differential braking and cutting engine torque (and in the near future, correction of wheel turn).

Some Active Safety Control Mechanisms

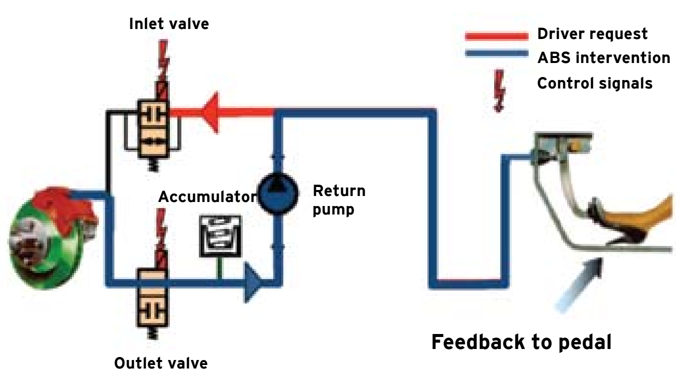
- Antilock braking systems (ABSs; available today)
- Traction control (TC; available today)
- Electronic stability control (ESC; available today)
- Automatic steering correction (future)

Antilock Braking

The first active electronic safety system was the anti-skid Sure-Brake system proposed by Chrysler and Bendix in 1971; a previous all-mechanical system was introduced by Dunlop in 1950 for aircraft. The first production use was in 1978 when Bosch mounted an ABS on trucks and the Mercedes-Benz S-Class.

The main objectives of ABS are to minimize stopping distance under braking and to avoid wheel locking to maintain the drivability of the vehicle. Since wheel locking occurs when the slip ratio between road and tire (that is, the normalized difference between the peripheral velocity of the tire and the longitudinal velocity of its hub) exceeds a maximum value, the ABS tries to avoid this situation.

As depicted in the figure below, the driver, through the brake pedal, imposes a certain pressure in the hydraulic system. The inlet and outlet valves initially work for normal braking, that is, open and closed, respectively (the opposite of the situation in the figure); in this case, the brake fluid (in the red branch) pushes the caliper into the braking disk. If this braking action determines a slip ratio on the wheel close to the maximal slip ratio, the control strategy changes the state of valves by closing the red branch and opening the blue one so that the pressure on the caliper decreases (and hence the slip ratio). The inversion of fluid flow causes a "feedback" vibration at the pedal. Notice that the opening/closing actions of the hydraulic system are cyclical (a form of high-frequency switching control) such that the slip ratio is kept close to its maximal value. The principal manufacturers are Bosch, Delphi, Continental Teves, and Kelsey-Hayes, which formed a group in 2000 called the ABS Education Alliance and estimated that almost 28% of the accidents on wet roads are caused by uncontrolled braking.



ABS in operation: the automatic release phase when the inlet/outlet valves are in closed and open status, respectively.

Contributor: Luigi Glielmo, Università del Sannio, Italy

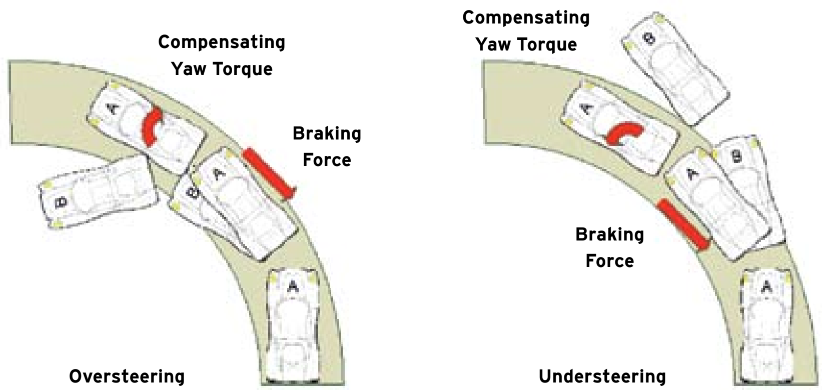
Cost-benefit analyses of these systems for EU-25 show that in the decade 2010-2020, the use of ESC can return benefits (in terms of accident avoidance) of €2.8-4.4 for each euro spent. This has convinced governments to make the installation of ESC systems on all cars in the European Union and the U.S. mandatory from 2012. ABS is not mandatory, but discussion is under way to make it mandatory for motorbikes in the U.S.

Traction Control

Traction control (TC) systems (or anti-slip regulators) have the opposite goal of ABS in that they try to keep the wheels from spinning in acceleration. This is done by maintaining the slip ratio (opposite in sign with respect to the braking situation) within a certain threshold, modulating the traction torque on the wheels. TC is available in two different versions: one, produced by Saab in collaboration with Teves and Hella, uses the braking system and engine torque variation; the other one, produced by Honda and Bosch, uses only the engine torque variation.

Electronic Stability Control (ESC)

ABS only works well during longitudinal panic braking and TC in start-up maneuvers; neither is effective when vehicle stabilization involves lateral dynamics (sideslips). ESC systems fulfill this need. They act on individual brakes and possibly engine torque, based on measurements or estimated errors of two vehicle variables and their respective (computed online) reference signals: the yaw velocity (the angular velocity around the vertical axis) and the sideslip angle (the angle between the longitudinal axis of the vehicle and the direction of the velocity vector). In particular, the yaw velocity must track a reference trajectory computed on the basis of the steering wheel angle and the vehicle velocity, and the sideslip angle must not exceed a certain threshold. The whole control action (estimation + actuator command generation) is performed in a strict sampling time (10-20 ms). Human drivers would not be able to simultaneously coordinate braking of four individual wheels and cutting of engine torque (if longitudinal velocity is too high) so as to correct the vehicle direction.



Correcting oversteer and understeer with ESC

The first commercial ESC was developed between 1987 and 1992 by Mercedes-Benz and Robert Bosch GmbH. Today ESCs are available under trade names such as AdvanceTrac, Dynamic Stability Control (DSC), Dynamic Stability and Traction Control (DSTC), Electronic Stability Program (ESP), Vehicle Dynamic Control (VDC), Vehicle Stability Assist (VSA), Vehicle Stability Control (VSC), Vehicle Skid Control (VSC), Vehicle Stability Enhancement (VSE), StabiliTrak, and Porsche Stability Management (PSM). These products differ in the combination of actuators used and the conditions for activating the control strategy.

The Future: Advanced Model-Based Control

Active safety control systems are typically designed using gain-scheduled single-input, single-output controllers whose calibration is obtained after extensive real-time simulations and tests on the track. Furthermore, the coordination among multiple subsystems is done through heuristic rules that determine activation conditions and manage shared resources. Limitations of this approach are that new actuators or sensors are difficult to integrate and it cannot take into account from the beginning the multivariable and constrained nonlinear nature of the global problem. Hence, research is under way to introduce more complex model-based and robust control design methods, exploiting the increased computational power available on board.

Automated Manual Transmissions

Shift buttons on the steering wheel of a FIAT Bravo (Source: www.fiat.it)

The automated manual transmission (AMT) is an intermediate technological solution between the manual transmission used in Europe and Latin America and the automated transmission popular in North America, Australia, and parts of Asia. The driver, instead of using a gear shift and clutch to change gears, presses a + or – button and the system automatically disengages the clutch, changes the gear, and engages the clutch again while modulating the throttle; the driver can also choose a fully automated mode. AMT is an add-on solution on classical manual transmission systems, with control technology helping to guarantee performance and ease of use.

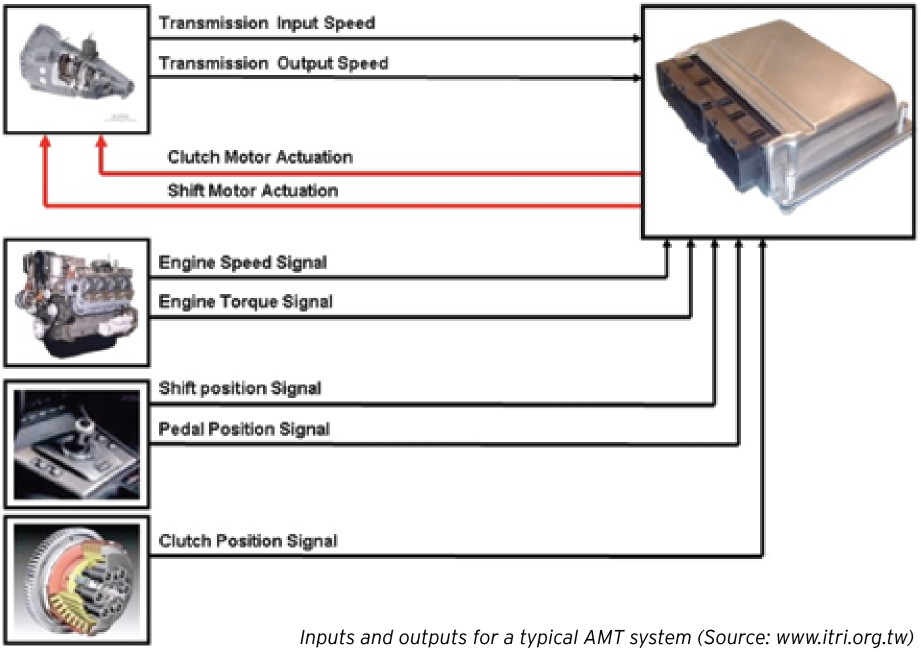
AMT Overview

An AMT is composed of a dry clutch, a gearbox, and an embedded dedicated control system that uses electronic sensors, processors, and actuators to actuate gear shifts on the driver's command. This removes the need for a clutch pedal while the driver is still able to decide when to change the gear. The clutch itself is actuated by electronic equipment that can synchronize the timing and the torque required to make gear shifts quick and smooth. The system is designed to provide a better driving experience, especially in cities where congestion frequently causes stop-and-go traffic patterns.

AMTs have been used in racing cars for many years, but only recently have they become feasible for use in everyday vehicles with their more stringent requirements for reliability, cost, and ease of use.

Benefits of AMT

- Changing gears without using a foot to operate the clutch
- No engine or gear modifications
- Less physical or psychological stress
- More comfortable than manual transmissions
- More "fun" factor compared to fully automatic transmissions



Inputs and outputs for a typical AMT system (Source: www.itri.org.tw)

Contributor: Luigi Iannelli, Università del Sannio, Italy

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

AMT systems are currently installed by several automakers under different commercial names, such as SeleSpeed by FIAT, Sequential Manual Gearbox by BMW, 2Tronic by Peugeot, SensoDrive by Citroen, and EasyTronic by Opel.

Commercial DCT systems include the Direct-Shift Gearbox by Volkswagen Group and the Dual Dry Clutch Transmission by FIAT Group.

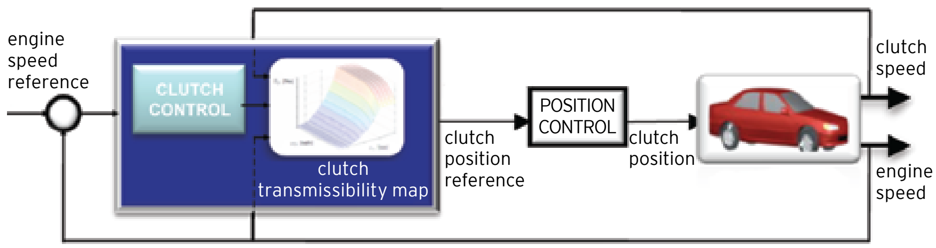
Inventions and Innovations

AMT is an interesting example showing the importance and potential of automatic control. The control of the clutch engagement on AMT systems must satisfy different and conflicting objectives:

- It should result in the same or better shifting times as with manual transmissions.
- It should improve performance in terms of emissions and facing wear.

In a typical AMT control scheme, a constant engine speed is requested during the engagement so as to equalize engine and clutch torques as well as possible. In this case, the clutch control provides a clutch torque reference, and through a suitable model (or map), the torque reference is converted into a position reference for the clutch actuator position control (see figure below).

Commercial implementations of AMT today rely on enhancements of PID controllers with feedforward actions and controller gain scheduling.



Future View: Toward Model-Based Control of AMTs

Model-based approaches are attracting increasing interest as evidenced by several control strategies that have recently been proposed in the literature. These strategies are based on optimal control, predictive control, decoupling control, and robust control.

Innovative AMT technology uses a dual-clutch transmission (DCT) consisting of one clutch for odd gears and another for even gears. The goal is to improve the speed and comfort of the gear shift. But effective AMT controllers, particularly for dual-clutch systems, are difficult to design without an accurate model of the clutch torque transmissibility characteristic, or the relationship between the clutch actuator position (or the pressure applied by the clutch actuator) and the torque transmitted through the clutch during the engagement phase.

The clutch transmissibility model, key to advanced control of AMTs, is difficult to attain: it depends on various parameters and phenomena, such as friction pad geometries, cushion spring compression and load, and slip-speed-dependent friction. Accurate clutch transmissibility models will allow the use of advanced model-based control strategies aimed at improving the overall behavior of the system with respect to current commercial solutions.

Mobile-Robot-Enabled Smart Warehouses

Order fulfillment is a multibillion-dollar business. Existing solutions range from the highly automated—whose cost-effectiveness is inversely related to their flexibility—to workers pushing carts around in warehouses and manually filling orders—which is very flexible but not very cost-effective. Kiva Systems uses a new approach to order fulfillment: operators stand still while the products come to them. Pallets, cases, and orders are stored on inventory pods that are picked up and moved by hundreds of mobile robotic drive units. As a result, any product can be moved to any operator.

Successful Installations Worldwide

Kiva Systems has deployed more than a dozen installations worldwide, including a 1,000-mobile-robot system for a retail company in the United States. Customers include:

- Staples
- Walgreens
- Boston Scientific
- Zappos (acquired by Amazon in 2009)
- Crate & Barrel
- Saks Incorporated
- The Gap
- Quiet Logistics
- Diapers.com

Contributor: Raffaello D'Andrea, Kiva Systems, Switzerland

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Winner of the 2008 IEEE-IFR Invention & Entrepreneurship Award in Robotics & Automation. Founders Mick Mountz, Peter Wurman, and Raffaello D'Andrea received this award, whose aim is to foster innovation paired with entrepreneurial spirit, and make the best possible use of synergies between science and industry in the fields of robotics and automation.

Kiva Systems ranked number 6 (number 1 in logistics) in the Inc. 500 ranking of the fastest-growing private companies in the United States.

System Description

Kiva uses hundreds of mobile robots and powerful control software to provide a complete fulfillment solution: storing, moving, and sorting inventory. Instead of being stored in static shelving, flow racks, or carousels, products are stored in inventory pods in the center of the warehouse while operators stand at inventory stations around the perimeter.

- When an order is received, robotic drive units retrieve the appropriate pods and bring them to the worker, who picks out the correct item and places it in the carton. Completed orders are stored on separate pods, ready to be picked up and moved to the loading dock when the truck arrives.
- The Kiva drive units are differential-drive two-wheeled robots with a patent-pending mechanism for lifting pods off the ground. This mechanism is essentially a large actuated screw; by rotating a drive unit underneath a pod and simultaneously counter rotating the screw, a pod can be lifted off the ground.
- A suite of sensors on the drive units and custom control software and algorithms allow the vehicles to safely navigate the warehouse. Coordination is aided by a hierarchical layer similar to that used in air traffic control systems.
- The drive units share information about their environment and use that knowledge to adapt. As a result, the performance of the vehicles, and hence the system, improves over time. In addition, adaptation and learning ensure that the system is robust to changes in the environment.

Select Customer Quotes

"Our customers expect to get great value and service from Crate & Barrel, but they also care about our carbon footprint. This played a role in our selection of Kiva Systems," said John Ling, vice president of supply chain management and logistics at Crate & Barrel. "Kiva's mobile robotic approach is not only the most cost-effective way to automate pick, pack and ship operations, but also the greenest. The robots themselves are energy efficient, plus the entire robot zone can be operated with almost no lighting."

"Using a flexible, automated order fulfillment system helped our Piperlime operations scale to increased capacity over the critical holiday season," said Chris Black, Vice President of Operations at Gap Inc. Direct. "The system freed up our employees' time, allowing them to focus on processing a higher volume of customer orders faster and to ensure more accuracy. We're looking forward to leveraging Kiva's system when we expand our online business internationally."

"Other material handling approaches would have required us to integrate different technologies to handle units, cases and pallets, as well as a wide range of product sizes," said Ken Pucel, Executive Vice President, Operations, at Boston Scientific. "Kiva is able to provide us a proven solution with the flexibility and ease-of-use of a single technology for our needs."

For further information: www.kivasystems.com

Dynamic Positioning System for Marine Vessels

The double-hulled dynamically positioned drillship DISCOVERY SPIRIT equipped with slx aquamaster thrusters

Development Driller III: Fifth generation, dynamic positioning semi-submersible ultra-deepwater drilling rig build by Keppel FELS Singapore.

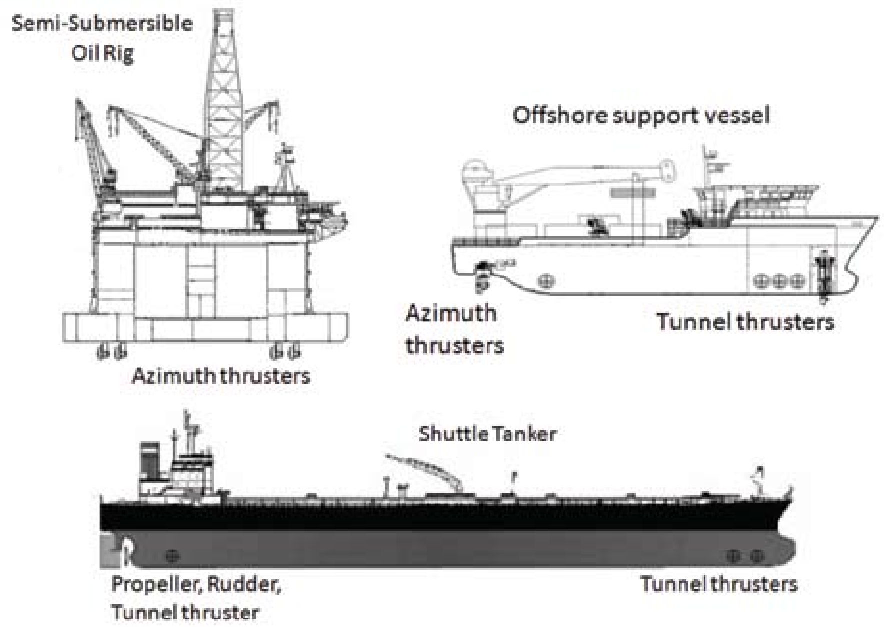
A dynamically positioned (DP) vessel maintains its position (fixed location or predetermined track) by means of active thrusters. The DP system can also be used in combination with mooring and anchoring to form position mooring systems for energy efficiency.

DP-operated vessels possess the ability to operate with positioning accuracy, safety, and reliability. Such systems have gained the trust and acceptance of the industry and the International Maritime Organization and have been successfully applied worldwide.

The advantages of fully DP-operated vessels include the ability to operate with positioning accuracy and the flexibility to establish position and leave location fast, without the need for mooring lines to be deployed. In addition, there may be restrictions on the deployment of anchors due to the already installed subsea structures on the seabed. For certain deepwater exploration and production scenarios, DP-operated vessels may be the only feasible solution due to the depth and length of mooring lines required.

A dynamic positioning system allows a vessel to automatically maintain its position and heading through the coordinated control of thrusters.

Various DP Vessels and Typical Actuator Setups



Implementation

Dynamic positioning systems have been installed on vessels used worldwide. Typical DP vessels include survey vessels, drilling ships, work boats, semi-submersible floating rigs, diving support vessels, cable layers, pipe-laying vessels, shuttle tankers, trenching and dredging vessels, supply vessels, and floating, production, storage and offloading vessels (FPSOs).

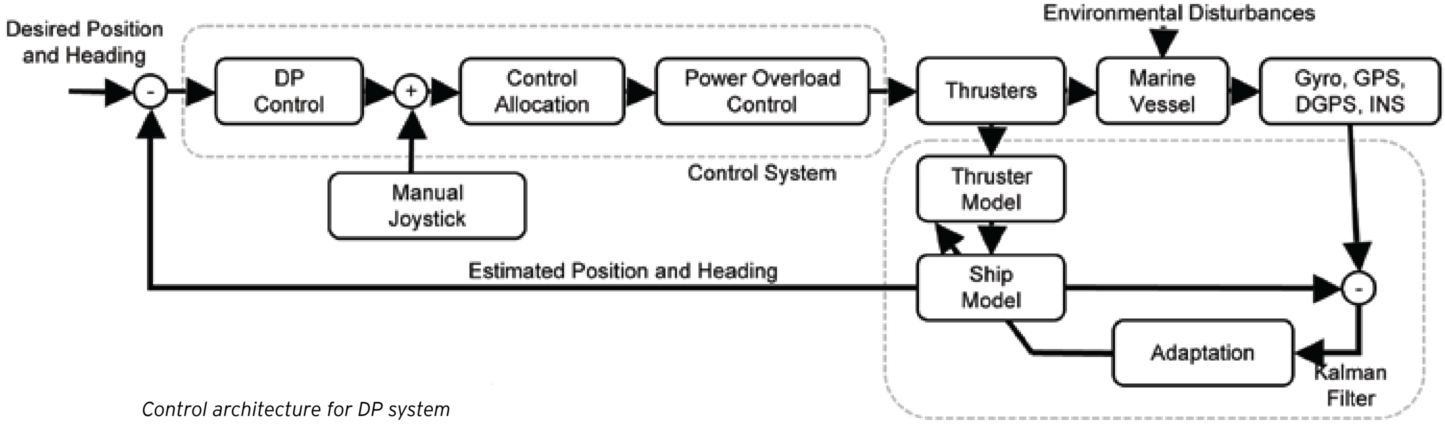
Contributors: Shuzhi Sam Ge, Choo Yoo Sang, and Bernard Voon Ee How, National University of Singapore, Singapore

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Position References

For oil and gas exploration and production, DP rigs can be configured to operate in water depths of up to 3000 m. At offshore locations, the most reliable form of position reference for the surface vessel is differential GPS (DGPS). Two or three separate and distinct DGPS systems provide the required redundancy through the use of differential correction links.

For drilling operations, it is important for the vessel to keep its position within a small envelope over the well such that the riser connecting the vessel to the well is nearly vertical.

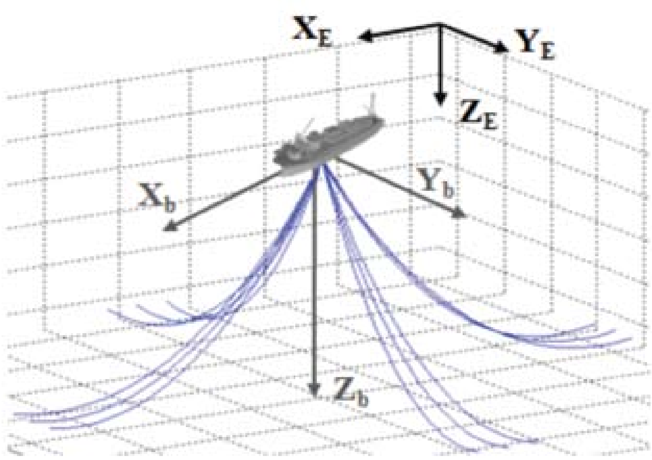


Control architecture for DP system

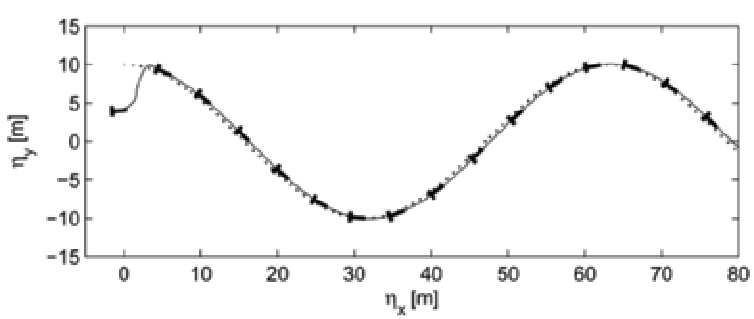
From PID to Advanced Control

The first DP systems introduced in the early 1960s used conventional PID controllers in cascade with low-pass and/or notch filters to suppress the first-order wave-induced motion components. From the mid-1970s, more advanced control techniques were proposed based on linear optimal control and Kalman filter theory.

With improvements and increasing sophistication in vessel control, the marine industry can look forward to more advanced control features such as DP-assisted position mooring systems, automatic maneuvering in shallow water and harbor areas, formation sailing, and automatic collision avoidance. These applications open new possibilities for the expansion of functionality in DP systems.



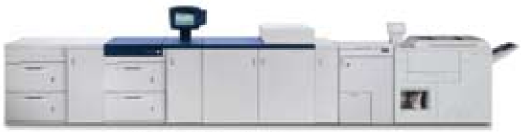
DP-assisted position mooring system and frames of reference



Tracking control of fully actuated ocean surface vessels

For further information: Social Robotics Laboratory, NUS, http://robotics.nus.edu.sg; Centre for Offshore Research and Engineering, NUS, http://www.eng.nus.edu.sg/core

Digital Printing Control: Print Shop in a Box



DocuColor® 8002

iGen4®

iGen4® 220 Perfecting Press

iGen3®

ColorQube™

The digital print process is remarkably challenging because it involves many process and digital actuators for applying a range of advanced control techniques. Many of the new challenges listed here opened the door for the insertion of new control theory. Numerous Xerox® printing systems (for example, iGen3®, iGen4®, DocuColor® 7002, DocuColor® 8002, DocuColor® 5000, DocuColor® 8000, iGen4® 220 Perfecting Press, Xerox® Color 800/1000 Presses, ColorQube™) produce high-quality prints using these control innovations to generate several billion dollars in revenue.

Control Challenges for Digital Printing

- Optimize job workflows via feedback from the press (streamline workflow and free operators to focus on running the print jobs)
- Increase productivity with automated color management tools
- Provide consistent color image quality (first page, between pages, job to job, operator to operator, machine to machine)
- Provide offset look and feel with best image quality, no nonuniformity, and no defects
- Provide automated calibration (completely hands-free), spot color (Pantone® matching), more stable color
- Allow mix-and-match of press configurations (any application to any printer, any finisher and feeder on any marking engine)
- Manage time-sensitive activities of various machine modules
- Adjust color dynamically using internal process control feedback loops
- Provide active control of registration of all color separations
- Compensate for sheet-to-sheet differences in the paper as well as drive system wear, temperature variations, and the like

Contributor: L.K. Mestha, Xerox Corporation, USA

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Digital printing today is a complex, high-technology process requiring advanced sensors and actuators and state-of-the-art control algorithms. Processing of images occurs at multiple levels within and outside a hierarchical printing and publishing system. Time-based separation is adopted at each level, with higher-level functions (spot colors, ICC profiles, gray balance) occurring at a slower time scale in the digital front end (DFE) and faster real-time control functions (Levels 1, 2, and 3) typically occurring in the print engine (PE).

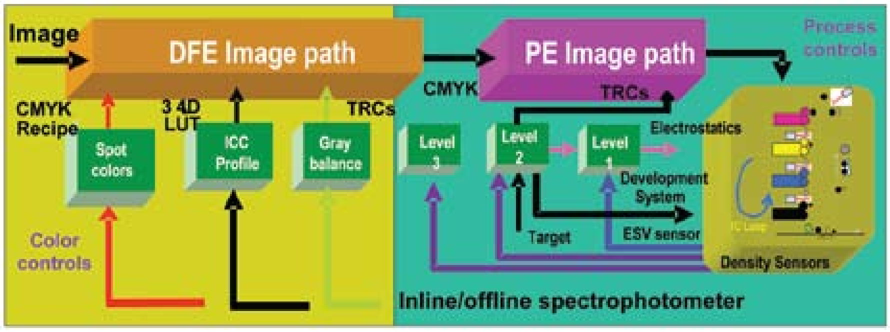


Figure legend:
LUT: look-up table, **ICC:** International Color Consortium, **TRC:** tone reproduction curve, **ESV:** electrostatic voltmeter, **TC:** toner concentration, **CMYK**: cyan, magenta, yellow, black colors.

Inventions and Innovations

Xerox's new printing systems are tours de force for control technology! Highlights include:

- A hierarchical automation architecture that distinguishes between different color control horizon levels.
- The control design includes classical single-input single-output (SISO) PID-type controllers, with delay and anti-windup compensation, for several subsystems.
- Toner concentration control, although a SISO system, is especially challenging. The solution integrates a Kalman filter to handle noisy and unreliable measurements, a Smith predictor to handle the delay, and an anti-windup compensator for constraint management.
- Several multivariable controllers are also part of the design. These include state feedback, pole placement, model predictive control, and linear quadratic regulator designs. Systems under multivariable control include electrostatic control, spot color calibration, and color management profiling (ICC profile).
- A learning algorithm is incorporated for paper registration control.
- Singular value decomposition is used for dimensionality reduction, to reduce gray-level samples while constructing spatial toner reproduction curves or functions on photoconductors.
- Ideas from cooperative control theory and simultaneous perturbation stochastic approximation have been used for gray-component replacement in the color management profiling (ICC profile) system.
- K-means clustering for spectral reconstruction in real time allowed the use of a low-cost LED-based spectrophotometer.
- For complex printing jobs, scheduling of paper sheets is a difficult operation; constraint-based scheduling algorithms are used to solve it.
- Motion and unevenness in motion can induce disturbances in the printing process. Repetitive control and adaptive feedforward control algorithms help mitigate the effects of these disturbances.
- The printers integrate diagnostic and self-monitoring features using statistical process control, among other methods.

For further information: L.K. Mestha and S.A. Dianat, Control of Color Imaging Systems, CRC Press, May 2009; L.K. Mestha and A.E. Gil, Color Controls – An Advanced Feedback System, in The Control Handbook, Second Edition, CRC Press, (to appear).

Trip Optimizer for Railroads

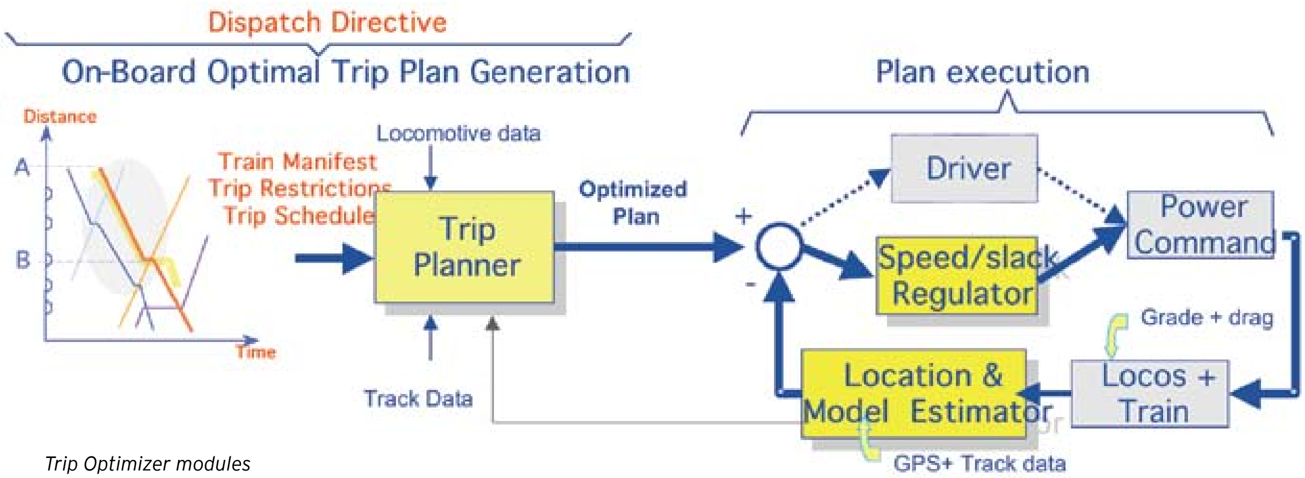
On-time arrival with the least fuel expenditure is a key priority for freight (and passenger) railroads worldwide. North American railroads consumed 4 billion gallons of fuel in 2008, accounting for 26% of operating costs.

GE's Trip Optimizer is an easy-to-use control system that allows the crew or dispatcher to achieve on-time arrival with the least possible fuel use.

Optimal driving solutions are computed onboard and executed in closed loop using GPS-based navigation. Train and track parameters are adapted online to reduce model errors. Fuel savings of 3% to 17% are realized.

Trip Optimizer Modules

- Trip Planner finds the driving strategy (speed and throttle) that minimizes fuel for the target arrival time and satisfies speed limit and other train and locomotive operating constraints.
- Speed Regulator closes the loop around the plan to correct for modeling errors and external disturbances and provides compensation for slack-action in the distributed dynamics of typical mile-long, heavy trains; both hands-off closed-loop and driver-in-the-loop "coaching" solutions are available.
- Location and Model Estimator provides precise location of the train and compensation for GPS dropouts, and adaptively tracks train parameters such as weight, length, and drag.



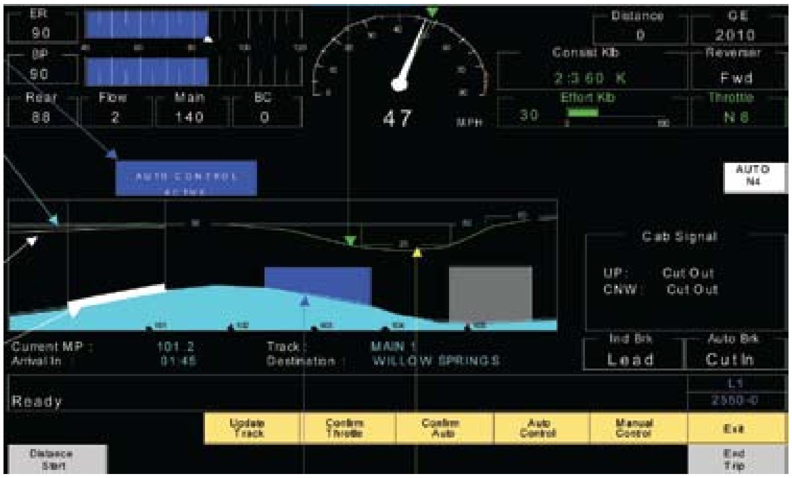
Trip Optimizer modules

Contributors: David Eldredge, GE Transporation, and Paul Houpt, GE Global Research, USA

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

For each Evolution locomotive on which it is used, Trip Optimizer can reduce fuel consumption by 32,000 gallons, cut CO_2 emissions by more than 365 tons, and cut NOx emissions by 3.7 tons per locomotive per year. If Trip Optimizer is deployed on the approximately 10,000 similar locomotives in service in North America, these savings equate to taking a million passenger cars off the road for a year.

Trip Optimizer is a product of GE Transportation, Erie, Pennsylvania.



Inventions and Innovations

Trip Optimizer provides innovative solutions to the optimization, estimation, control, and operator interface requirements for achieving fuel savings and emissions reductions for freight railroads.

Computation of the driving plan requires solving a math program with thousands of constraints and decision variables in seconds, with time- and fuel-based objectives.

Robust speed regulator design relies on a loop-shaping algorithm to maintain stable operation and deal with variations in intercar separation and the resulting forces.

Location estimator provides precise coordinate tracking via Kalman-filter-based compensation for GPS dropouts. Model-based methods adaptively track key train parameters using GPS and other locomotive data. Tools for extensive offline analysis were also developed to produce high-integrity database sources for use in control and estimation.

Innovative displays bring intuitive mode awareness and ease of use to the underlying optimal control strategy. Experienced drivers can learn the system in minutes.

Robust satellite communication from the locomotive provides rapid access to train data (and updates) directly from railroad mainframes with backup from a dedicated 24/7 GE facility.

More Than 5 Million Miles of Successful Revenue Service on North American Railroads

- Canadian Pacific
- Burlington Northern Sante Fe
- CSX
- Canadian National
- Total fuel savings to date of over 3 million gallons!

For further information: http://ge.ecomagination.com/products/trip-optimizer.html

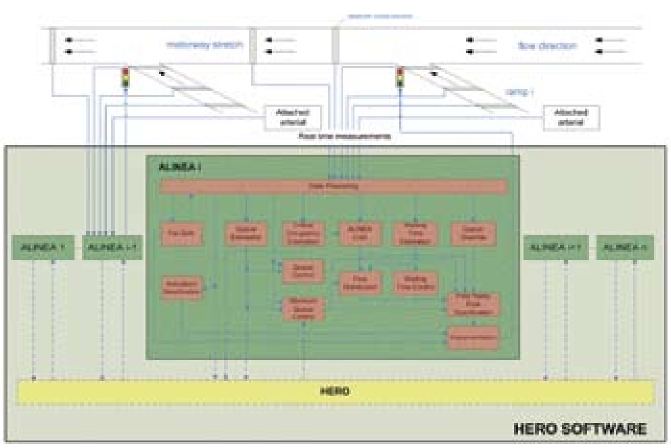
Coordinated Ramp Metering for Freeways

Freeways were originally conceived to provide virtually unlimited mobility to road users. However, the continuous increase in car ownership and demand has led to a steady increase (in space and time) of recurrent and nonrecurrent freeway congestion, particularly within and around metropolitan areas. Freeway congestion causes excessive delays, increases fuel consumption and environmental pollution, and deteriorates traffic safety.

Ramp metering, the most direct and efficient way to control freeway networks, aims at improving traffic conditions by appropriately regulating inflow from the on-ramps to the freeway mainstream. Coordinated ramp-metering strategies make use of measurements from a freeway network to control all metered ramps included therein. A new traffic-responsive feedback control strategy that coordinates local ramp-metering actions for freeway networks was developed by Prof. Markos Papageorgiou and Dr. Ioannis Papamichail at the Dynamic Systems and Simulation Laboratory of the Technical University of Crete, Greece. The proposed coordination scheme is named HERO (HEuristic Ramp metering coordination) and has been extensively tested via simulation as well as in field implementations. The developers are the recipients of the 2010 IEEE CSS Transition to Practice Award, a prize awarded by the IEEE Control Systems Society to recognize outstanding university-industry collaboration that enables the transition of control and systems theory to practical industrial or commercial systems.

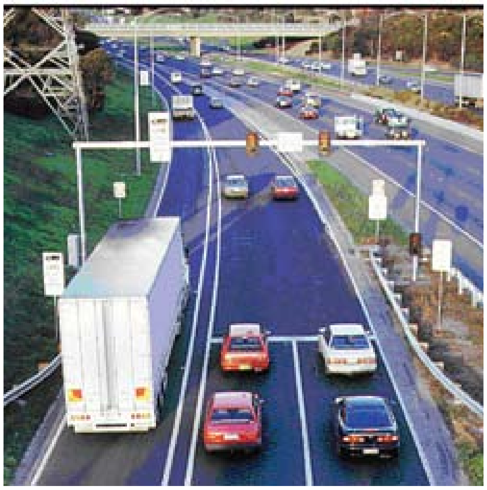
Solution Overview

HERO is simple and utterly reactive, that is, based on readily available real-time measurements, without the need for real-time model calculations or external disturbance prediction. HERO is modular in structure and includes many interacting and cooperating feedback control loops (such as mainstream occupancy control, ramp queue-length control, waiting time control) as well as two Kalman filters for estimation of ramp queue length and mainstream critical occupancy. Generic software has been developed that implements the HERO coordination scheme for any freeway network via suitable input configuration.



Contributors: Markos Papageorgiou and Ioannis Papamichail, Technical University of Crete, Greece

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.



Field Application at the Monash Freeway, Australia

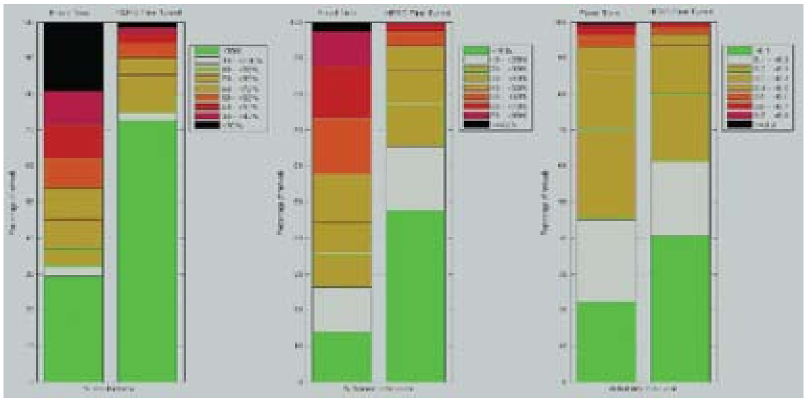
Since 2000, Melbourne's freeways have become heavily congested with extended periods of flow breakdown. The Monash Freeway, a six-lane dual carriageway that carries in excess of 160,000 vehicles per day, of which up to 20% are commercial vehicles, experiences long periods of congestion lasting between 3 and 8 hours a day.

To address this congestion problem, in early 2008 the responsible road authority, VicRoads, implemented HERO at six consecutive inbound on-ramps of the Monash Freeway. This $1M (Australian) pilot project was part of the Monash-CityLink-West Gate (MCW) upgrade and received two 2009 Victorian Engineering Excellence Awards, one for Technology and one for Engineering Innovation (http://veea09.realviewtechnologies.com). Significant benefits were demonstrated over the previous metering policy. The control logic has proven to be robust and transparent to traffic engineers. Transition to HERO has been seamless to motorists and provides significant flexibility and capability to operate the freeway close to optimal conditions. The pilot project economic payback period was just 11 days. The successful implementation and evaluation of HERO has led to its rollout during 2009/2010 at 63 sites across the entire 75-km route of the MCW upgrade project.

An evaluation of HERO's field performance was undertaken by VicRoads. HERO sensibly reduced the space-time extent of freeway traffic flow breakdown and provided significant improvements in throughput and travel speed. The a.m. peak evaluation revealed a 4.7% increase in average flow (over the previous system) and a 24.5% increase in average speed, whereas the p.m. peak evaluation showed an 8.4% increase in average flow and a 58.6% increase in average speed.

Other Field Applications

HERO was also field-implemented in a 20-km stretch of the inbound A6 freeway in the south of Paris, France, in 2006, albeit in simplified form due to lack of real-time on-ramp data in the control center. Nevertheless, results indicated a clear improvement over the existing system. In addition, HERO has been adopted for field implementation at the urban on-ramps of the A10 ring-road around Amsterdam by the responsible road authority (Rijkswaterstaat); the related implementation work should start soon. Several authorities in North America and Australia have expressed interest in adopting HERO for their freeway networks as well.



Results from the HERO implementation on the Monash Freeway using Austroads National Performance Indicators (ANPI). Three indicators are shown, with side-by-side before-and-after comparisons for each. Left to right, the indicators are: productivity (a combination of high speed and high volume on the freeway), mean speed deviation from the posted speed limit, and reliability (reflecting travel time differences from day to day).

For further information: I. Papamichail and M. Papageorgiou, Traffic-responsive linked ramp-metering control, IEEE Transactions on Intelligent Transportation Systems, vol. 9, pp. 111–121, 2008; I. Papamichail, et al., HERO coordinated ramp metering implemented at the Monash Freeway, Proc. 89th Transportation Research Board Annual Meeting, Washington, D.C., 2010.

Control in Mobile Phones

Mobile phones have made a huge impact on the world in a short time period. They are now affordable for those with daily incomes as low as a dollar, and they have brought communication infrastructure to new areas. In addition to enabling convenient and low-cost telephone services, mobile phones have also made information available at subscribers' fingertips. For many, their first contact with the Internet is with a mobile phone, not a computer.

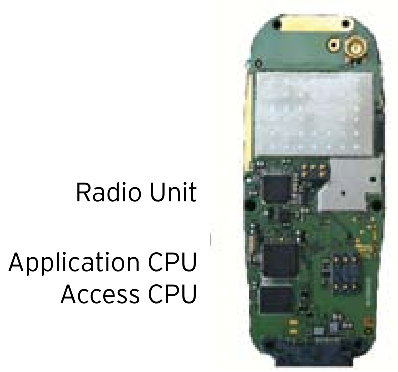
Mobile phones as affordable and attractive consumer products would not be possible without control. Each phone has at least a half dozen function-critical control loops. Control is used to reduce cost, size, and power consumption to levels where mass-produced, battery-operated products are feasible.

Control has been embedded in mobile telephones since the first large, bulky, barely portable handsets and continues to be a key technology for today's smart phones (images not to scale).

With a world penetration of 4 billion users, the number of control loops in mobile phones is in the range of 10^{10} to 10^{11}. If you choose any control loop in the world at random, it is likely located in a mobile phone, making the application area one of the major success stories of control in recent times. The area is heavily patented, with thousands of new patents granted each year, a large share of them describing control inventions.

Access Control

Each phone contains a transceiver unit that makes radio access possible with one or several base stations. Designing a low-cost transceiver that is easy to mass produce and has sufficient power efficiency, receiver sensitivity, and linearity is a major technical challenge. Some of the control loops that have enabled transceiver design with the technology components available today are automatic gain control (AGC), automatic frequency control (AFC), transmission power control, timing control, and feedback control of coding and modulation.



Radio Unit Clock Rate: ~2 GHz
Application and Access CPU: ~500 MHz
Memory: 512 MB RAM + 1 GB Flash
(data for high-end phones)

Circuit Design Level Control

Control loops are also heavily used on the electronic circuit design level, for example, in the design of low-noise amplifiers (LNAs), voltage conversion units, operational amplifiers, and power-efficient sigma-delta analog-to-digital and digital-to-analog converters. Feedback control on the circuit level is typically used to compensate for component variations due to temperature, voltage, and aging.

Application Control

In mobile telephones, application control refers to the control of on-device resources. Boundaries between mobile phones and computers are disappearing. A major challenge is to facilitate distributed application development on scalable architectures, where the amount of available computational resources, memory, and power is unknown until runtime. Thus, feedback control loops are also becoming important for controlling computational resources in mobile phones. Reliable temperature control is also important for products that lack the ability to survive critical situations by starting a cooling mechanism such as a fan.

Contributor: Bo Bernhardsson, Lund University, Sweden

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Power Control

In the most-used version (WCDMA FDD) of the 3G radio standard introduced at the beginning of the millennium, all mobile phones in a radio cell transmit simultaneously on the same frequency. A clever design of the coding scheme makes it possible to filter out and amplify the wanted part of the received signal. All other transmissions will act as noise. Thus, controlling the power of all transmitted signals is critical; failed power control in one mobile can destroy the operation of an entire cell.

The base station (BS) and mobile phones cooperate to control both the downlink signal power (BS to mobile) and uplink signal power (mobile to BS) using two control loops. An interesting coupling between these loops arises because failure in downlink control will have an impact on the communication of control commands for the uplink power, and vice versa.

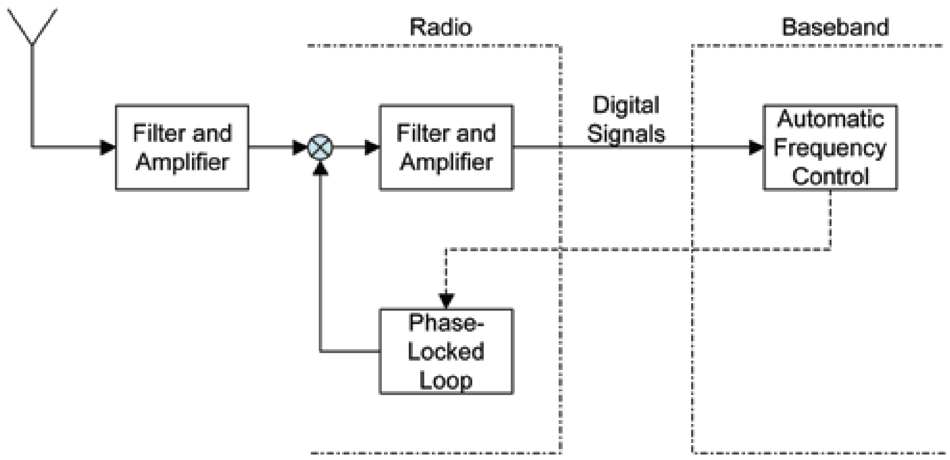
Because the controller includes integral action, anti-windup control must be used when the control loop is broken. The 3G standard includes tests for proper anti-windup. An interesting control situation also occurs during the so-called "soft handover," when several base stations simultaneously try to control the output power of the mobile.

Automatic Frequency Control (AFC)

For correct reception of radio signals, the local oscillator in the mobile phone must have the same frequency as the signal to be received. The relative frequency accuracy targeted for good reception is on the order of 0.01 to 0.1 parts per million (ppm).

Without feedback control, achieving this specification would require crystal oscillators with high power consumption. The crystal oscillators would also be large and expensive. The accuracy achievable with open-loop control and at reasonable cost is on the order of 10 ppm today. Thus feedback extends the technological frontier by a factor of 1000. The main disturbances for which feedback is essential are due to temperature variations, the Doppler effect for moving users, variations over battery voltage, and oscillator frequency and aging.

The AFC control loop locks the oscillator phase to the phase of the received radio signal using known transmitted signals, digital "pilot" symbols. The controller can be of proportional-integral (PI) type, and the main design tradeoff is to achieve good noise rejection and fast tracking of frequency variations simultaneously. Gain scheduling is typically used, with faster control for rapidly moving phones.



Mobile phones include a radio unit, which works with analog signals at high frequencies using analog circuits, and a baseband unit, which works with digital signals using digital hardware blocks and special-purpose digital signal processors. Automatic frequency control thus controls the analog phase-locked loop using digital symbols.

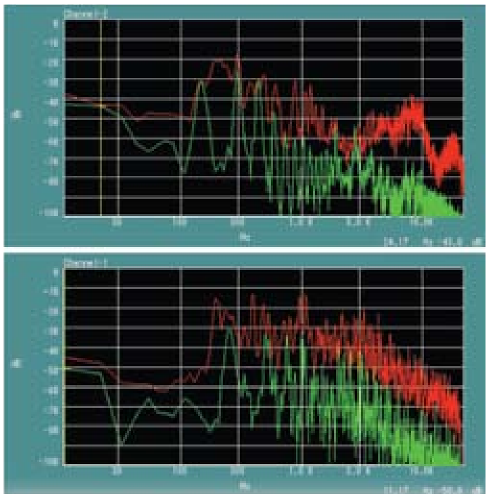
Automatic Gain Control (AGC)

The strength of the received signal shows large variations depending on the distance from the transmitter to the receiver. The receiver must show linear behavior for an operating range between -25 and -115 dBm (where 0 dBm equals 1 mW); that is, a 10^9 power change on the input (comparable to the power ratio between a lamp and a nuclear power plant).

Low-cost electronic components with such dynamic range are not feasible today. Feedback is used in several stages to control the gain of each block in the receiver chain so that the output signal fits the dynamic range of the succeeding block. The AGC loops must be sufficiently fast to track channel propagation variations. A PI controller with gain scheduling is often used.

Improved Audio Reproduction with Control Theory

To the discerning listener, sound quality in CDs and other standard digital formats leaves something to be desired. It turns out that this "something" is related to high-frequency signal elements. A new advance based on sampled-data control theory, the YY filter, has overcome this problem with audible advantages!



Fast Fourier transforms (0–20 kHz) of an analog record (top) and CD reproduction (bottom). The green trace is the FFT; the red trace is the peak FFT value over the past 10 seconds. The record exhibits a range that extends well beyond 20 kHz; the CD has a sharp cutoff at 20 kHz. (The traces are not from the same sound.)

Conventional Sound Encoding

The audible range is widely accepted to be limited to 0-20 kHz, and anything beyond is sharply cut (filtered out) by a low-pass filter. This is based on the well-known Whittaker-Shannon sampling theorem; all frequencies beyond the Nyquist cutoff are regarded as noise. However, the Shannon formula is noncausal and hence not directly applicable to sound reconstruction/recovery.

In addition, the high-end frequency (so-called Nyquist frequency) 22.05 kHz (half of the sampling frequency used in digital audio) may not provide sufficient margin against the audible range. Digital filters used today usually cut the frequency components beyond 20 kHz very sharply. But this has the side effect of inducing (1) a large amount of phase distortion (phase error is not considered in the conventional Shannon paradigm); and (2) ringing around 20 kHz due to the sharp-cut characteristic of the filter (Gibbs phenomenon). The latter induces a very "aggressive," sharp, and metallic sound that is likely the main reason for the audiophile's complaints about CD recordings. Undesirable distortions intrude below the Nyquist frequency too.

An Application for Sampled-Data Control Theory

High-frequency components are intersample signals. This observation suggests that modern sampled-data control theory can offer solutions to the problems of sound processing today. Based on recent theoretical results, filters can be designed that optimally interpolate the intersample content (that is, the lost high-frequency components). The optimal continuous-time (analog) performance is recovered.

Commercial Example

The figure below is an example from a mini-disc (MD) (similar format as MP3) player. The horizontal axis is the audio frequency on a linear scale of 0–22 kHz. The top graph shows the frequency response with the MDLP4 standard at 66 kbps. The bottom graph shows the response at the same bit rate with the YY filter implemented. The improved high-frequency response is evident.

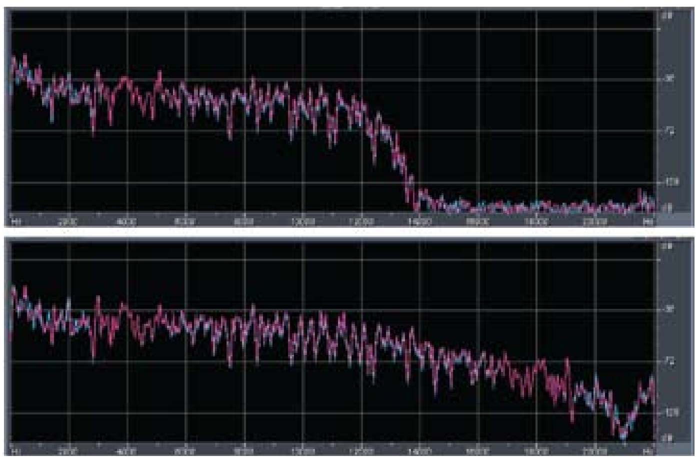
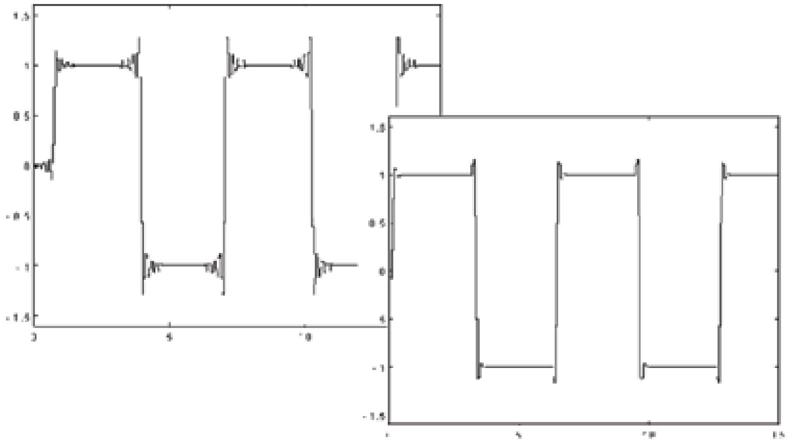


Figure courtesy of Sanyo Corporation.

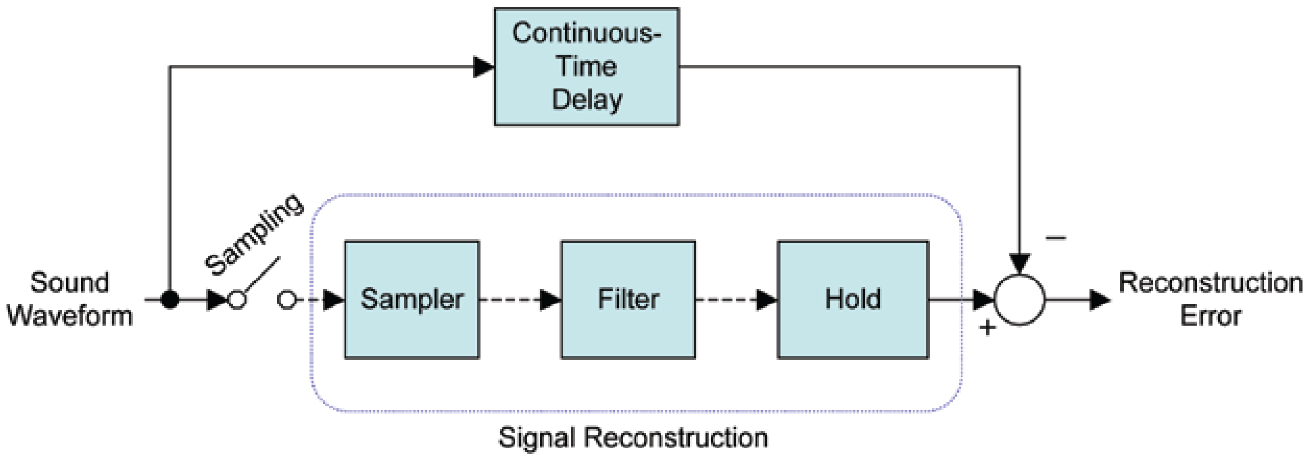
Contributor: Yutaka Yamamoto, Kyoto University, Japan

By the Millions!

The YY filter has been implemented in integrated circuits produced by SANYO Semiconductor for expanding the effective range in such devices as CDs, MP3s, mobile phones, digital voice recorders, and car audio systems. The sound quality has proven superior to the original according to the PEAQ (Perceptual Evaluation of Audio Quality) index; the filter enhances the quality by almost 30% for MP3 128 kbps and by over 30% for advanced audio coding (AAC) on average. Cumulative production has reached 16 million chips during the period 2005-2010.



The effect of the YY filter is particularly evident in these reconstructions of a sampled square wave. Conventional reconstruction results in significant high-frequency distortion (the "ringing" observed at the corners of the signal.) The YY filter substantially reduces the distortion.



The YY filter design process mathematically optimizes the filter to ensure that the reconstruction error across a desired frequency range—not determined solely by the Nyquist frequency—is less than a design parameter. This is a sampled-data H_∞ control problem.

Applications to Image Processing

The YY filter can be applied to images as well, as illustrated here. Left image: interpolation via a bicubic filter; right image: interpolation with the YY filter. Visit http://www-ics.acs.i.kyoto-u.ac.jp/~yy/sound.html for high-resolution images.

For further information: Can you hear the difference? Visit http://www-ics.acs.i.kyoto-u.ac.jp/~yy/sound.html.

Part 3

CROSS-CUTTING RESEARCH DIRECTIONS

Networked Decision Systems

Munther A. Dahleh and Michael Rinehart

Decision and Communication Networks: Overview and Challenges

A *decision network* can be broadly characterized as a distributed system of locally controlled agents whose dynamics and/or objective functions have a neighborhood structure that can be described by a graph. The decision network is supported by an underlying *communication network* that may consist of both wired and wireless networks of varying quality and whose connectivity structure need not align with the decision network topology. We refer to the combination of the two networks as a *networked decision system*. A schematic networked decision system is shown in Fig. 1.

A familiar example of a networked decision system is a formation of unmanned aerial vehicles (UAVs). Each UAV has a local controller to control its flight, but it must also follow commanded trajectories while avoiding collisions and the like. This may require information from other nearby UAVs, ground bases, or other information sources. In addition, a leader UAV may need to provide trajectory or waypoint commands to the formation. These decisions can be communicated through the formation itself (as a multihop routing network) or through other nodes. Other examples of networked decision systems include distributed emergency response systems, interconnected transportation, energy systems, and even social networks.

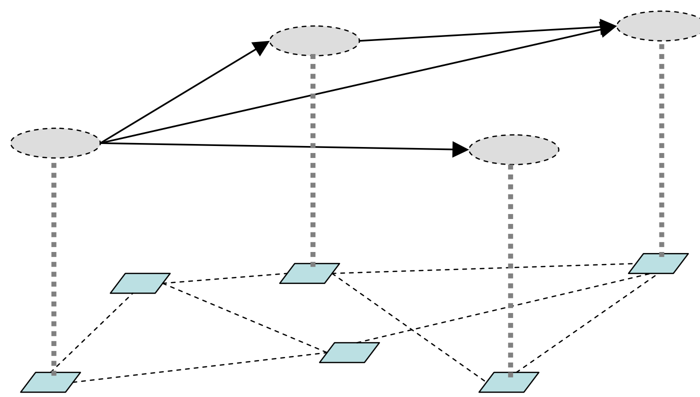


Figure 1. Illustration of a networked decision system. The upper-level nodes represent the decision network component (such as UAVs) and the lower-level nodes represent the communication network component (such as a multihop network).

Examples of networked decision systems include UAV formations, distributed emergency response systems, interconnected transportation, energy systems, and even social networks.

Networked decision systems are pervasive, and society and industry are becoming increasingly dependent on them. However, decentralized decision making over imperfect networks is fraught with difficulties. Issues and challenges are especially pronounced when dynamics are involved as the stability of the network also becomes a top priority. It is precisely these areas that the controls research community, with its history of designing robust and optimal dynamic systems, can address.

The ultimate objective of controls-related research in networked decision systems is a general analysis framework that can be used to derive fundamental performance limitations. The variety of realistic complications that such a framework must accommodate—communication delays, uncertainty in

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

information, competitive environments, limitations of communication and computational resources, learning and adaptation, mobility in agents or infrastructure nodes—points to the ambitious nature of this goal. We begin our discussion with a description of the latest research in networked control systems. Although these efforts have revealed many fundamental limitations of these systems, generalizations of them lead to a general formulation of interest. We conclude with some broad considerations related to a unified theory of networked systems.

A research objective is to characterize the fundamental limitations and capabilities of networked systems by deriving performance bounds that are functions of the underlying topologies of the networks, the capacities of the communication links, the dynamics of each node, and the computational and storage resources available to each node.

Decisions Networks: Fundamental Limits and Open Questions

Challenges in Networked Decision Systems

If one is able (willing) to assume a priori that the decision network and communication network do not interact except though an interface of constraints and requirements, the two networks can be analyzed and designed essentially independent of one another, allowing for a classical analysis of the system. In particular, the communication network can be abstracted as a set of static constraints (such as channel capacities or delays) on the operations of the decision network, whereas the decision network can be seen by the communication network as imposing requirements or preferences such as performance guarantees or utility functions. However, this assumption is rarely true in practice. For example, the decision network may take actions that disconnect the communication network or the communication network may not efficiently route critical information to portions of the decision network quickly enough, affecting performance or even stability.

Complicating matters are the dynamics of the decentralized decision network itself. Even if the underlying communication network is perfect (infinite capacity and no latency), the decision network possesses performance limitations that are missing in the centralized single decision agent. In fact, the analysis and design of distributed systems with different information patterns is still an open problem. Resource constraints that necessitate practical protocols and algorithms, and even fundamental challenges in control theory such as delays, further complicate this setting.

Control theory, information theory, optimization and game theory, and graph theory considered aspects of these applications in isolation and were able to provide basic limitations such as those captured by Bode's integral formula, Shannon's information transmission, Myerson-Satterthwaite's result on bilateral trade, and the spectral theory of graphs. However, no general analysis framework exists that is capable of addressing the interplay of these factors. In fact, the very paradigms for control and communication systems are incompatible. For example, although information theory has focused on zero-error transmission with possibly large delays, control systems tend to be very sensitive to delays while being less sensitive to static and dynamic errors—both consequences of the use of feedback. Even in the context of a single agent, the interplay between the physical space (where the agent is expected to perform) and the information space (which is described by the ability to communicate) when the agent has limited resources and when the success of communication depends on the actual dynamical behavior is still to be investigated. A network of such cooperating agents creates an even more challenging set of problems in terms of the fundamental limits. Finally, a network of autonomous agents that have possibly conflicting interests is still more difficult to analyze and coordinate.

The results presented in this section provide insight into the fundamental performance limits of decision networks. Some of these limitations arise from dynamics of the decentralized nature of the system and others from the agent's uncertainties about the system, in part due to delay and channel rate limitations.

Single Agent: The Value of Side Information in Static Decisions

The network can often be part of the overall design of a system. In complex applications such as transportation systems or the power grid, which involve humans in the loop, it is critical that only select information be communicated to the decision maker. Otherwise decisions will be delayed substantially until the decision maker sorts through the massive data sets. In short, information must be compressed and filtered so that only the information that most influences decision making is communicated. Also, because gathering, transmitting, and storing information can be costly, the minimum amount of information that is required to reach a certain level of performance should be determined.

To begin to understand the relationship between information type, information quantity, and decision making, consider a simple but prototypical problem: a single agent traversing the shortest path of a graph [1], [2]. Although only a single agent is considered, the information on which this agent bases its decisions—uncertain, intermittent, partial information about traversal delays along different edges—is analogous to the problems that would be faced if such information were being communicated by distributed sensors over an imperfect communication network. Information-theoretic bounds and other results from the single-agent scenario carry over to networked decision systems.

The standard stochastic shortest-path problem can be described briefly as follows. An agent wishes to traverse a graph along the shortest path in that graph. The delays on the edges are random (with a known distribution), and the agent may or may not know some information about the edge delays in advance of choosing a path. Now if the agent has limited resources with which to gather information about the edge delays in advance of its travel (for example, it has a limited budget for purchasing sensors), relevant questions in this context are: What types of sensors should the agent purchase, and on which edges should the agent place the sensors to best improve its overall performance? Beyond shortest-path optimization, we may more generally seek to provide a simple, intuitive framework for studying decision making under limited information conditions as well as to provide algorithms that (sub)optimally allocate information resources (such as sensors or bandwidth) to best improve the agent's performance.

In this static, centralized, and performance-centric setting, the optimal information is not characterized by mutual information quantities or bit rates, but rather as a measure of the degree to which that information is concentrated to the agent's decision subspace (termed the *actionable information*). In particular, the agent uses the information to estimate the edge delays, and the variance of this estimate in a particular subspace is the sole determinant of the agent's performance [1]. In fact, the agent's overall estimation error is irrelevant and can be arbitrarily high. Furthermore, under certain conditions, a practical scheme exists by which the agent can guarantee that the information it receives is concentrated (that is, without additional processing) to its actionable component: place all sensors to at most two paths of the graph. Generally, this scheme may contain some irrelevant information, but the performance resulting from this configuration can be shown to be acceptable.

This setting can be further generalized to a quasi-dynamic setting [2] where information is gradually revealed to the agent as it traverses the graph. In this case, the actionable information changes with each step the agent takes. If future information is reconcentrated to these subspaces, the agent's performance can be shown to further improve. However, if the information is blindly broadcast, the agent's performance can only degrade.

Designing a network with limited capacity to support decision making brings to light important research questions that generalize this framework:

- Inclusion of dynamics: The amount of actionable information determines the agent's performance. How can this notion be generalized in a non-performance-centric setting where the agent has dynamics and is concerned with stability?
- Algorithms for computing the actionable information set: The actionable information was shown to correspond to a subspace in the single-agent shortest-path problem. In more general settings with nonlinear objectives and multiple agents, the actionable information set may not have such a simple characterization. Can general techniques be developed for efficiently approximating the set of actionable information in this case?

Single Agent: Stability and Asymptotic Performance Under Communication Constraints

Understanding the fundamental limitations of performance in a feedback system is critical for effective control design. Substantial progress has been made in this direction, addressing questions of stability and performance tradeoffs in feedback systems. One of the most powerful results capturing performance tradeoffs in a stable linear feedback system is Bode's integral formula [3], which captures performance limitations in terms of the unstable modes of the plant.

In the context of centralized control under communication constraints, generalizations to this result as well as other results were obtained using information-theoretic concepts. For example, research has shown that the minimum bit rate through a discrete, error-free channel between the plant and controller that is required to stabilize a linear system is expressed purely in terms of the unstable modes of the plant [4], [5]. Furthermore, practical communication schemes can be developed that provide that base rate. A performance-centric variation of this problem is considered where the plant and controller have perfect communication but track a reference that is communicated over a channel [6]. Furthermore, the controller is to provide good model-matching performance subject to this limited reference. Research shows that there is an inherent tradeoff between communication delay and performance which forces the design of the encoder/decoder and the controller to be performed simultaneously [6].

In the two cases above, communication constraints were treated as bit-rate constraints on a discrete channel. A different representation for communication constraints is considered whereby a communication channel between the plant and controller is characterized solely by its capacity [7], [8]. A nonclassical analysis using information-theoretic quantities is used to examine the flow on entropy in the feedback loop as a means of obtaining fundamental asymptotic performance limitations. The result is a generalization of Bode's integral formula that provides conditions under which this limit can be improved by using side information.

To apply an entropy-flow analysis, properties of the controller must be characterized in terms of information-theoretic constraints. The causality of the controller and overall stability of the plant are expressed, respectively, in terms of a mutual information equality and a variance constraint [7], [8].

Generally, such abstract representations of the system allow for an asymptotic analysis that can reveal fundamental performance limits.

Although these results bridge the gap between control and communication, much remains to be explored. Following are some interesting open problems:

- Notions for information: What is the correct notion of "information" when communication supports a decision system? The notion of information captured by Shannon in point-to-point communication is not adequate in this setting. In the context of channel coding, block codes perform optimally in transmitting a message with small probability of error; however, such codes can be detrimental to a control system due to large delays.

- Tradeoff between bit rate and delay: How do we address the interplay between control and communication? The summary above assumes that the system dynamics are decoupled from the communication channel. In many situations, the bandwidth or capacity of the channel depends directly on the state of the underlying dynamic systems, such as in a mobile system where the communication depends on its actual physical location. Since the mobile system can choose to deploy itself at a particular location, the power consumed is shared with the power available for communication. Such examples where communication directly interferes with the control strategy are not very well understood.

Network of Cooperating Agents: Decentralized Computation Under Communication Constraints

We now move beyond the centralized decision maker setting to a decentralized setting, specifically decentralized decisions over unreliable networks. Examples of such networks include ad hoc wireless networks, satellite networks, and noisy social and human networks. Such networks can severely limit the capabilities of decision makers as their ability to estimate the underlying states of the systems is limited by the ability to faithfully communicate with the other agents in a timely fashion. The research objective is to characterize the fundamental limitations and capabilities of such networked systems by deriving performance bounds that are functions of the underlying topologies of the networks, the capacities of the communication links, the dynamics of each node, and the computational and storage resources available to each node.

When nodes can have unlimited computational power, research has shown that the conductance of the network graph—a measure of how "well knit" the graph is—plays a critical role in characterizing the performance of consensus-type problems where nodes are trying to compute a function of a set of initial values that are distributed over the network [9]. In particular, the time needed for each node to compute an accurate estimate of its function scales as the inverse of the conductance. For example, a ring network that communicates with neighbors with probability 1/4 scales as the inverse of the number of vertices, which implies a linear growth in convergence time for the estimates. Networks that communicate with all agents with the same probability have no bottlenecks and their conductance is constant regardless of the network's size. For example, the preferential model of the Internet has this property, which indicates that the Internet is a good medium for distributed computation. Another example of such a network is the ad hoc wireless model of Gupta-Kumar [10], which allows two wireless devices to communicate simultaneously only if they are outside a disk of a certain radius (this is often referred to as the disk model). In this case, the computation can be obtained accurately at a rate not faster than the square root of the number of vertices.

A natural generalization of this framework is one where evolving functions need to be communicated. This problem is further complicated in the realistic case of agents having dynamics. For example, if agents communicate with other agents over channels with capacities that depend on their locations and resources, the graph connecting them may change dynamically. Putting aside the agent's own dynamics, the stability of the distributed function estimation itself is put in danger as the graph changes.

Previous work has also explored some of the mechanisms for computation in the presence of varying time delays and changes in network connectivity [11], [12], but only relatively simple operations such as consensus protocols have been fully explored. Conversely, some work has been done in maintaining robust communications topologies, but without regard for the most effective utilization of network resources or the details of the desired information flow and possible effects of latency. These problems are particularly difficult in the case where local decisions are made at the network's nodes, requiring global properties to either be represented in a distributed fashion or estimated by individual nodes (including receivers and transmitters in the network).

In addition to the above, further research areas include:

- Architectural limitations on distributed problems: Consider, for example, a network where agents can only communicate their decisions (or the values of the functions they are computing). In this context, we think of these functions as utilities. Communicating utilities gives only aggregate information about the underlying state of the system and imposes severe limitations on the ability to learn the state. How can these limitations be characterized?

- Robustness: In this regard, it may be beneficial to search for the right topology (or metric) on the set of graphs that is amenable to perturbation analysis. Under what perturbation conditions is asymptotic estimation possible?

Network of Competitive Agents: Information Aggregation and Asymptotic Learning

Social networks are attracting substantial attention within the research community. In particular, a tremendous opportunity exists for bringing in quantitative tools to analyze the formation of such networks as well as to study the impact of such networks on decision making. What differentiates such networks from standard decentralized networks is the human presence. A question that arises in the investigation of networks with human actors is how game-theoretic interactions modify the well-known existing results on dynamic aggregation of decentralized information over networks with non-autonomous agents (for example, see the literature on consensus [13]-[17]).

Results have been reported that begin to address this framework [19]. They show that when selfish agents are sequentially detecting an underlying binary state of the world, information may not aggregate properly. The loss of collective wisdom is due to the "herding" phenomenon often witnessed in technology and fads. A realistic framework for learning in a multi-agent system must model the structure of social networks with which individuals observe and communicate with each other; however, such generalizations turn out to be challenging to analyze. One difficulty with this class of models is that to determine how beliefs will evolve, we need to characterize the perfect Bayesian Nash equilibrium, which involves rather complex inferences by individuals. To this end, we will consider a simplified model where the agents observe the actions of a neighborhood of individuals that are randomly chosen from the entire set of past actions of this neighborhood. Although actual social leaning can involve very complicated dynamics not captured in this simplified model, it does provide a first-order approximation for which definitive statements can be made, and the fundamental limitations of this model may hold in more complicated models.

More recent work addressed this problem and established exact conditions under which herding is impossible [20]. In this work, these effects are captured in terms of characteristics of the graph's interconnectedness and the properties of the underlying random process. In particular, under certain conditions, an excessively influential group can emerge within a social network if interconnectedness among individuals is not rich enough.

These results consider an idealized situation where all agents have the same utility and where there is an absence of disruption. Furthermore, they only address asymptotic learning as the size of the network increases. Hence, several interesting research directions in this field have not been pursued or have provided only partial results:

- Sequential decisions and feedback: Analysis was simplified by allowing agents to fix a decision once it is made, but repeated decision making better reflects real-world dynamics. How do repeated decisions and endogenous sequencing of actions affect asymptotic learning over time?
- Perturbations: The influence of external effects (such as media, injecting outside agents, changing the network topology) on the propagation of beliefs is relevant because such outside effects can serve as either control inputs or as adversarial influences on the system. For example, what types of networks allow a "reversal" in the beliefs of individuals?
- Forward-thinking agents: In social learning models studied to date, individuals care only about their immediate payoffs. What general approach should one take toward analyzing the perfect Bayesian Nash equilibrium in the case where agents' payoffs depend on the future decisions of other agents?

Broad Considerations in Decision and Communication Networks

The natural generalizations considered for each of the previous works seem to quickly lead to common problems of high importance. Although the specifics of these problems still vary (each has different objectives and algorithms), a general analysis framework could be established that can be used to derive fundamental performance limitations. Below we discuss several research areas that may be helpful in developing such a framework.

- Network separation principle: The separation principle from classical control theory offers conditions under which a feedback control signal cannot improve the controller's estimate of the plant's state. When it applies, the optimal performance of the system can be directly analyzed by constructing an optimal estimator and controller. However, if these conditions are not met, even a simple feedback system can have a complex optimal controller. Under what conditions are system uncertainties independent of the agents' decisions? A "degree of separation" may be useful in establishing approximate results in this challenging area.
- Dynamic notions for actionable information: In learning and centralized-feedback control, an entropy-flow analysis was used to study the dynamic exchange of information between agents. The results were algorithmically free, asymptotic fundamental limits for performance. However, in the performance-centric setting of shortest-path optimization, it was the amount of information concentrated to the actionable subspace of decisions that affected the quality of

decision making. In a dynamic setting, does there exist a similar set to which information should be concentrated and that changes over time? How should agents track it? The flow of actionable information content over the network may yield tighter, more useful fundamental limits than entropy flow alone.

- Representations for abstract computation: An effort to link decision making to information flow may require developing a representation for algorithms in the language of information flow. In the case of a centralized-controller feedback system, causal, stabilizing controllers can be represented by imposing information-theoretic constraints on the feedback system. Can such formulations be extended to decentralized and nonlinear settings? Additional constraints that may be useful to develop are those that capture limited computational capability.

- Representations for communication: The notion of information captured by Shannon in point-to-point communication is not adequate for analysis. As noted earlier, although block codes perform optimally in the context of transmitting a message with small probability of error, such codes can be detrimental to a control system due to large delays. How can we efficiently represent causality across a network with many information flows? Notions of mutual information and information rates do not completely capture the interactions of multiple causal dependences.

- Robustness to perturbations: Perturbations in network topology, computation, or communication may propagate errors throughout the network that can degrade performance or, worse, result in positive feedback loops in the system that may amplify the effect of the errors, destabilizing the system. To illustrate the types of perturbations that need to be specially considered in dynamic agent networks, consider the case where the interaction between the agent's dynamics and the graph are carefully designed, but a time-varying perturbation in the graph results in a transient cycle in information flow. If the network is a Bayesian learning network, these cycles may destabilize learning. Even in the simplest case where the dynamics of the nodes can be modeled as linear input/output systems (including time delays), the static graph structure is known to be crucial for determining its overall stability [21].

Selected recommendations for research in networked decision systems:

- What is the correct notion of "information" when communication supports a decision system, and how do we address the interplay between control and communication? Fundamental problems of analysis and design in cases where communication directly interacts with the control strategy need to be investigated.
- Our understanding of the fundamental limitations and capabilities of decentralized networked systems under uncertainties is incomplete. Performance bounds that are functions of the underlying topologies of the networks, the capacities of the communication links, the dynamics of each node, and the computational and storage resources available to each node would be useful for many applications.
- Connections with game theory are an important research area, with several open problems. For example, how do repeated decisions and endogenous sequencing of actions affect asymptotic learning over time in a game-theoretic network of competitive agents?

References

[1] M. Rinehart and M.A. Dahleh. "The value of side information in shortest path optimization," to appear in *IEEE Transactions on Automatic Control,* 2010.

[2] M. Rinehart and M.A. Dahleh. "The value of sequential information in shortest path optimization," submitted to *Automatica,* 2010.

[3] H.W. Bode. *Network Analysis and Feedback Amplifier Design.* Princeton, NJ: D. Van Nostrand, 1945.

[4] S. Tatikonda and S. Mitter. "Control over noisy channels," *IEEE Transactions on Automatic Control,* vol. 49, pp. 1196-1201, July 2004.

[5] S. Tatikonda and S. Mitter. "Control under communication constraints," *IEEE Transactions on Automatic Control,* vol. 49, pp. 1056-1068, July 2004.

[6] S. Sarma and M.A. Dahleh. "Remote control over noisy communication channels: A first-order example," *IEEE Transactions on Automatic Control,* vol. 52, pp. 284-289, February 2007.

[7] N.C. Martins and M.A. Dahleh. "Feedback control in the presence of noisy channels: Bode-like fundamental limitation of performance," *IEEE Transactions on Automatic Control*, vol. 53, pp.1604-1615, August 2008.

[8] N.C. Martins, M.A. Dahleh, and J.C. Doyle. "Fundamental limitations of disturbance attenuation with side information," *IEEE Transactions on Automatic Control,* vol. 52, pp.56-66, January 2007.

[9] O. Ayaso, D. Shah, and M.A. Dahleh. "Information theoretic bounds for distributed computation," submitted to *IEEE Transactions on Information Theory,* 2008.

[10] P. Gupta and P.R. Kumar. "The capacity of wireless networks," *IEEE Transactions on Information Theory,* vol. 46, pp. 308-404, March 2000.

[11] A. Jadbabaie, J. Lin, and A.S. Morse. "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control,* vol. 48, no. 6, pp. 988-1001, June 2003.

[12] R. Olfati-Saber and R. M. Murray. "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control,* vol. 49, pp. 1520-1533, September 2004.

[13] J.N. Tsitsiklis. "Problems in Decentralized Decision Making and Computation," PhD thesis, Massachusetts Institute of Technology, 1984.

[14] J.N. Tsitsiklis, D.P. Bertsekas, and M. Athans. "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control,* vol. 31, no. 9, pp. 803-812, September 1986.

[16] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. "Gossip algorithms: Design, analysis, and applications," in Proceedings of IEEE INFOCOM, vol. 3, pp. 1653-1644, March 2005.

[17] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis. "Convergence in multiagent coordination, consensus, and flocking," in Proceedings of IEEE CDC, pp. 2996-3000, December 2005.

[18] A. Nedić and A. Ozdaglar. "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control,* vol. 54, pp. 48-61, January 2009.

[19] A. Nedić, A. Olshevsky, A. Ozdaglar, and J.N. Tsitsiklis. "On distributed averaging algorithms and quantization effects," *IEEE Transactions on Automatic Control,* vol. 54, pp. 2506-2517, November 2009.

[20] S. Bikchandani, D. Hirshleifer, and I. Welch. “A theory of fads, fashion, custom, and cultural change as information cascades,” *Journal of Political Economy,* vol. 100, pp. 992-1026, 1992.

[21] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar. Bayesian Learning in Social Networks, LIDS Technical Report 2779, 2008.

[22] J.A. Fax and R.M. Murray. “Information flow and cooperative control of vehicle formations,” *IEEE Transactions on Automatic Control,* vol. 49, no. 5, pp. 1465-1476, September 2004.

Cyber-physical Systems

Radhakisan Baheti and Helen Gill

Introduction

The term *cyber-physical systems* (CPS) refers to a new generation of systems with integrated computational and physical capabilities that can interact with humans through many new modalities. The ability to interact with, and expand the capabilities of, the physical world through computation, communication, and control is a key enabler for future technology developments. Opportunities and research challenges include the design and development of next-generation airplanes and space vehicles, hybrid gas-electric vehicles, fully autonomous urban driving, and prostheses that allow brain signals to control physical objects.

Over the years, systems and control researchers have pioneered the development of powerful system science and engineering methods and tools, such as time and frequency domain methods, state space analysis, system identification, filtering, prediction, optimization, robust control, and stochastic control. At the same time, computer science researchers have made major breakthroughs in new programming languages, real-time computing techniques, visualization methods, compiler designs, embedded systems architectures and systems software, and innovative approaches to ensure computer system reliability, cyber security, and fault tolerance. Computer science researchers have also developed a variety of powerful modeling formalisms and verification tools. Cyber-physical systems research aims to integrate knowledge and engineering principles across the computational and engineering disciplines (networking, control, software, human interaction, learning theory, as well as electrical, mechanical, chemical, biomedical, material science, and other engineering disciplines) to develop new CPS science and supporting technology.

In industrial practice, many engineering systems have been designed by decoupling the control system design from the hardware/software implementation details. After the control system is designed and verified by extensive simulation, ad hoc tuning methods have been used to address modeling uncertainty and random disturbances. However, the integration of various subsystems, while keeping the system functional and operational, has been time-consuming and costly. For example, in the automotive industry, a vehicle control system relies on system components manufactured by different vendors with their own software and hardware. A major challenge for original equipment manufacturers (OEMs) that provide parts to a supply chain is to hold down costs by developing components that can be integrated into different vehicles.

The increasing complexity of components and the use of more advanced technologies for sensors and actuators, wireless communication, and multicore processors pose a major challenge for building next-generation vehicle control systems. Both the supplier and integrator need new systems science that enables reliable and cost-effective integration of independently developed system components. In particular, theory and tools are needed for developing cost-effective methods to: (1) design, analyze, and verify components at various levels of abstraction, including the system and software architecture levels, subject to constraints from other levels; (2) analyze and understand interactions between the vehicle control systems and other subsystems (engine, transmission, steering, wheel, brake, and suspension); and (3) ensure safety, stability, and performance while minimizing vehicle cost to the

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

consumer. Increasingly, new functionality and the cost of vehicle control systems are major differentiating factors for business viability in automobile manufacturing.

Need for CPS Research

CPS research is still in its infancy. Professional and institutional barrier have resulted in narrowly defined, discipline-specific research and education venues in academia for the science and engineering disciplines. Research is partitioned into isolated subdisciplines such as sensors, communications and networking, control theory, mathematics, software engineering, and computer science. For example, systems are designed and analyzed using a variety of modeling formalisms and tools. Each representation highlights certain features and disregards others to make analysis tractable. Typically, a particular formalism represents either the cyber or the physical process well, but not both. Whereas differential equations are used for modeling physical processes, frameworks such as Petri nets and automata are used to represent discrete behavior and control flows. Workforce expertise is similarly partitioned, to the detriment of productivity, safety, and efficiency. Although this approach to modeling and formalisms may suffice to support a component-based "divide and conquer" approach to CPS development, it poses a serious problem for verifying the overall correctness and safety of designs at the system level and component-to-component physical and behavioral interactions [1]. In the following paragraphs, research needs in CPS are briefly discussed.

> *Although the diversity of models and formalisms supports a component-based "divide and conquer" approach to CPS development, it poses a serious problem for verifying the overall correctness and safety of designs at the system level.*

Abstraction and Architectures

Innovative approaches to abstraction and architectures that enable seamless integration of control, communication, and computation must be developed for rapid design and deployment of CPS. For example, in communication networks, interfaces have been standardized between different layers. Once these interfaces have been established, the modularity allows specialized developments in each layer. The overall design allows heterogeneous systems to be composed in plug-and-play fashion, opening opportunities for innovation and massive proliferation of technology and the development of the Internet. However, the existing science and engineering base does not support routine, efficient, robust, modular design and development of CPS. Standardized abstractions and architectures are urgently needed to fully support integration and interoperability and spur similar innovations in cyber-physical systems [2].

Distributed Computations and Networked Control

The design and implementation of networked control systems pose several challenges related to time- and event-driven computing, software, variable time delays, failures, reconfiguration, and distributed decision support systems. Protocol design for real-time quality-of-service guarantees over wireless networks, tradeoffs between control law design and real-time-implementation complexity, bridging the gap between continuous and discrete-time systems, and robustness of large systems are some of the challenges for CPS research. Frameworks, algorithms, methods, and tools are needed to satisfy the high reliability and security requirements for heterogeneous cooperating components that interact through a complex, coupled physical environment operating over many spatial and temporal scales.

Verification and Validation

Hardware and software components, middleware, and operating systems need to be developed that go beyond existing technologies. The hardware and software must be highly dependable, reconfigurable, and, where required, certifiable, from components to fully integrated systems. Such complex systems must possess a trustworthiness that is lacking in many of today's cyber infrastructures. For example, certification is estimated to consume more than 50% of the resources required to develop new, safety-critical systems in the aviation industry. Similar efforts are needed in the medical, automotive, energy systems, and other application domains. Overdesign is currently the only path to safe and successful system certification and deployment. Yet this approach is rapidly becoming intractable for complex designs and for systems where interoperability is needed. Testing "until the money runs out" is not a viable strategy, and science- and evidence-based methods are needed for reasoning about system reliability. New models, algorithms, methods, and tools are needed that will incorporate verification and validation of software and systems at the control design stage.

Challenges and Opportunities: Industry-Academia

Advances in CPS research can be accelerated by indentifying needs, challenges, and opportunities in several industrial sectors and by encouraging multidisciplinary collaborative research between academia and industry. The objective is to develop new systems science and engineering methods for building high-confidence systems in which cyber and physical designs are compatible, synergistic, and integrated at all scales. Current and past industry investments in CPS technology research have been significant but focused on shorter-term, quicker-payoff proprietary technologies. Recently, governments and some industry sectors are investing in longer-term, precompetitive technologies and innovative testbeds. For example, the European Union has initiated a major joint technology initiative with public-private funding by European nations and industry called Advanced Research and Technology for Embedded Intelligence Systems (ARTEMIS). Similarly, based on recommendations in the August 2007 report of the U.S. President's Council of Advisors on Science and Technology (PCAST), the U.S. National Science Foundation has been funding fundamental CPS research and education [3]. Related initiatives are being pursued in other countries, including Japan, China, South Korea, and Germany.

CPS grand challenges are being articulated in many industry sectors. The U.S. National Academy of Engineering has listed 14 grand challenges that relate environmental, health, and societal issues; these issues will clearly benefit from advances achieved in cyber-physical systems. The control engineering research community can play a leading role in the development of cyber-physical systems. Some of the opportunities are described below.

> *The U.S. National Academy of Engineers has listed 14 grand challenges that relate environmental, health, and societal issues; these issues will clearly benefit from advances achieved in cyber-physical systems.*

Biomedical and Healthcare Systems

CPS research is revealing numerous opportunities and challenges in medicine and biomedical engineering. These include intelligent operating rooms and hospitals, image-guided surgery and therapy, fluid flow control for medicine and biological assays, and the development of physical and neural prostheses. Healthcare increasingly relies on medical devices and systems that are networked and need to match the needs of patients with special circumstances. Thus, medical devices and systems will be needed that are dynamically reconfigured, distributed, and can interact with patients and caregivers in complex environments. For example, devices such as infusion pumps for sedation, ventilators and oxygen delivery systems for respiration support, and a variety of sensors for monitoring patient condition are used in many operating rooms. Often, these devices must be assembled into a new system configuration to match specific patient or procedural needs. The challenge is to develop systems and control methodologies for designing and operating these systems that are certifiable, safe, secure, and reliable.

Research challenges in medical technology and healthcare were considered in a series of workshops that are summarized in a U.S. National Information Technology Research and Development (NITRD) report [4]. The report recommends research for new system science and engineering with the following goals:

- Interoperable and open medical systems;
- Distributed monitoring, distributed control, and real-time wireless networks for hospital intensive-care facilities;
- Certification methods for medical device software and systems and networked patient monitoring and assistance;
- Model-based frameworks that support component-based modeling, design, testing, and certification using patient-specific models.

Another challenging area for CPS research is cognition and neuroscience for understanding the fundamental principles of human motor functions and exploiting this understanding in engineered systems. Examples include brain-machine interfaces, therapeutic and entertainment robotics, orthotics and exoskeletons, and prosthetics. Humans and animals seamlessly integrate sensing, computing, and motor control functions. These highly coupled systems do not satisfy simple modularity principles, but are composed of multifunctional elements, computation, and feedback loops at different time and length scales, noisy signals, parallel processing, and redundant fault-tolerant architectures. Recent research has suggested that animals use some form of optimal filtering, stochastic control algorithms, and large-scale probabilistic computing structures in dealing with uncertainty. Control researchers working with biologists, neurophysiologists, and computer scientists may be able to make further progress.

Next-Generation Air Transportation Systems (NextGen)

Cyber-physical systems research is likely to have an impact on the design of future aircraft and air traffic management systems, as well as on aviation safety. Specific research areas include (1) new functionality to achieve higher capacity, greater safety, and more efficiency, as well as the interplay and tradeoffs among these performance goals; (2) integrated flight deck systems, moving from displays and concepts for pilots to future (semi)autonomous systems; (3) vehicle health monitoring and vehicle health management; and (4) safety research relative to aircraft control systems. One of the key technical challenges to realizing NextGen involves verification and validation of complex flight-critical systems with a focus on promoting reliable, secure, and safe use for NextGen operations. As the complexity of systems increases, costs related to verification and validation and safety assurance will likely increase the cost of designing and building next-generation vehicles. The broader aeronautics community has identified verification and validation methodologies and concepts as a critical research area [5].

The goals of research in verification and validation of aviation flight-critical systems include providing methods for rigorous and systematic high-level validation of system safety properties and requirements, from initial design through implementation, maintenance, and modification, as well as understanding tradeoffs between complexity and verification methods for supporting robustness and fault tolerance. Some of the control engineering challenges include:

- Large-scale, real-time, deterministic robust or stochastic optimization algorithms;
- Multiple-objective, multiple-stakeholder optimization frameworks;
- Design of automation with graceful degradation modes;
- Safety diagnosis/health monitoring methods;
- System architectures that facilitate distributed decision making;
- Data fusion from heterogeneous sensors and assessment of the value of the derived information.

Smart Grid and Renewable Energy

Smart grid and renewable energy research and development has been in the forefront of public interest and is therefore a high priority for policy makers. The goal is to improve energy efficiency by investing in modernization of the energy infrastructure. The geopolitical drivers for renewable energy and smart grids are that (1) electricity demand is expected to increase more than 75% by 2030; (2) generation of electricity contributes to more than 40% of greenhouse gas emissions; (3) the cost of generating 1 KWh is four times greater than the cost of saving 1 KWh. Government funding agencies have partnered with industry, utilities, and local government in technology development and demonstration projects dealing with smart grids. For example, “Energy Smart Florida” is a groundbreaking public/private alliance of the City of Miami, Florida Power and Light, General Electric, Silver Spring Networks, and Cisco. This project is using federal economic stimulus funds as part of an $800 million investment in smart grid technology and renewable energy over the next two years. An estimated 4.5 million smart meters will be installed in U.S. homes and businesses to develop and demonstrate technology for demand management, distribution automation, substation intelligence, distributed generation, and information technology. The goal is to demonstrate the increase in energy efficiency through demand optimization and distributed automation by significantly reducing peak load.

Advances in flexible AC transmission devices (FACTS) and phasor measurement units (PMUs) have opened new opportunities for wide-area control of smart grids. The U.S. Department of Energy-sponsored North American SynchroPhasor Initiative (NASPI) has been heavily investing in PMU hardware. Future efforts will be needed to focus on data fusion and analysis for real-time dynamics monitoring, prediction, and system control. With increased reliance on wide-area communications and control to improve system operation, tight coupling is needed between cyber systems and the components of physical systems in smart grids. Critical gaps and shared challenges pertain to advances in system science, particularly hybrid digital-analog systems, complex emergent systems, and advanced software systems for large-scale, time-varying, geographically distributed systems. Advances in optimization of multiscale stochastic dynamic systems as well as in distributed control are necessary to improve smart grid performance with respect to security, efficiency, reliability, and economics. These issues have been identified by the research community in several workshops dealing with information technology and smart grids.

Conclusions

Cyber-physical systems are expected to play a major role in the design and development of future engineering systems with new capabilities that far exceed today's levels of autonomy, functionality, usability, reliability, and cyber security. Advances in CPS research can be accelerated by close collaborations between academic disciplines in computation, communication, control, and other engineering and computer science disciplines, coupled with grand challenge applications.

Selected recommendations for research in cyber-physical systems:

- Standardized abstractions and architectures that permit modular design and development of cyber-physical systems are urgently needed.
- CPS applications involve components that interact through a complex, coupled physical environment. Reliability and security pose particular challenges in this context—new frameworks, algorithms, and tools are required.
- Future cyber-physical systems will require hardware and software components that are highly dependable, reconfigurable, and in many applications, certifiable . . . and trust-worthiness must also extend to the system level.

References

[1] A. Rajhans, S.W. Cheng, B. Schmerl, B.H. Krogh, C. Aghi, and A. Bhave, "An Architectural Approach to the Design and Analysis of Cyber-Physical Systems," Third International Workshop on Multi-Paradigm Modeling, Denver, CO, October 2009.

[2] S. Graham, G. Baliga, and P.R. Kumar, "Abstractions, Architecture, Mechanism, and Middleware for Networked Contro,l" *IEEE Transactions on Automatic Control,* vol. 54, no. 7, pp. 1490-1503, July 2009.

[3] President's Council of Advisors on Science and Technology. Leadership Under Challenge: Information Technology R&D in a Competitive World [Online], Aug. 2007. Available at http://www.nitrd.gov/Pcast/reports/PCAST-NIT-FINAL.pdf.

[4] Networking and Information Technology Research and Development Program. High-Confidence Medical Devices: Cyber-Physical Systems for 21st Century Health Care [Online], Feb. 2009. Available at http://www.nitrd.gov/About/MedDevice-FINAL1-web.pdf

[5] National Science and Technology Council. National Aeronautics Research and Development Plan [Online], Feb. 2010. Available at http://www.whitehouse.gov/sites/default/files/microsites/ostp/aero-rdplan-2010.pdf.

Cognitive Control

Martin Buss, Sandra Hirche, and Tariq Samad

Introduction

As the field of control engineering has evolved, its horizons have continually broadened. From regulation with simple proportional-integral-derivative (PID) loops, to model-based control and multivariable schemes, to explicit incorporation of uncertainty in robust and modern control theory, to hybrid and hierarchical architectures, and most recently, to control of and via networks, both theoretical foundations and application scope have seen dramatic advancement.

What's next, we might wonder? In this section we outline one prospective answer: "cognitive control." We believe that the incorporation of properties we usually associate with cognition—including reasoning, planning, and learning—within control systems holds the promise of greatly expanding the scope and impact of the field.

The scope and impact of control systems could be substantially increased with the incorporation of properties we usually associate with cognition, such as reasoning, planning, and learning.

We consider cognitive control to be an enabler for novel technologies in many diverse application areas. Field robotics, space and sea exploration systems, and next-generation unmanned aerial vehicles will achieve a higher degree of autonomy through cognitive function. Cognitive control systems for manufacturing plants will be partners to plant operators and engineers; less human intervention will be necessary even as the safety and performance of plants improve. Similar benefits can also be expected from cognitive systems assisting or ultimately replacing human operators in supervisory control applications (for example, in power generation/distribution, traffic control, and similar infrastructure-oriented domains). Search and rescue missions, especially in environments that are remote or inhospitable for humans, will also be an important application domain. Assistive technologies for the elderly are another target, and an increasingly important one given aging populations in many developed countries—cognitive control systems can help overcome both physical and cognitive impairments by enabling the elderly and infirm to live independently as well as by assisting human health workers in caring for them.

The behaviors, functions, and features required of envisioned cognitive control systems have always been part of the vision of control engineering—as articulated in motivating research in areas such as adaptive, robust, and intelligent control. This vision, however, is not much in evidence in the conferences and journals in the field. Specific research in control has focused on narrower—and better defined—problem formulations. Yet the relevance of control methodologies and tools to the broader vision is not in question. The rigor and "systems" orientation of control will be instrumental for realizing cognitive control systems in practice, and by virtue of both its intellectual depth and its record of success across all engineering fields, the controls community is ideally positioned to spearhead the development of cognitive control systems.

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Below we first discuss what motivates cognitive control as a research field. We then explain in broad terms what we mean by cognitive control. Related work in other fields is outlined, and we highlight the crucial role of control science and engineering. We conclude with discussion of some challenge problems and associated research questions for cognitive control.

Motivation: Why Cognitive Control?

Current automated systems function well in environments they are designed for, that is, around their nominal operating conditions. They also function well in environments with "predictable" uncertainties as treated, for example, in the advanced adaptive and robust control frameworks—and as demonstrated in modern engineering systems such as unmanned aerial vehicles (UAVs) and process plants without on-site operators. However, control systems of today require substantial human intervention when faced with novel and unanticipated situations—situations that have not been considered at the controller's design stage. Such situations can arise from discrete changes in the environment, extreme disturbances, structural changes in the system (for example, as a result of damage), and the like. To illustrate, future autonomous robots in search and rescue operations, in mining, in the service domain, and in autonomous driving will regularly encounter novel situations that require perception, reasoning, decision making, fact generalization, and learning. Such cognitive control aspects will play a major role in future automated and autonomous systems and will advance "automation" to the next level.

But fully autonomous systems represent just one direction for cognitive control research. Today's control systems for applications such as aircraft, chemical factories, and building systems automate many operational functions while simultaneously aiding human operators in doing their jobs. More cognitive abilities in such control systems will enable safer and higher performance semiautonomous engineering systems.

This human-automation interaction aspect suggests another important focus for cognitive control: social and group environments. Multiagent coordination and control, cooperative execution of complex tasks, effective operation in competitive or mixed competitive-cooperative situations all require the participating agents, whether human or machine, to have cognitive capabilities. In this context, communication takes on added importance and complexity. Agents will need linguistic sophistication. Shared semantic models and ontologies will be necessary. Beyond semantics, just as people rely on pragmatics in their use of language—much of what we convey through speech or writing is not directly related to the literal meaning of our utterances—so will cognitive control systems.

Definition/Description of the Topic: What Is Cognitive Control?

Attempting to define the notions of "cognition" and "cognitive system" is a controversial endeavor, as shown dramatically by the 40-plus diverse definitions of cognition that were collected within the "euCognition" project funded by the European Commission [1]. Rather than attempt a necessary and sufficient definition, we describe several fundamental ingredients of cognitive control, without any claim of completeness.

A system under cognitive control

- exhibits goal-oriented behavior in sensing, reasoning, and action;
- flexibly changes its goals and behavior depending on situational context and experience;

- is able to act in unstructured environments without human intervention and robustly responds to surprise; and
- is able to interact with humans and other cognitive systems to jointly solve a complex task.

To achieve these properties, a system under cognitive control needs to

- understand the present situation (including awareness of itself, its environment, and other agents)—to this end, the cognitive control system must implement several functions, such as (active) sensing, the extraction and abstraction of relevant information, acquisition of semantic knowledge, comparison with previous experience, and knowledge updating;
- purposefully act to modify the current situation and react to any unpredicted changes in a reasonable (not necessarily optimal) way—components required include decision making, planning, reasoning, learning, and adaptation.

An important characteristic is that full information is rarely available to construct models. Hence, the mechanisms for estimating the current state as well as for purposeful modification of this state need to operate on partial/uncertain information.

In Fig. 1, a cognitive control system architecture is proposed showing the possible components of the system:

- Perception includes the acquisition of low-level sensor data, data fusion, information processing and abstraction, and the interpretation of the information for decision making. The question is, how can important (that is, task-relevant information) be reliably filtered from the vast amount of noisy and incomplete data. Major challenges are the inclusion of contextual/semantic knowledge for more robust signal processing and interpretation and the development of active (multi-modal) sensing and signal processing strategies.

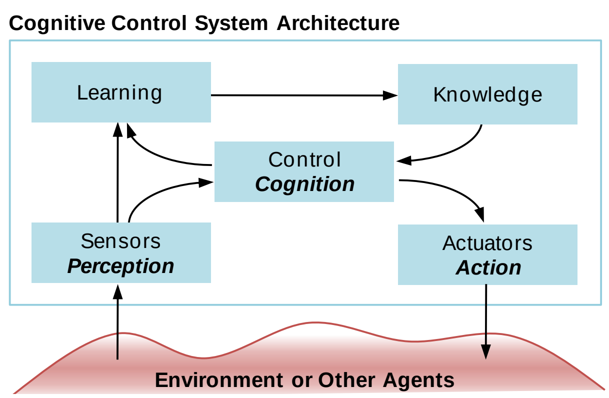


Figure 1. The perception-cognition-action loop—a proposal for a cognitive control system architecture (CoTeSys®) [2].

- Control maps percepts onto actions using existing knowledge/experience. One of the major challenges is to combine semantics with continuous and discrete signal-based representations and to produce a reasonable control decision in the presence of incomplete and/or uncertain information.
- Actions implement the output of the control element, thereby affecting the external environment of the cognitive control system. Both symbolic and continuous actions may be required, similar to the structure of the control output.
- Learning is essential to updating existing knowledge, resulting in the online adaptation of cognitive functionalities to changing environmental situations and contexts. Learning under

partial/incomplete information, hierarchical learning, and learning of symbolic temporal sequences, relations, and concepts present some of the major challenges in the area.

- Knowledge or memory/experience represents a fundamental feature of cognitive control systems. In contrast to classical approaches, this knowledge is continuously updated and modulates the task execution at runtime. An important aspect is the representational formalism for knowledge, such as the choice of representational primitives, compositions, and structure.

One limitation of Fig. 1 is that it does not show interagent interactions separately from the inputs and outputs associated with the environment. At some level of abstraction, other agents and the environment are both part of the external world of an agent, but an agent's ways of engaging will be very different with both. These differences need to be explicitly addressed in a more complete architectural design.

Relevant Neighboring Disciplines and Rationale for a Leadership Role for Controls

The area of engineered cognitive systems has so far been dominated by the artificial intelligence (AI) and computer science communities. These disciplines, together with areas of neuroscience, cognitive science, and psychology, represent the most relevant neighboring disciplines. Their contributions so far and their role in cognitive control are highlighted below. In addition, the contributions of operations research, embedded real-time systems, signal processing, and pattern recognition have been helping to advance the field and are expected to continue to do so in the future.

Artificial Intelligence and Computer Science

Within artificial intelligence and computer science research, advanced methods for reasoning, planning, decision making, and learning have been investigated over the past decades and successfully applied in information-based systems. However, their impact on systems interacting with the physical world has been limited. Such "cyber-physical systems" certainly require a deep understanding of dynamical systems (including hybrid systems that combine continuous and discrete dynamics) and feedback loops, concepts that are fundamental to control. Accordingly, existing theories need to be reformulated to include dynamical system properties. Relevant topics from AI for the area of cognitive control include

- theories of reasoning under uncertainty, sequential logic reasoning, rule-based systems, and inference machines;

- knowledge representation, reasoning about knowledge, and use of prior knowledge;

- machine learning, probabilistic learning methods, reinforcement learning, and statistical learning.

The state of the art is regularly demonstrated in benchmarking competitions such as the DARPA Grand Challenge (2005), Urban Challenge (2007), Grand Cooperative Driving Challenge (2011), and the RoboCup (yearly since 1997). (Control technologists have also been involved in, and in several cases have successfully led, entries in these competitions.)

Neuroscience, Cognitive Science, and Psychology

Neuroscience, cognitive science, and psychology can stimulate research in cognitive control by providing insights on fundamental mechanisms of natural (biological) cognition. Progress in technology for measuring brain activity has provided and will continue to provide results that are useful for engineering purposes concerning the function and architecture of the brain and their relationship to human behavior. These results are relevant for the area of cognitive control from two standpoints:

- Natural cognition as a role model for artificial cognition: The understanding of the principal mechanisms of decision making, learning, abstraction, and other functions may guide the development of artificial cognition.

- Joint human-machine cognition: To design machines with cognitive functionalities that help humans perform their tasks efficiently, the mechanisms of human perception, decision making, and action, as well as their fundamental limits, must be clearly understood. A major challenge is to obtain quantitative dynamical models suitable for cognitive control design.

The Role for Control

Given the contributions of the neighboring disciplines, what are the envisaged contributions of the controls community? As mentioned above, a fundamental ingredient of a cognitive system is goal-oriented behavior in unstructured environments. This is hardly a novel concept for control—achieving goal-oriented behavior is the basis of almost all control designs! Furthermore, control technology includes efficient and effective methods for addressing issues such as stability, optimality, and robustness. Formulations and solutions for modeling and control for uncertain, stochastic, and hybrid dynamical systems have been developed.

The controls community can contribute greatly to the area of cognitive control by using its strengths in the understanding of dynamical systems, advanced modeling concepts, feedback system analysis methods, and control synthesis tools. The methodical, system-oriented approaches and the mathematical rigor of control methods will be required for deriving provably correct results and for ensuring the safety and performance of engineering products. Even the critical importance of properties such as stability, controllability, and robustness are best appreciated, and the realization of these properties best assured, by experts in control. Without the rigor and analysis that are hallmarks of control science, we cannot expect to develop reliable, high-confidence cognitive control systems for complex applications. These arguments not only justify a role for control in cognitive control; they suggest that the controls community should adopt a leadership role in shaping the research agenda.

These arguments suggest that the controls community should take a leadership role in shaping the cognitive control research agenda.

Challenge Problems for the Field

To provide a better and more specific sense of how a cognitive control system might bring novel capabilities to automation technology, and of the multidisciplinary aspects of such a system, we outline two broad challenge problems below.

Adaptive Management of Cognitive Resources in Real-Time Systems

In today's complex automation systems, human operators play the crucial roles of aggregating and consolidating information, balancing long-term and immediate priorities, and shifting attention dynamically as circumstances dictate. Such capabilities are especially important in large-scale systems, where hundreds, thousands, or more sensors and actuators must be managed. Examples include building automation, manufacturing or process control systems, and traffic management, but an everyday example can help make the point. We are all able to drive a car on a highway while carrying on a conversation with a passenger and listening off and on to the car radio. In the background, we know the route we are taking and effect appropriate actions. However, if another car suddenly cuts in front of us or some other emergency event occurs, we immediately divert our attention to focus on the urgent need of ensuring safety. Our cognitive resources are rescheduled flexibly and at a moment's notice. This flexible, robust behavior is in contrast to the scheduling of tasks in today's computational real-time systems, which is typically static and predefined.

The difference between biological cognition and computer-based attention management becomes more pronounced as the scale of the system under control increases. Learning becomes increasingly important with problem scale. Human operators learn over time what information is important to attend to and what (huge amount of) other information can be safely ignored. The performance improvement, in terms of the ability to monitor and control complex systems, that operators achieve as a result of experience is, in part, a consequence of improved attention management strategies that they have acquired over time.

As these examples illustrate, biological cognition suggests how much better our engineered systems can be in terms of resource management, learning, and adaptation. Questions such as what new control methods are needed, how can generic platforms be developed, how can they then be specialized for critical applications, and how can we have some assurance that flexible, adaptive, learning-endowed cognitive control systems will operate reliably and consistently over extended time periods . . . these remain to be addressed by researchers in controls in collaboration with other disciplines.

Control Response to Rare and Sudden Events

Currently, almost all control systems are designed around structured nominal conditions. At the lowest level, a PID controller will regulate to a setpoint, using an error signal to determine how to move a valve or a motor. Although mathematically much more sophisticated, a multivariable predictive controller is conceptually similar—it processes sensor data with a fixed algorithm (in this case, model-based) and provides an output to a lower-level controller or an actuator. Little else is required for the operation of the control loop under nominal conditions, but what about sudden, and unmodeled, events: sensor or actuator failure, a drastic change in the plant, or a major disturbance?

Automation systems have strategies in place, from redundant devices to fault detection systems to safety shutdown systems, to deal with many such eventualities, but there is a qualitative difference between how expert human operators will respond to an unforeseen event and how today's automation systems respond. Partly as a result of training (often heavily reliant on simulators), pilots and process plant operators can continue the operation of an affected complex system in situations that would be beyond the scope of a fully automated system, based on the best of off-the-shelf technology.

One recourse, of course, is to explicitly model emergency conditions and to "program" appropriate responses to each. To undertake such a project for all conceivable situations would be impossible, but this strategy does not need to be an all-or-nothing one. So questions arise: Can one develop a systematic control design methodology weighing the human resource effort required for the design of fail-safe algorithms with performance when sudden events occur, given likelihoods of events as best they can be estimated? Is there a continuous progression of controls capability with increased human design effort? Can the design effort be automated or adapted online to such sudden events? Can control systems learn online when faced with rare events? Is this knowledge interchangeable through a rare event database with all local control systems feeding knowledge into this database? Therein lies more grist for the cognitive control research mill.

Acknowledgments

This section is based on the discussion on cognitive control at the Berchtesgaden Workshop. Participants in the group included Michael Branicky, Munther Dahleh, Klaus Diepold, Ulrich Konigorski, Henk Nijmeijer, Christoph Stiller, Dawn Tilbury, Georg von Wichert, and Janan Zaytoon. Their constructive comments and contributions to this section are highly appreciated. We also thank Anuradha Annaswamy for comments on an earlier version.

Selected recommendations for research in cognitive control:

- Control strategies for the adaptive management of cognitive resources in real-time systems need to be developed. Cognitive control systems will need to aggregate and consolidate information, balance long-term and immediate priorities, and shift attention dynamically as circumstances dictate.
- Human operators are still the preferred recourse for responding to rare and sudden adverse events. Research is needed to develop automation systems that can exhibit humanlike capabilities in such situations.
- Modeling and estimation take on added dimensions in cognitive control, with representations of self, the environment, objectives, and other elements required. Such representations must often be developed from partial and uncertain information.

References

[1] euCognition, the European Network for the Advancement of Artificial Cognitive Systems [Online], 2010. Available at http://www.eucognition.org/euCognition_2006-2008/definitions.htm.

[2] CoTeSys. Cluster of Excellence: Cognition for Technical Systems [Online], 2010. Available at http://www.cotesys.org.

Systems of Systems

Tariq Samad and Thomas Parisini

Introduction

The increasing scale and complexity of control applications have reached a point that imposes qualitatively new demands on control systems technology. Just as the transition from single-input, single-output systems to multivariable control required new theories, tools, and techniques, similarly, the new imperatives cannot be satisfied by evolutionary extensions of the state of the art.

Emerging applications are not just large-scale and complex; they are also characterized by decentralized, distributed, networked compositions of heterogeneous and (semi)autonomous elements. These new "systems" are, in fact, "systems of systems" (SoS) [1]. The term has arisen from the systems engineering community and reflects the interest in concepts and developments such as smart grids, integrated supply chains, collaborative enterprises, and next-generation air traffic management.

The challenges associated with designing, building, and operating systems of systems are not limited to control science and engineering. Yet the relevance of SoS to control, and vice versa, is apparent—as acknowledged by the frequent use of terms such as "systems and control" and "control systems." It can even be said that the importance of control increases as our conception of systems is broadened by encompassing consideration of multiscale and hybrid dynamics, cooperative and competitive architectures, multicriteria optimization, semiautonomous and autonomous systems, self-diagnosing and self-repairing systems, and so on. Systems of systems thus offer exciting opportunities for research in control (or systems and control) [2].

Below, we first note some properties of SoS. Next, we briefly compare and contrast SoS with another emerging research focus in the controls community, cyber-physical systems (CPS). We include a few examples to illustrate the increasing levels of interest in SoS. Before concluding, we discuss a few SoS-relevant research topics in control.

What Are Systems of Systems?

As the term implies, a system of systems is a composition; it consists of components that are themselves systems. But the term gains specificity with two properties that the whole must possess for it to be considered a system of systems [3]:

- *Operational independence of components.* The component systems fulfill valid purposes in their own right and continue to operate to fulfill those purposes if disassembled from the overall system; and
- *Managerial independence of components.* The component systems are managed (at least in part) for their own purposes rather than the purposes of the whole.

The "independence" aspect implies that autonomy is inherent in SoS—not just in the function of the SoS but also in the function of component systems. Autonomy in this context does not necessarily mean human-free operation; the human element may be part of the component system. But this subsystem

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

must be able to function independently on occasion and yet be a cog in a larger machine on other occasions. Dynamics in the evolving structure is a peculiarity of SoS.

The prospect of developing large, functionally rich, behaviorally complex SoS *ab initio* is unrealistic, especially given the requirement that component systems be useful entities in their own right. Systems of systems tend to exhibit evolutionary development—intermediate systems are developed that perform useful functions and are then integrated into larger systems. SoS will typically evolve through stable intermediate forms [4].

Other characteristics of systems of systems can be highlighted as well:

- SoS will be heterogeneous. From components to subsystems to systems, different technologies and implementation media will be involved.
- SoS will exhibit emergent behavior. Given their architectural complexity, the interaction of the SoS component elements will inevitably result in behaviors that are not predictable in advance.
- SoS will be large-scale systems. "Scale" should be interpreted more in a logical than necessarily a geographical sense—a system of systems can be a local entity with collocated subsystems.

Although these defining properties and characteristics do not explicitly invoke control, the relevance of the technology to SoS is evident given the dynamics involved in the component systems and compounded by the meta-system. Individual components will require control applications within them, and these control applications will interact explicitly (e.g., through coordination signals) or implicitly (e.g., through physical influences). Information technologies will provide the integration infrastructure, which is an enabler for closing the loop and optimizing design and operations (Fig. 1).

Although the defining properties and characteristics of systems of systems do not explicitly invoke control, the relevance of the technology is evident given the dynamics involved in the component systems and compounded by the meta-system.

Dynamics and control aspects of SoS are also critical for nonfunctional properties; SoS requirements cannot be limited solely to their core performance-related functions. Systems of systems have to be designed so as to provide assurances for predictability, dependability, and safety. Verification at several levels of abstraction will be required given the safety- and mission-criticality of engineered systems of systems, and such verification will need to be informed by the dynamics of SoS.

Systems of Systems and Cyber-physical Systems

SoS and CPS both represent exciting new vistas for control, in the eyes of the controls community as well as its government and industry sponsors. The influence of the revolutionary advances in information technologies is prominent in both areas, and thus it is not surprising that overlap exists. However, there are significant differences between CPS and SoS as well—differences reflected in the semantics of the terms.

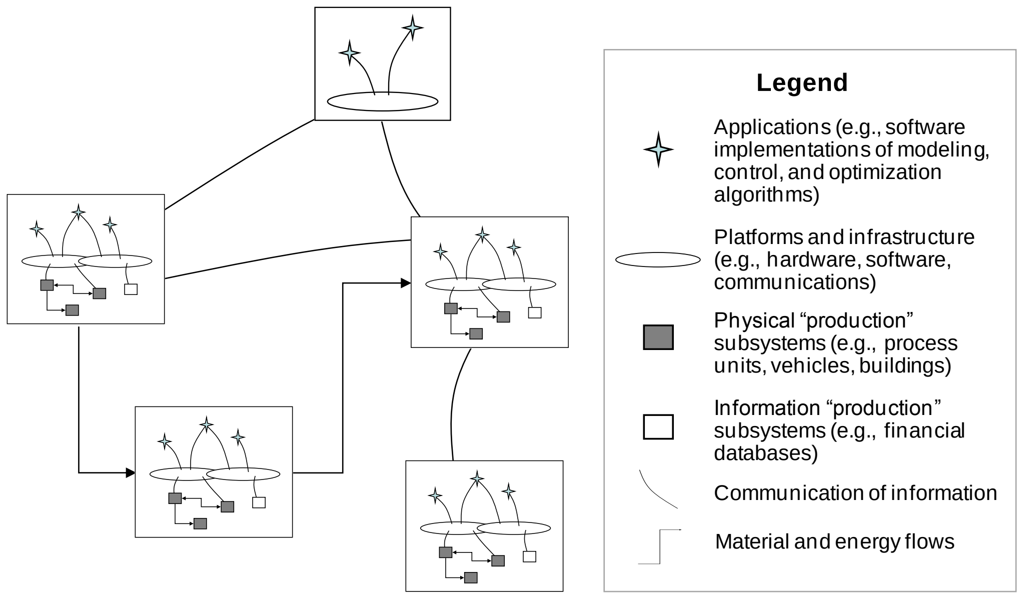


Figure 1. A schematic representation of a system of systems. Each component system may consist of applications, platforms, and "production" elements, the last of which can be physical systems or information systems—dynamics are very different in the two cases. Component systems may be integrated through information and/or material/energy interconnections. Additional component systems may be employed for coordination and control (the top system), and coordination can also occur among production systems.

First, the connection with physical systems is a defining feature of CPS. The interconnection of computer-based control algorithms and mechanical, chemical, or other processes governed by scientific laws has been exemplified by control systems since they became digital decades ago. With progress in computer science and related fields, new opportunities have arisen for control applications, and CPS research is attempting to capitalize on these opportunities.

Not all control applications are connected with the physical world, however. One point of divergence between CPS and SoS relates to applications that are purely in the information space. Control technologists are working in financial industries or otherwise developing applications to economic and market systems. Similarly, enterprise applications are a fertile target for control technology and do not necessarily require closing the loop in the "real" world. These applications are generally considered outside the CPS realm but not, at least necessarily, outside the SoS one. (It might be argued that even in these applications, the underlying processes are physical ones—what ultimately must be modeled is human psychology, for example. However, at least today this ultimate reduction is not being pursued.)

Second, many control applications, and many complex control applications, are not focused on distributed, hierarchical, and compositional mega-/meta-systems. Even a single-input, single-output PID controller, in a digital implementation, suggests opportunities for CPS research. Certainly the development of reconfigurable multivariable controllers running on sophisticated real-time platforms with adaptive scheduling would constitute significant progress in CPS. The connection with SoS is minimal at best.

These are definitional differences, and an obvious question is whether they lead to differences in research agendas and methodologies. This is a difficult question to answer at this stage of development of these fields, especially of SoS. Whereas the controls community has been instrumental in establishing

CPS (the field of more recent vintage) over the last several years, the engagement with SoS has only just begun. In any case, we do not expect significant qualitative contrasts, but rather variations in emphasis and prioritization. For example, topics such as verification and validation of real-time controller implementations or control over wireless links will likely be more prominent in CPS than in SoS. Conversely, game-theoretic negotiation algorithms strike us as more SoS territory.

SoS Examples

Examples of systems of systems, either existing or proposed, can be found in all societal sectors [5]: air and road transportation, power grids, healthcare, water management, industrial processes, building complexes, critical infrastructures, enterprise systems, smart homes and cities, and others. We discuss a few examples here, highlighting control connections.

Manufacturing Supply Chains

A large-scale manufacturing facility is a system of systems in itself, and today connectivity with upstream and downstream entities is being explored. The focus is largely on IT integration—platforms and communications that can, for example, automate ordering from suppliers based on inventory and production levels in a factory. Although the benefits of such automation are significant, the real value of the infrastructure is as a foundation for the optimization of the overall supply chain—enabling responsiveness to market conditions, maximizing energy efficiency, coordinating inventories with production plans dynamically, and the like.

Control loops exist within the entities in a supply chain (even suppliers and distributors that do not have manufacturing operations have feedback processes operating to service requests, accommodate inputs, and manage inventories; these are typically discrete-event processes, with simpler dynamics than a production operation). An interconnected supply chain establishes additional control structures with complicating factors. Different business entities are involved with their own, and often competing, priorities. Centralized or global optimization is not feasible. See Fig. 2 for an illustrative sketch.

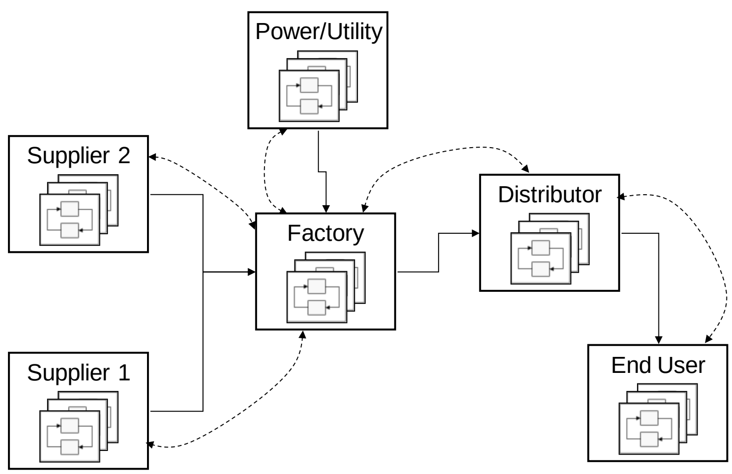


Figure 2. A control-centric view of an enterprise-level system of systems. Solid arrows show material and energy flows; dashed lines show information (including measurement, estimation, and command) flows. Individual systems contain optimization and control loops, and intersystem interactions realize higher-level control loops.

Embedded Automotive Systems

Today's cars are collections of embedded systems on wheels. Much of the innovation in the automotive industry in the last decade or two has been as a result of onboard computing, control, and communication, and this innovation has dramatically improved safety, fuel economy, emissions, and reliability. A number of separate embedded systems exist in a modern automobile—just those related to safety include collision impact warning, airbag deployment and seatbelt pre-tensioners, antilock and differential braking, intelligent cruise control, and traction and stability control. Often designed independently, these systems are nevertheless interdependent through the physics of the vehicle and the environment and the actions of the driver. Thus arose failure modes such as cars that locked themselves if the driver got out with the engine running and shut the door, or cars whose antitheft systems disengaged and doors unlocked if the cars were rocked side-to-side, triggering rollover detection.

The solution, evidently, is to adopt an SoS viewpoint when designing automotive systems (Fig. 3). Standard network protocols and buses have already been adopted in vehicles. Some level of algorithmic integration has also occurred—some systems coordinate traction control and antilock braking, for example. But much remains to be done, and with the continuing rollout of X-by-wire systems (e.g., active steering), more opportunities will arise.

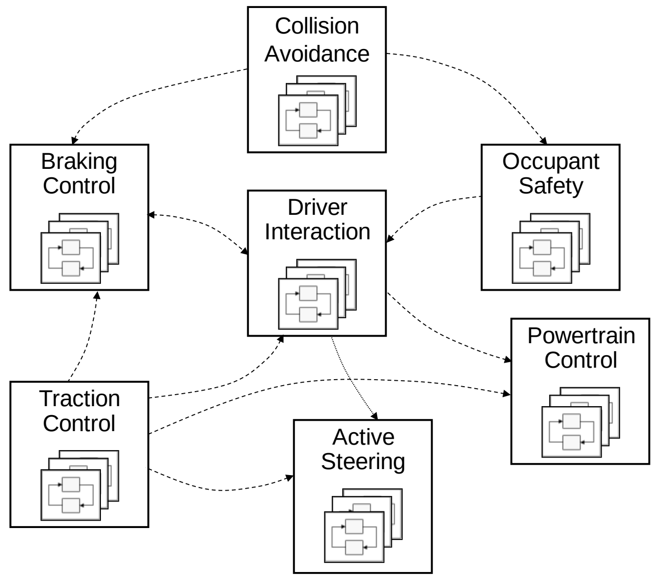


Figure 3. An "embedded" system-of-systems example (automotive) illustrating control applications and their dependencies. The dependencies can be realized through the physics of the vehicle-driver-environment SoS or through explicit control commands. (Not all control-related embedded systems are shown and not all possible interactions are depicted.)

With developments in intelligent road transportation systems, communication and coordination among vehicles and between vehicles and infrastructure elements (road signage, traffic lights, etc.) will further increase the SoS web. We have focused here specifically on intra-automobile systems to make the point that (unlike most examples that are discussed) the SoS vision, and its strong control connections, are also relevant in localized embedded electronic domains.

Smart Grids

Smart grids are a topic of tremendous interest worldwide. They represent a revolutionary advance over today's power grids enabled by two-way flow of both electricity and information. Smart grids incorporate an overlay of communication and control over a modernized power system infrastructure, resulting in a cyber-physical system extending from generation to consumption and facilitating the integration of distributed storage and generation (especially from renewable sources) and electric and plug-in hybrid vehicles.

Today, electricity consumption is, for the most part, independent of the exigencies of supply. Adjustment of consumption may be desired for several reasons—generation shortfalls, desires to ramp down use of polluting or expensive generation assets, better use of renewable generation, bottlenecks in the transmission system—but no systemwide infrastructure exists to realize such adjustment. Similarly, opportunities to effect optimized control of transmission and distribution grids, accurately monitor and communicate system state, closely connect power markets with power flows, and achieve other advanced power system capabilities are limited by the existing infrastructure.

A smart grid, as a system of systems, will enable such functions. One example is depicted in Fig. 4. Autonomous control units manage generation (including renewables and combined heat and power (CHP)), storage devices such as fuel cells, electric vehicles, and building loads—and the future may bring as-yet-unknown technologies. Utilities and system operators can interact with these master controllers and also with market entities. Individual control systems must satisfy local objectives, but they also need to cooperate to ensure the reliable and efficient operation of the power system. In the figure, the coordination is through a central node—this is closer to today's situation; in the future we can anticipate less hierarchical, more collaborative decision and control architectures.

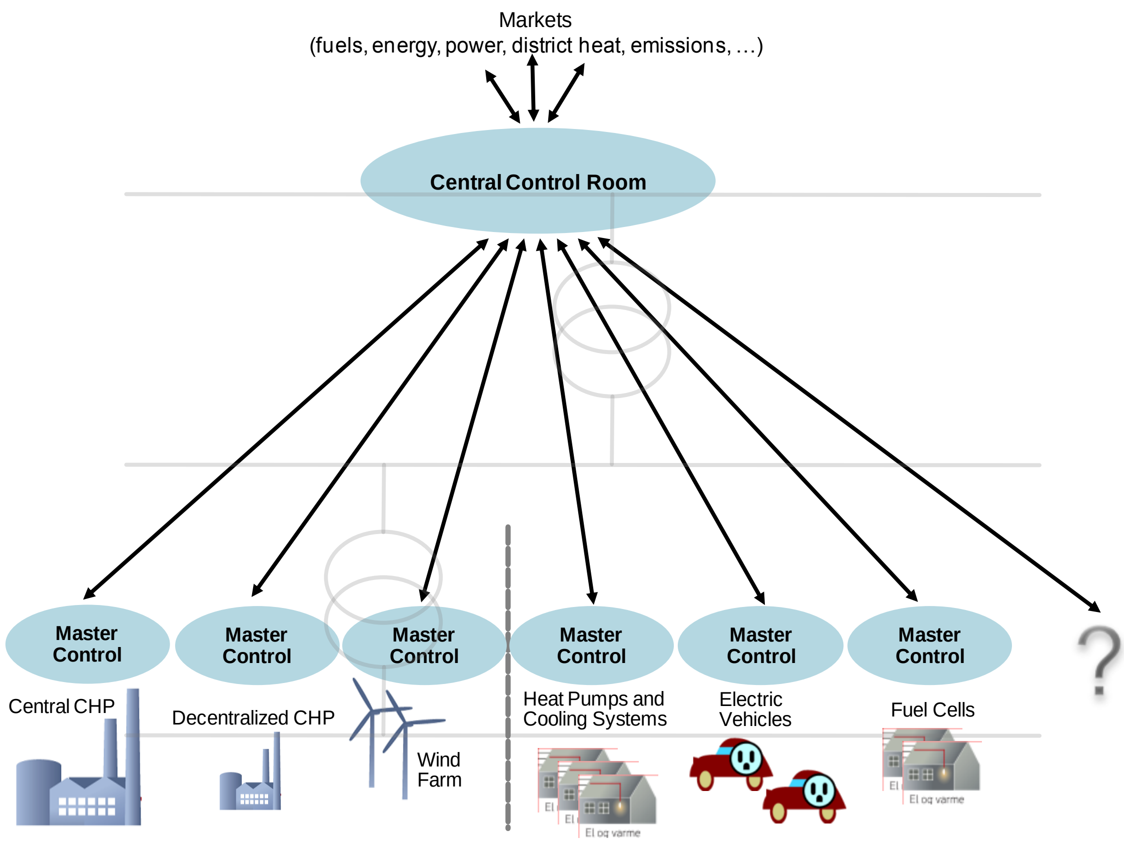


Figure courtesy of Dong Energy

Figure 4. Smart grid SoS example.

Research Opportunities for the Controls Community

Although systems of systems is not the research preserve of any one field—it is truly a multidisciplinary research frontier, as noted earlier—systems and controls constitutes a core enabling discipline. Here we discuss some of the SoS research implications for the controls community.

Cooperative/Coordinated/Collaborative Control

> *Systems of systems is not the research preserve of any one field—it is truly a multi-disciplinary research frontier—but control is a core enabling discipline.*

The adjectives overlap and are often used interchangeably, but altogether this is currently one of the most active topics of research in control systems. Formation flight and coordinated robot motions are the main application targets; the focus is on vehicular/mobile agents and geometric relationships. Notable theoretical results have been generated. A system-of-systems perspective can further enrich research in the area by broadening the space of applications.

For agents that are capable of autonomous or semiautonomous operation, cooperation and collaboration imply task-level interactions. Indeed, in the SoS context, it should be expected that component systems have their imposed goals but might also generate (in an evolutionary way) their own goals—causing dynamic interactions with other component systems. The relevance for control becomes especially prominent when temporal aspects must be considered—whether at the level of individual tasks or of the interaction. Both continuous-time behaviors and discrete decisions can be involved, and the interactions between the two offer particular opportunities. Consider, for example, scenarios in which the usefulness, for one agent's objectives, of a task being undertaken by a second agent varies based on both the degree of completion of the task and the elapsed time; the first agent must continuously decide whether to wait or to incorporate partial results. Expanded over a large scale, the complexity can be overwhelming, and new approaches will be needed.

In general, richer formulations for cooperative/coordinated/collaborative control are needed for the potential of this area to be realized. Theoretical and algorithmic contributions to systems of systems will be spurred by such broadening of perspective.

Identification, Learning, and Adaptation

Embedded models are a prerequisite for advanced control. However, modeling for systems of systems brings complexities that are often not encountered at the subsystem level. Effective techniques are available for developing models in general, and control-relevant models in particular, at a component level, but these cannot straightforwardly be extended to SoS. In particular, a distributed assemblage of independent, heterogeneous elements renders the prospect of centralizing knowledge about it problematic. Thus, instead of modeling as we often know it, with its first-principles orientation, the emphasis with systems of systems is likely to shift toward more data-driven, empirical techniques such as identification, learning, and adaptation.

Empirical modeling is subject to theoretical limitations based on partial information—as is the use of derived models for decision-making and control. There is no gainsaying these limitations, but they can provide guidance for research. In this context, we offer a few suggestions below.

- Generic, prepackaged solutions for large-scale applications are not a realistic objective. The "one size fits all" model is not a scalable one. Knowledge of problem domains, even if it is heuristic in nature, must be incorporated in customized approaches and algorithms.

- Levels of uncertainty in modeling and control are likely to be significantly higher with SoS. Greater attention must be paid to the stochastic aspects of learning and identification. Uncertainty and risk must be rigorously managed. This challenge plays to the strengths of control science and engineering.

- Autonomy, in the sense of operator-free automation, is not a viable prospect for many systems of systems. Given the likelihood of model-reality mismatch and the safety- or performance-critical nature of systems of systems, ultimate decision authority is likely to lie with humans, not machines. The human-in-the-loop aspect must be considered—it lessens the responsibility that otherwise would rest with automation and at the same time opens new opportunities for research in learning and adaptation.

Monitoring, Fault Diagnosis, and Fault-Tolerant Control

Increased scope and scale can imply a proportionate increase in risk. Adverse impacts from a low-level component failure can be managed. Failures at the system-of-systems level can have truly catastrophic consequences—for individuals, societies, system owners and operators, and the environment. Extreme levels of safety, reliability, and dependability will be required of systems of systems. With the large numbers of components and interconnections, individual failures will be unavoidable. Instead, methodologies and tools will be needed that can ensure safe and reliable SoS operation even in the face of component faults.

These arguments suggest renewed emphasis on monitoring, fault detection and diagnosis, and fault-tolerant control. Recently, the controls community has made significant progress in these areas. Rigorous, scalable theoretical results are now available, and sophisticated algorithms have been developed. Yet the SoS perspective further raises the bar on requirements and will provide further motivation for continued research.

Automation and Control Architectures for Systems of Systems

"System of systems" refers not only to a physical application configuration, but also to the automation and control infrastructure required to support the coordinated operation of heterogeneous autonomous and semiautonomous elements. Large-scale complex physical systems exist today and are supported by large-scale complex automation systems. These latter, however, tend toward centralized command-and-control architectures. The strictly hierarchical approach is untenable for SoS; "cooperation and coordination" is the appropriate metaphor, not command and control (traditionally construed). New developments in platforms and architectures will be needed.

Middleware is one central infrastructure need. Flexibility and adaptation will be essential to SoS and will need to occur in real time. Plug-and-play features will also be increasingly important. Traditional real-time systems restrict online flexibility, a conservative strategy that ensures stability at the cost of agility. An open question for the controls research community is whether this conservative approach can be overcome—whether, based on advances in real-time systems, wireless networks, embedded intelligence, and componentized software (for example), new control architectures can be developed that are not static and hierarchical but are constituted by flexible, adaptive, dynamic networks of cooperating objects.

The envisioned features for SoS automation and control must be attained without compromising safety, reliability, or security. Of especially critical importance is cyber security. The autonomy and heterogeneity of SoS imply a lack of centralized control. Multiple platforms and communication protocols will be involved, software and hardware components will be dynamically updated, and greater system responsiveness will be demanded. These are desirable, even revolutionary, benefits, but diligence must be exercised in building security features into the devices, software, protocols, and operational procedures.

Conclusions

Systems of systems are an exciting vision for the engineering community in general and for controls researchers in particular. The concept represents the culmination of developments in complex systems engineering over the last few decades, with obvious and broad-based advantages to society and industry. At the same time, SoS also represents a set of fresh research challenges. The multidisciplinary nature of SoS will require close collaboration between control researchers and researchers from several other fields—a positive development from several perspectives.

The control-related research required for systems of systems covers the basic-to-applied spectrum: advances are needed in areas ranging from fundamental systems science to the development of specific applications. What is crucial, however, is for basic research to be informed by application prospects and, conversely and at least at this early and immature state of the field, for application-oriented explorations to be conducted with awareness of the broader scientific issues.

Acknowledgments

We would like to thank Anuradha Annaswamy, Svetlana Klessova, Bruce Krogh, and Jorge Pereira for comments and suggestions.

Selected recommendations for control-related research in systems of systems:

- For SoS visions to be realized, robust, scalable algorithms for cooperation and coordination among heterogeneous autonomous and semiautonomous components—that can effectively balance local and global objectives—must be developed.
- Given their scale, systems of systems will always be faced with component-level faults; fault-tolerant and fault-adaptive methods are needed to ensure safe and reliable operation nonetheless.
- New automation and control architectures, hierarchical and heterarchical, are required that are based on dynamic networks of cooperating, flexible, and adaptive objects.

References

[1] M. Jamshidi (Ed.). *System of Systems Engineering: Innovations for the 21st Century.* New York: John Wiley & Sons, 2009.

[2] T. Parisini. Control systems technology: Towards a systems-of-systems perspective, *IEEE Trans. on Control Systems Technology,* vol. 18, no. 2, p. 249, March 2010.

[3] M.W. Maier. Architecting principles for systems of systems, *Systems Engineering,* vol. 1, no. 4, pp. 267-284, 1998.

[4] E. Rechtin and M.W. Maier. *The Art of System Architecting.* Boca Raton, FL: CRC Press, 1997.

[5] DG INFSO G3. Report of a workshop on systems of systems (working draft), Brussels [Online], 21 September 2009. Available at ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/esd/workshop-report-v1-0_en.pdf.

Part 4

GRAND CHALLENGES FOR CONTROL RESEARCH

High-resolution versions of the flyers in this section are available at www.ieeecss.org/main/IoCT-report.

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 sugar, much the same way the driver uses the brake and gas pedals in an automobile. Insulin functions as the "brake pedal," lowering the blood sugar by stimulating the uptake of glucose from muscle, fat, and kidney cells. Balancing this is the counter-regulatory hormone glucagon, which acts primarily to break down glycogen in the liver, yielding glucose and an elevation in blood sugar levels (acting as the "gas pedal"). Insulin and glucagon are both produced by the pancreas.

Although type 1 diabetes is currently incurable, the development of a reliable artificial pancreas would considerably improve the lifestyle of subjects with this disease.

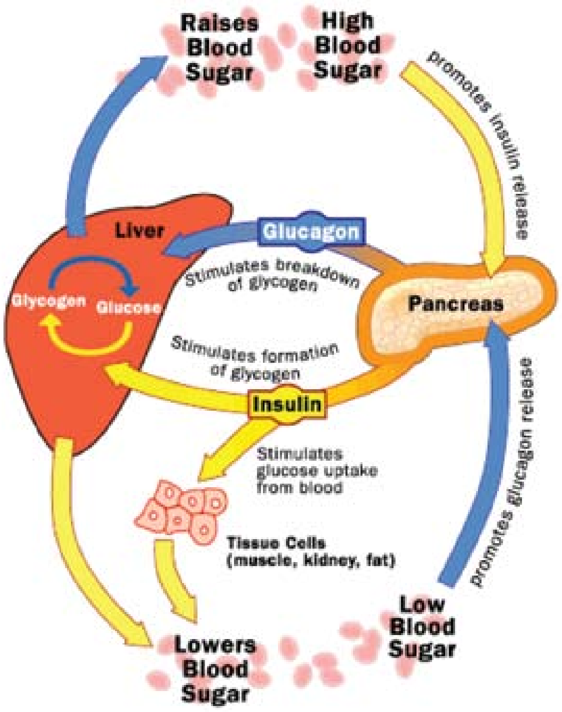
Why Is Systems and Control Relevant?

The systems and control community can play a critical role in developing architectures for the reliable automation of blood glucose monitoring 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

The Artificial Pancreas Vision

In addition to control algorithms, an automated, fully closed-loop device will require sensors and actuators. 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!



Type 1 Diabetes

- Type 1 diabetes is an autoimmune disease leading to insufficient or no production of insulin by the pancreas. The disease causes wide swings in 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 and frequently result in hypoglycemia (low blood sugar) events 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 continuous subcutaneous insulin infusion (CSII) pump.
- Type 1 diabetes affects 1 million individuals in the U.S. with associated annual medical costs of $15 billion.

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

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

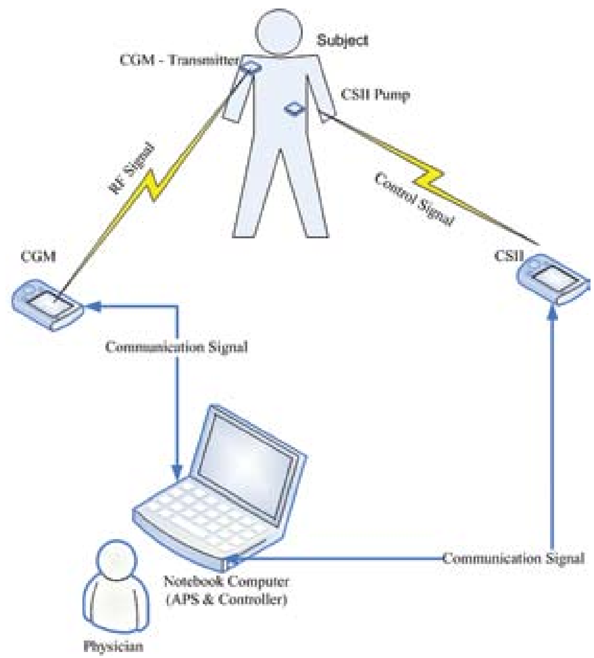
Barriers and Challenges

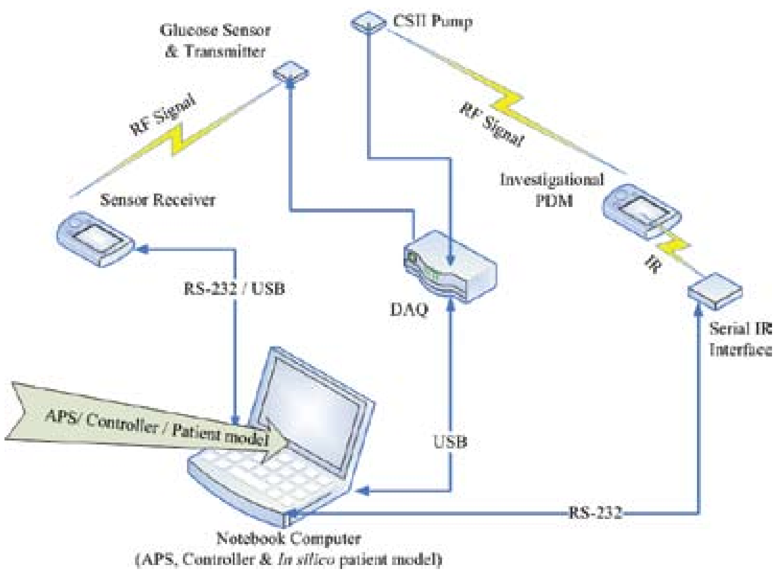
Several technical, policy, and cultural barriers must be addressed before a viable artificial pancreas can be developed:

- 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 face regulatory hurdles.

Several specific engineering challenges must also be resolved:

- Arrays of glucose sensors that are based on a different operating concept than today's sensors
- Dual-chamber pumps that will allow delivery of both insulin and glucagon
- Accurate predictive patient models that can be customized for the individual
- 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. 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.

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.

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. It is also being used to link glucose sensors and insulin pumps using wireless protocols (top figure). Combining the APS© with a feedback control algorithm enables both hardware-in-the-loop testing (bottom figure) and closed-loop human clinical trials.

For further information: http://www.artificialpancreasproject.com/

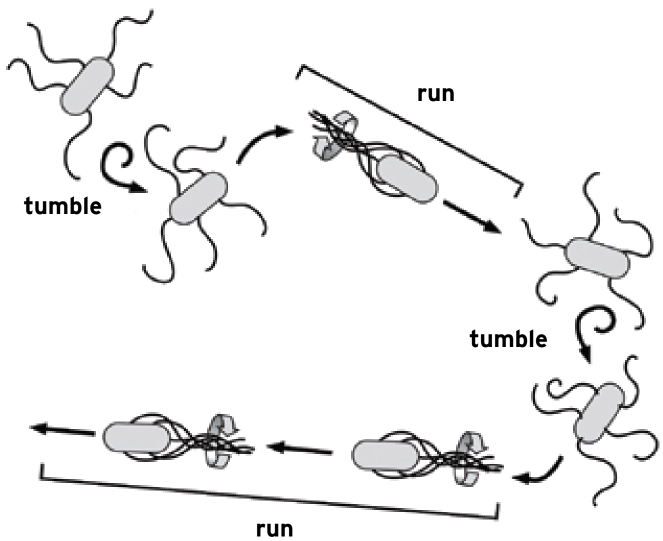
Redesigning a Bacterium Control System

Systems Engineering for Biology

"Systems biology" and "synthetic biology" are two major growth areas within biology. For these efforts to be successful in the long run, a systems engineering framework for biological circuit design must be built. Recent successes in building synthetic circuits that provide novel biological function (such as an oscillator or a programmable switch) demonstrate that the basic technology is at hand. However, initial attempts to systemize synthetic biology have not yet succeeded in building working systems from libraries of standard parts. The challenge lies in exploiting the modulator of molecular biology while at the same time gaining enough insight into the fundamental processes to understand key issues in building larger and larger systems from individual components.

Target Problem: Bacterial Chemotaxis

Chemotaxis is the process by which bacteria and other microorganisms sense chemical signals in the environment and adjust their motion to either move toward the signal (chemoattractants) or away from the signal (chemorepellants). The chemotaxis system in E. coli consists of a sensing system that detects the presence of nutrients, an actuation system that propels the organism in its environment, and control circuitry that determines how the cell moves in the presence of chemicals that stimulate the sensing system. Each of these subsystems is implemented via proteins inside the cell, with communication, computation, and control intermingled through the various molecular reactions that occur. Many chemotaxis mechanisms are stochastic in nature, with biased random motions causing the average behavior to be either positive, negative, or neutral (in the absence of stimuli).



Physical Biology of the Cell
© Garlance Science 2009

Scientists are able to genetically modify microbiological organisms so that they produce certain chemicals or change their behavior. Can we redesign the control systems in bacteria (including implementation!) so that we can program their behaviors in response to external stimuli? Possible applications include new types of medical treatments, new methods for environmental remediation, and in vivo sensing systems. Initial demonstrations have successfully modified the sensing system, but true reprogramming would include systematic methods for designing the control system to have specified closed-loop properties, including stability, performance, and robustness.

Contributor: Richard Murray, California Institute of Technology, USA

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Chemotaxis Control System

The components that implement chemotaxis are becoming increasingly well understood. For example, flagellar motors have been imaged in detail (Figure 1), showing the incredible structure present in these systems. The cryo-electron micrograph in the top figure shows the molecular components of the motor, with the components labeled in the lower figure. The motor sits between the outer and inner membranes (OM, IM) and consists of a rotor stator and other elements.

The basic control circuitry (Figure 2) consists of membrane-bound proteins for sensing chemical signals (ligands) external to the cell, signal transduction pathways that communicate the presence or absence of ligands, and feedback regulation mechanisms that modify the sensitivity of the sensor and act as a form of integral feedback. The effect of this sensing and control system is to modulate the flagella such that they spin either clockwise or counterclockwise, resulting in either tumbling or nontumbling motion. Dynamical models have been developed that describe the various processes present in chemotaxis and give insight into the system-level properties.

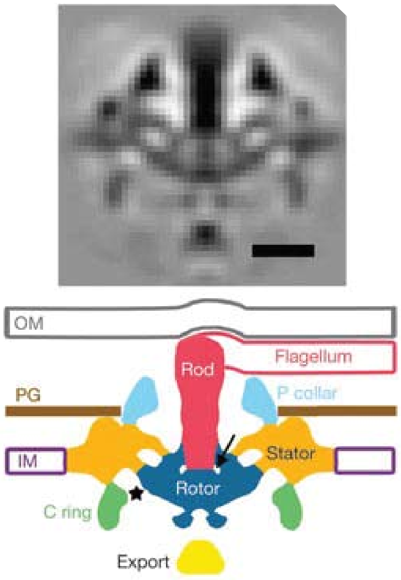


Figure 1: Detailed image of the flagellar motors

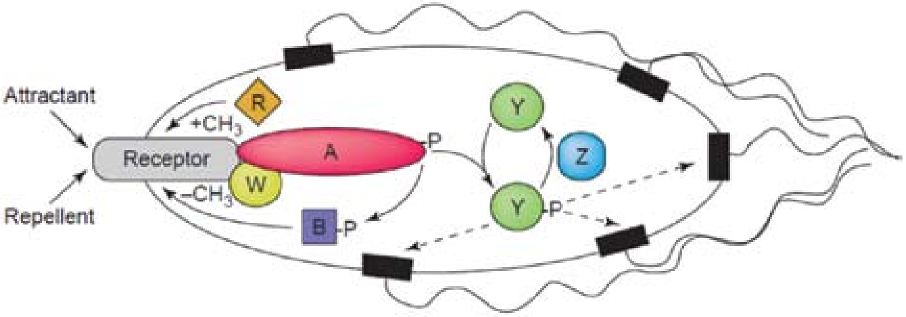


Figure 2: Basic control circuitry

Design Specification and Properties

We seek to design a control circuit that can be inserted into a bacterium and that will modify its behavior so that the bacterium can be used to deliver a drug to a specific area in an organism. The condition the bacterium responds to should be substantially more specific than current drug treatments, targeting a very localized area in the body or an organ, for example. The drugs that are released can either be synthesized within the bacterium or stored in some inert form and released at an appropriate time and location.

Rather than just experimenting with existing cellular organisms and genetic mutations, the circuitry used to accomplish this task should be designed in the usual engineering sense. This includes developing a model that predicts the behavior of the control circuit in the system and can be used to iterate on the design before synthesizing the DNA that encodes the circuit. Current technology already allows simple biological circuits to be built and inserted into bacterium, and there have been many genetic engineering demonstrations to modify the behavior of existing systems. The goal in this grand challenge is to move from a culture of experimentation and invention to one of systematic modeling, analysis, and design.

And there's much more!

Bacterial chemotaxis is just one of many interesting processes implemented using biomolecular feedback systems.

For further information: Nature special issue on synthetic biology, http://www.nature.com/nature/journal/v438/n7067/index.html; Registry of Standard Biological Parts, http://parts.mit.edu/

High-Performance Control with Slow Computing

Current techniques for the design of software-enabled control systems rely on the existence of high performance sensing, actuation and computational devices that can be embedded within a physical system at modest cost. Driven by Moore's law, the success of this paradigm can be seen through the broad usage of feedback controllers in modern application areas. The goal of this challenge lies at the other end of the computational spectrum: Can we develop new principles and tools for the design of closed loop control systems using highly distributed, but slow, computational elements?

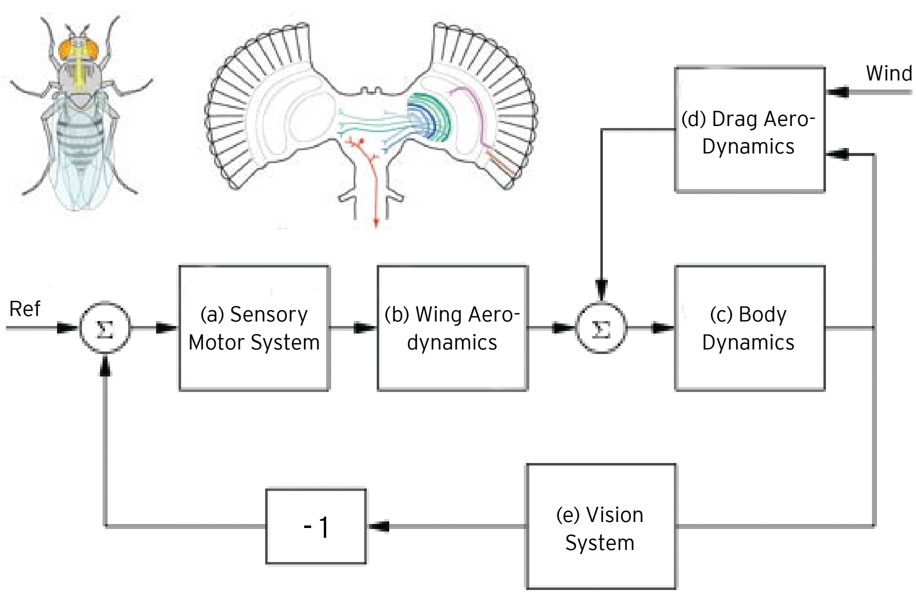
The motivation for control design using slow computing is to develop control system architectures for applications where computational power is extremely limited. One important class of such systems is that for which the energy usage of the system must remain small, either due to the source of power available (e.g., batteries or solar cells) or the physical size of the device (e.g., microscale and nanoscale robots). A longer term application area is the design of control systems using novel computing substrates, such as biological circuits. A critical feature is the fact that the speed of the basic computing elements is similar to the underlying dynamics, leading to tight coupling between dynamics and computing.

Design of feedback systems using slow computing is particularly challenging because of the performance limitations associated with computational delays that are comparable to the underlying plant dynamics. Highly parallel, non-deterministic architectures are likely to be needed to achieve what is normally accomplished through the tightly synchronized, serial interconnections of sensing, filtering, estimation, planning, regulation, and actuation that are common in traditional control systems. Unfortunately, current techniques for systematic design of control systems assume a mostly serial processing architecture and techniques that make use of parallel architectures (such as neural networks) do not provide sufficiently systematic design methods. New research is needed to develop the architectures, theory, and tools required to design controllers where computational delay does not allow current techniques to be used.

Context

Current approaches to the design of software-controlled systems make use of a combination of abstractions and design techniques that are often implicitly based on the assumption that significant computational capacity is available to implement computations and communications. This is a good assumption for many application areas where substantial amounts of computing can be embedded within a physical system to control the dynamical behavior of the underlying process. As a consequence, many of the approaches that are available for designing complex, "cyberphysical" systems rely on large amounts of computing to achieve complex and robust behavior.

As a complementary approach, consider instead the control system for a fruit fly, depicted in the figure below. This system uses approximately 300,000 neurons with typical time constants in the range of 1 – 100 msec (10 – 1000 Hz) and is approximately the size of a sesame seed. Yet it is able to take off, land, avoid obstacles, find food and mate (among other things), often with performance that is beyond what we can do in engineered systems at this size scale. As just two specific instances, the control system of a fly is capable of executing saccades (rapid changes in direction) that occur at angular rates of up to 1800 deg/sec and it can fly in wind gusts that are up to twice its flight speed in air.



Control System for a Fruitfly

Contributor: Richard Murray, California Institute of Technology, USA

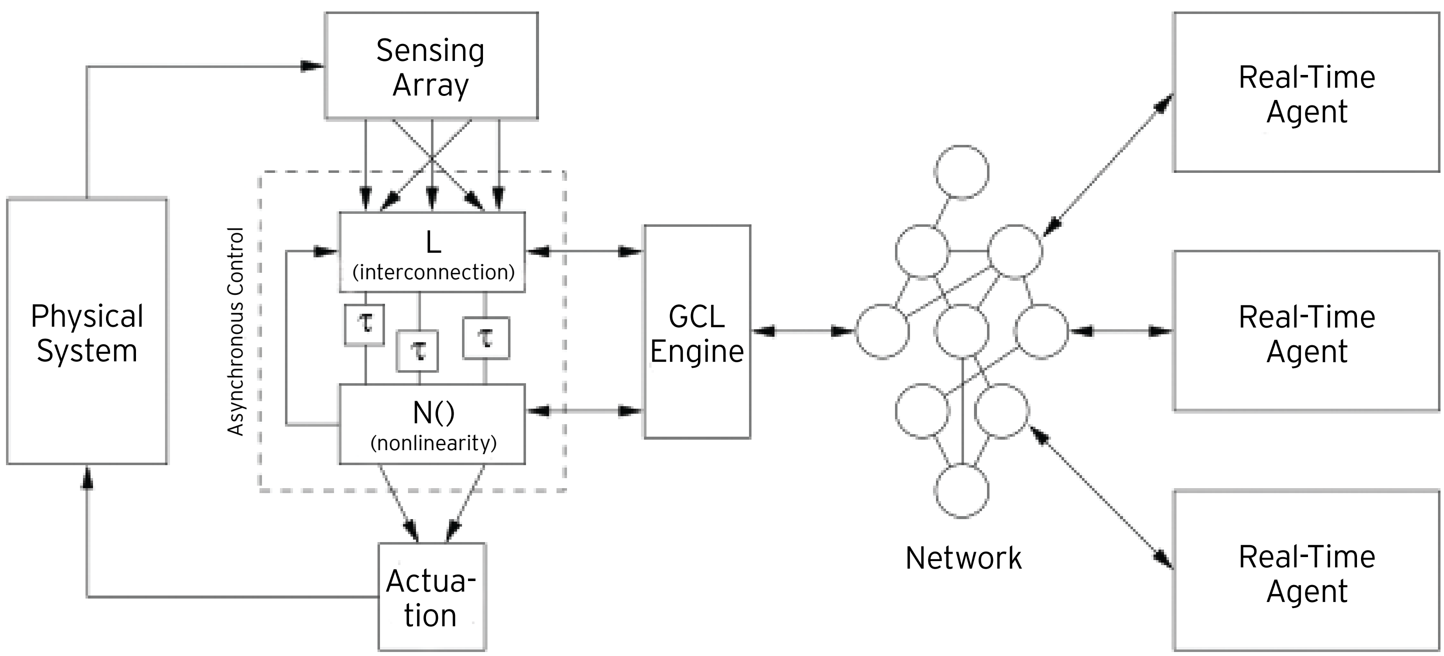
From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Challenge

The goal of this challenge is to develop the fundamental insights and tools that will allow us to design control systems that can perform the tasks of modern high-speed control systems but using an architecture that is compatible with the speed of computation used by an insect. The development of such an architecture has the possibility of providing new ways of integrating control into systems where large amounts of fast computation are not easily available, because of limitations on power, physical size, or choice of computing substrate. It is likely that many of the tools and insights required to design such systems will prove to be central to the design of other classes of systems in which the effects of time delay, asynchronous execution of parallel computations, and highly complex interconnections play a defining role.

A candidate architecture for the implementation of such a system is shown below. The system consists of a set of agents that are interconnected through a network (the portion of the figure to the left of the network illustrates the architecture within an agent). Each individual agent has a highly structured inner loop, which interacts with a guarded command language (GCL) protocol engine that triggers rules to change the discrete state of the system. The "inner loop" control system makes use of an interconnection matrix L, a set of asynchronous delays (represented by the blocks labeled τ) and nonlinear elements N(). Internal feedback between the nonlinear block and the interconnection block allows a general set of dynamical systems to be formed from this simple structure. The protocol-based feedback system modifies the inner loop dynamics, but also controls communications between other agents, using a packet-based communications network. Multiple agents interact with each other across the communications network, using packet-based communication protocols. This network introduces another layer of variable time-delays and asynchronous behavior, enabling complex behaviors by the multi-agent control system.

The complexity of the architecture and of the resulting behavioral space will require the development of formal tools for specification, design and verification.



Biophysical Networks

Systems biology refers to the understanding of biological network behavior through the application of modeling and simulation tightly linked to experiment. The "network" aspect of biological systems has become especially prominent recently as a result of the understanding that complex phenotypes (disease states such as cancer or diabetes, or the infection of a host by a virus or bacterium) are governed by the behavior of genetic networks rather than single genes.

The network-centric approach to biology is already yielding insights into complex disease pathways and shows great promise in the identification of novel disease readouts, as well as potential (vectoral) drug targets for implementation of control measures by small molecules, RNAi's, monoclonal antibodies, and other approaches.

In the case of bacterial infections, the network analyses that detail the attacking pathogens and subsequent hijacking of host cells span multiple scales (temporally and spatially), including sequence knowledge of viruses and bacteria, gene regulation, protein-protein interactions between hosts and pathogens, immune-receptor signaling, and ultimately organ-level analyses of physiological responses.

Why Is Systems and Control Relevant?

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

- Understanding and treating diseases,
- Disease inference and tracking of progression using novel assays,
- Novel (molecular) drug treatment, and
- Developing systems methodologies for implementing personalized medicine.

Complex Biophysical Networks

Control and dynamics tools can be used to interrogate complex signaling networks, such as those responsible for apoptosis (programmed cell death; Figure 1). Key modules within the network can be isolated, such as the crosstalk between two different TNF receptors (Figure 2) and elementary control principles can be used to elucidate the role of positive and negative feedback between receptors on the dynamic response of the system to perturbations.

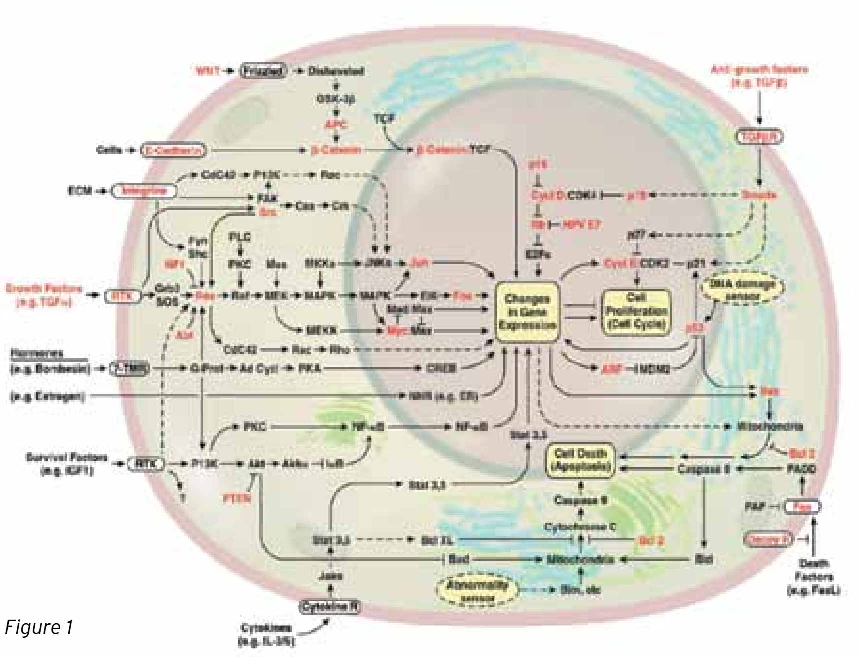


Figure 1

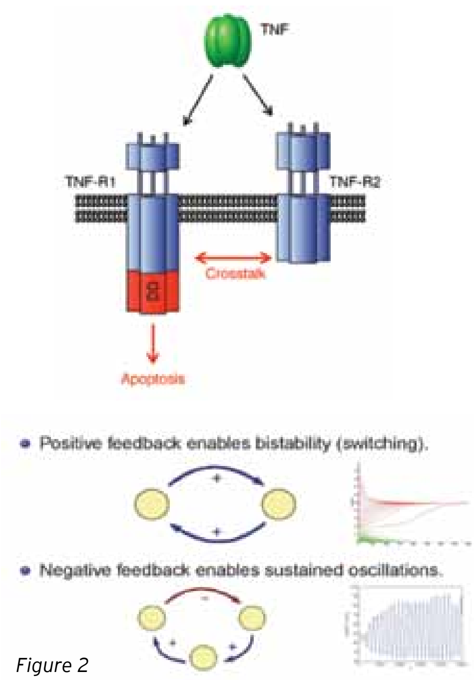


Figure 2

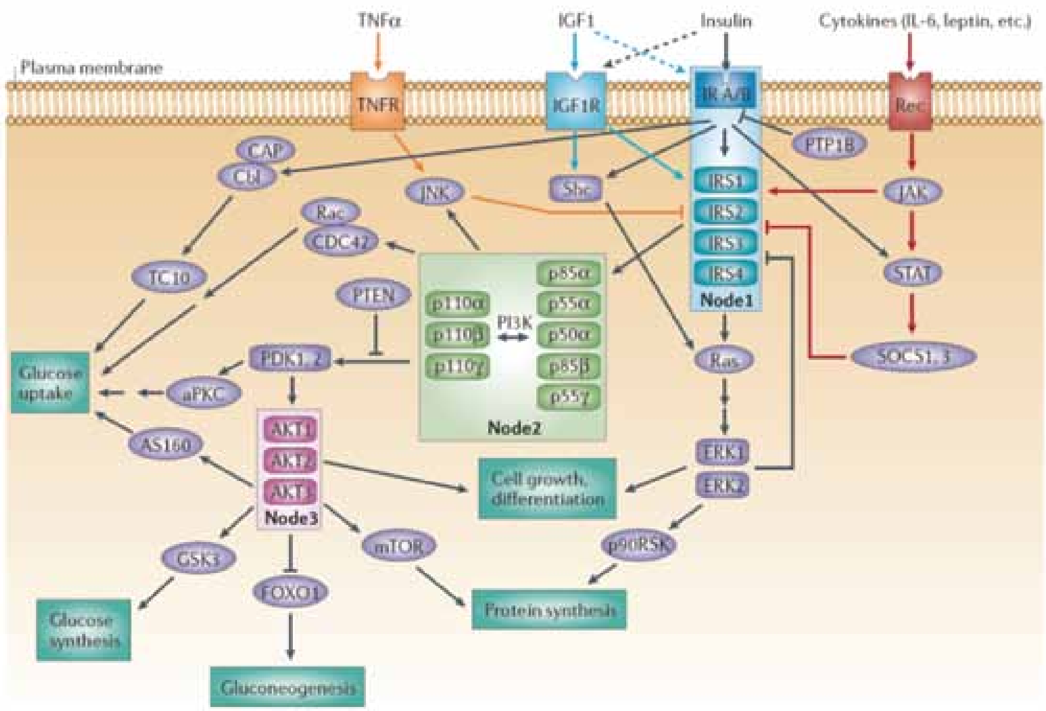
Contributors: Francis J. Doyle III, University of California at Santa Barbara, USA and Richard Middleton, National University of Ireland Maynooth, Ireland

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Example: Type 2 Diabetes

Type 2 diabetes is a metabolic disorder primarily characterized by hyperglycemia and insulin resistance. An estimated 350 million people worldwide will be affected by the year 2030. In the U.S. today, 14 million have type 2 diabetes, with associated annual medical costs of $132 billion. The disease is linked to obesity from high caloric intake combined with low physical activity. Current drugs are ineffective, and the consensus is that we do not understand the disease well enough (from a molecular network level) to choose appropriate drug targets. Biophysical networks are extremely "noisy" due to molecular-scale fluctuations.

The combination of systems modeling of the biophysical network with robustness analysis (sensitivity of candidate drug targets in the face of molecular noise) is a visionary and exciting prospect.

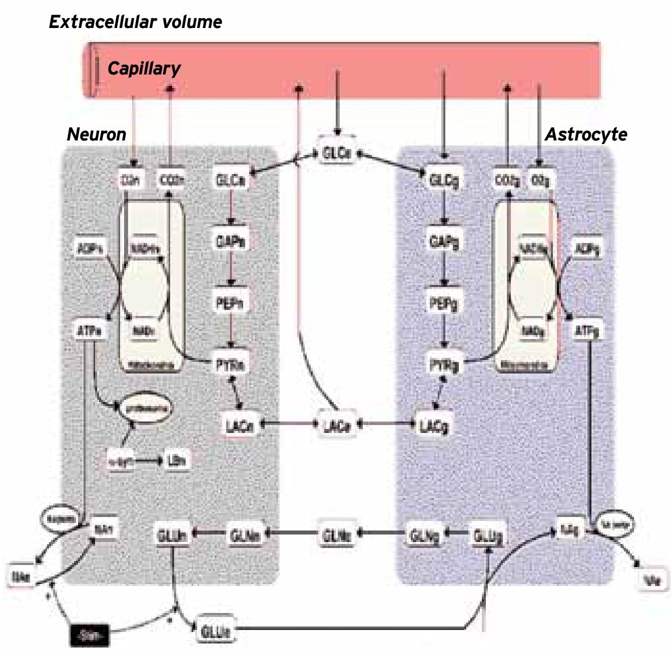


Example: Parkinson's Disease

Worldwide approximately 6.3 million people are living with this disease. A conservative cost estimate for treatment is $14,000 per patient year. Total costs are poised to rise sharply with population aging.

The causes of neuronal degeneration in Parkinson's are poorly understood: no clear genetic marker has been identified; toxin exposure and stress have been postulated with mixed results; metabolic changes with aging are clearly correlated, but precise mechanisms are unclear.

Systems biology can help integrate diverse insights into the disease and thereby elucidate key dynamic interactions and causative factors.

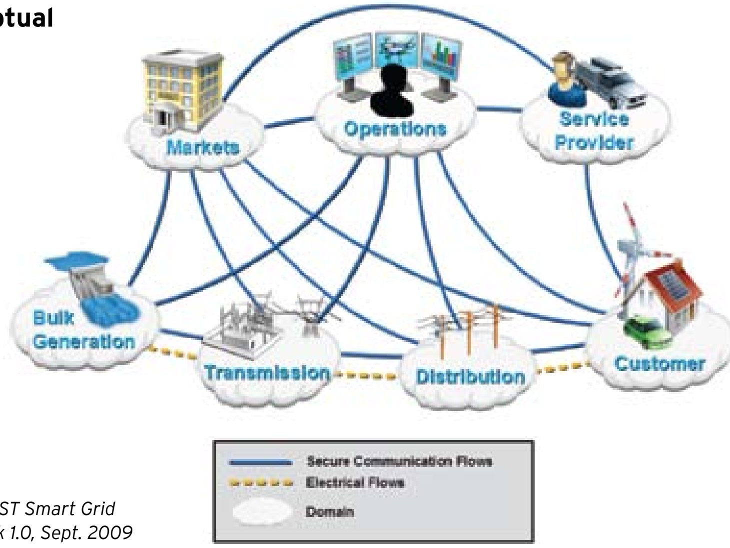


J. Comput Neurosci DOI 10.1007/s10827-09-0152-8

Control for Smart Grids

Today's electric power infrastructure is ill suited for dealing with global energy and environmental concerns. The "smart grid" promises a solution to this predicament. By incorporating a communication, computing and control overlay on the power grid, we can integrate large-scale renewable generation and emerging storage technologies, provide (direct and indirect) control signals to loads to match supply, dramatically improve energy efficiency and reduce consumption in homes, buildings, and industries, and increase the performance and reliability of transmission and distribution networks.

Conceptual Model



Source: NIST Smart Grid Framework 1.0, Sept. 2009

Dynamics, feedback, stability, optimization—these and other concepts that are core to control science and engineering are at the heart of the smart grid. It is perhaps only a slight exaggeration to say that the smart grid is, in essence, a controls problem!

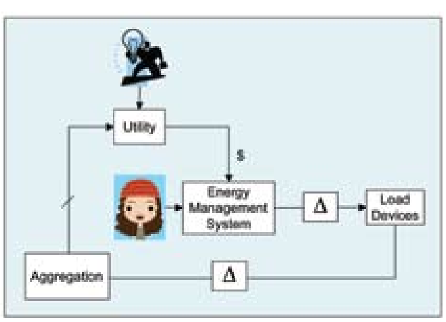
Demand Response

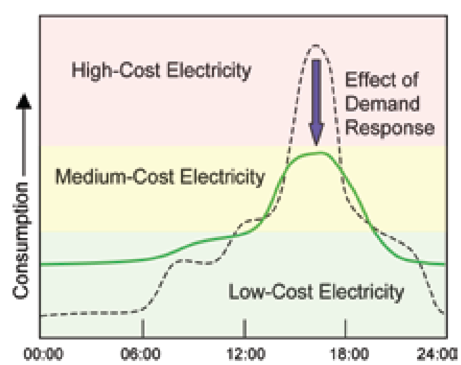
Electricity demand varies significantly over the course of a day (and over longer-term cycles), and the cost of servicing the demand varies even more dramatically. For example, the marginal cost of generating the electricity needed to satisfy demand on a hot summer afternoon can be more than an order of magnitude greater than the baseload generation cost.

Demand response refers to mechanisms that can enable electricity consumption to be better aligned with generation cost structure—and with other imperatives such as the use of renewable sources. Demand response comes in different forms, such as direct load control in which the utility directly switches loads in consumer premises, remote adjustment of set points of equipment, and, perhaps most interestingly, dynamic pricing signals to consumers (see figure below).

Demand response is practiced today in very limited fashion—for example, hourly day-ahead prices may be communicated to energy management systems in commercial buildings. Real-time pricing and more dynamic demand response schemes have garnered much interest, but this tighter closing of the control loop must be carefully thought out.

- Models are needed that capture the dynamics of customer behavior in response to price signals.
- The effect of delays (the Δs in the diagram below) in adjusting loads must be understood—this is critical from a stability perspective too.
- Control strategies will need to tolerate the high level of noise and uncertainty in this application.





Contributor: Tariq Samad, Honeywell, USA

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Renewables, Distributed Generation, Storage

For all the potential of renewable generation, it complicates the overall balance of supply and demand. Traditionally, power companies have been able to control generation with reasonably complete authority, thereby addressing variation in demand. With renewables, variability is extended to the supply side as well, compromising control of generation. Distributed generation—for example, small-scale renewable sources owned and operated by power customers—further complicates the picture.

Better predictive models of renewable generation are needed at all levels; the management of generation and load is now a multi-variable, non-convex optimization problem with numerous constraints, and decision and control schemes are needed that factor in this complexity.

One solution to the problem of intermittent generation is storage, and approaches are being pursued at different scales and with different technologies—from compressed air underground to flywheels to novel battery materials. Electric vehicles, including plug-in hybrids, may also be useful for storage. The inclusion of storage devices brings that many more assets to be coordinated and optimized.

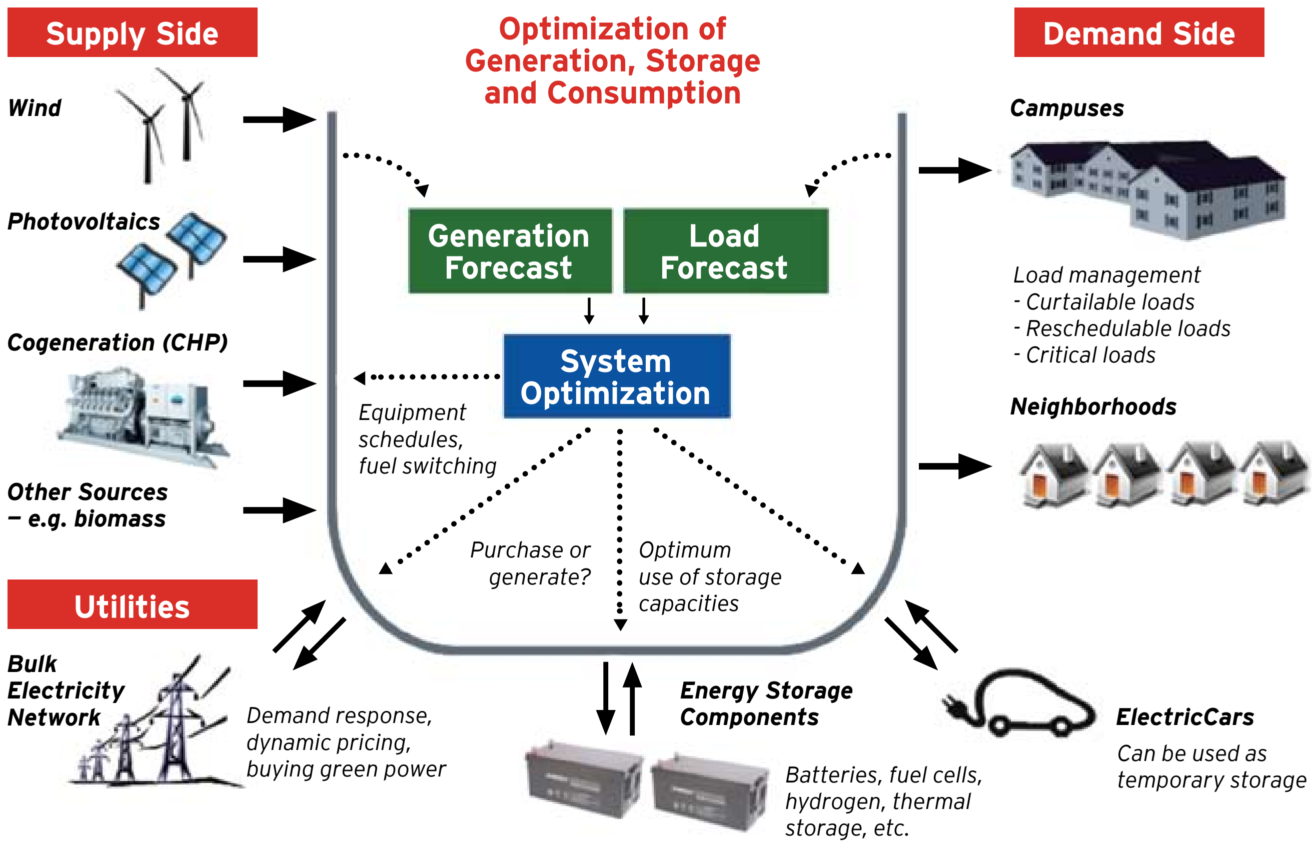


Figure courtesy of Petr Stluka, Honeywell

And there's much more!

The smart grid is an emerging technology domain with numerous sub-problems that require sophisticated treatment of dynamics, modeling, and control. Opportunities also exist in the transmission and distribution area that are not presented here. The time scales range from the milliseconds in which power electronic devices and converters must react to the days or weeks over which some loads must be scheduled (and policy-making with its even longer cycle could benefit from advanced control concepts too).

The bottom line: Control is critical for the smart grid, and there's no shortage of outstanding problems for control scientists and engineers to address.

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.

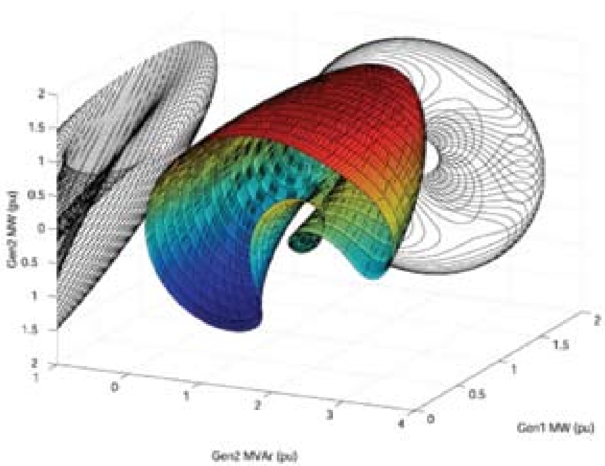
Smart grids are forcing the evolution of grid operational strategies. The variability inherent in large-scale renewable generation challenges existing regulation approaches. Plug-in electric vehicles, if adopted in large numbers, will introduce charging loads that must be carefully coordinated to avoid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

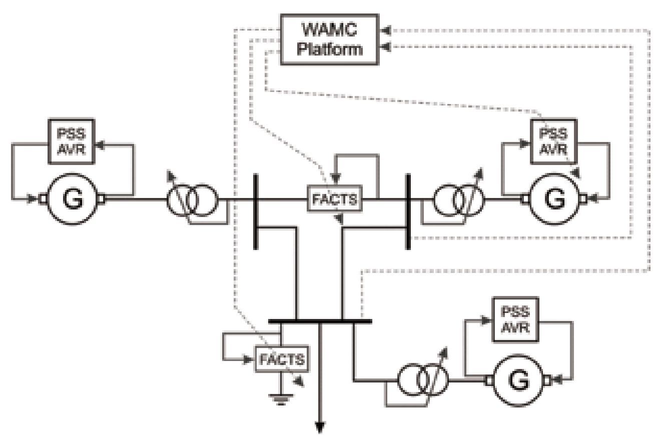
Enhanced grid responsiveness will rely on a range of available and emerging technologies. Phasor measurement units (PMUs) provide fast, accurate, time-stamped measurements that facilitate wide-area monitoring and control. Flexible AC transmission system (FACTS) devices use power electronics to control active and reactive power flows. Load control must be used for grid regulation as well. 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 (PMUs) 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 system-wide events. Possibilities range from enhanced damping of inter-area oscillations to power flow modulation following large disturbances. In order to realize these benefits, however, controller designs must take into account signal latency and reliability.

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





For the example power system on the right, The colored figure shows the boundary of the power flow solution space for all combinations of the active power of one generator and the active and reactive 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, and that cannot be avoided in real-world analysis and control applications. (AVR: automatic voltage regulation; PSS: power system stabilizer)

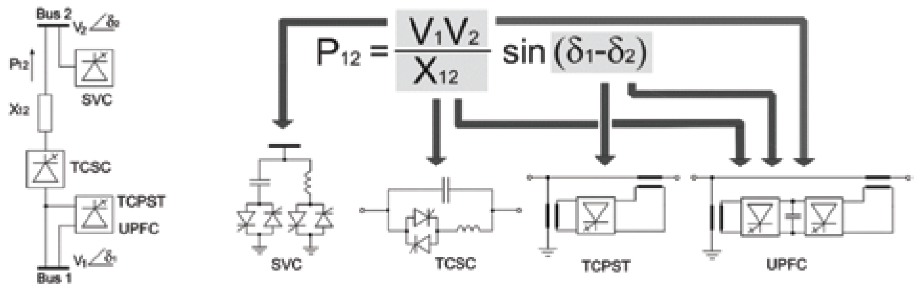
Contributor: Ian Hiskens, University of Michigan, USA

Flexible AC Transmission Systems (FACTS)

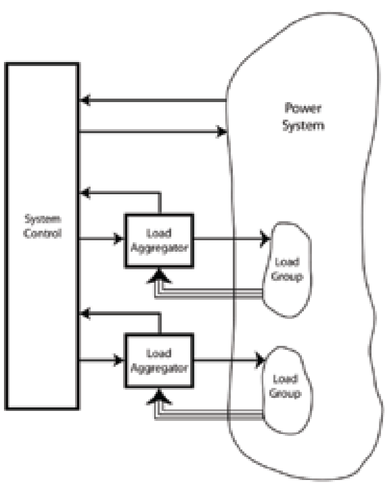
FACTS devices use the switching capability of power electronics to control voltages and currents in an AC grid. The most common FACTS devices are used to regulate bus voltages, for example at the collector bus of a wind farm. FACTS devices are, however, also capable of controlling power flow over transmission 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.

Examples of FACTS devices include static var compensators (SVCs, for regulating voltage magnitudes), thyristor controlled series capacitors (TCSCs, effective line impedances), thyristor controlled phase shifting transformers (TCPSTs, phase angle differences), and unified power flow controllers (UPFCs, 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.



Various FACTS devices



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

Load Control

Power system operation has traditionally relied upon generation to balance supply and demand. 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 huge numbers of autonomous devices. Centralized control seems impractical, with hierarchical control structures more likely to succeed. Many outstanding control questions remain to be addressed.

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 are suited to the analysis and design of large-scale power systems.

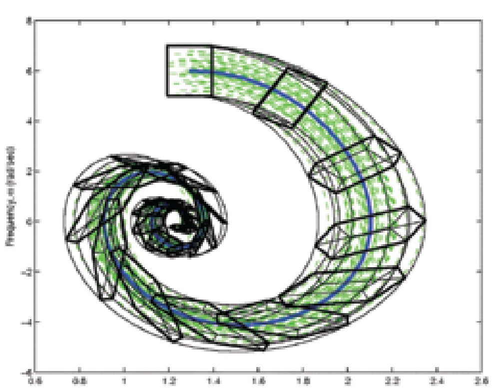


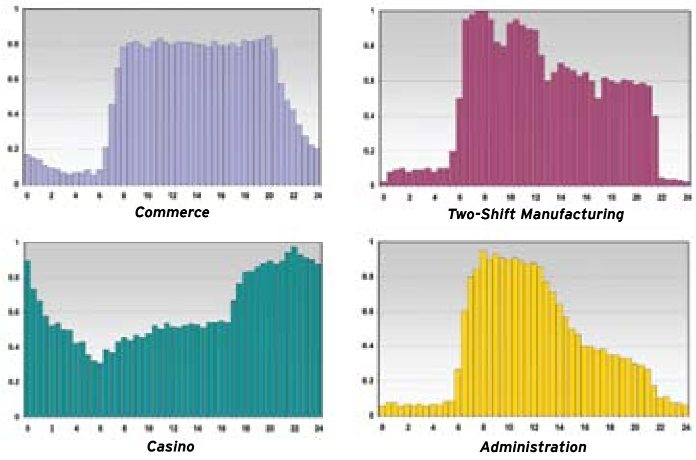
Illustration of phase angle (horizontal axis) and frequency (vertical axis) evolution in a power system showing nonlinear effects of parameter uncertainty. The complete uncertainty set generates a time-varying parallelotope that is mapped along with the nominal trajectory (the blue curve).

Control for Energy-Efficient Buildings

The building sector is responsible for about 40% of energy consumption and more than 40% of 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. New developments in building models and in building automation systems are addressing these and other challenges.

Daily Consumption Profiles: *Every Building Has a Unique Consumption Pattern*



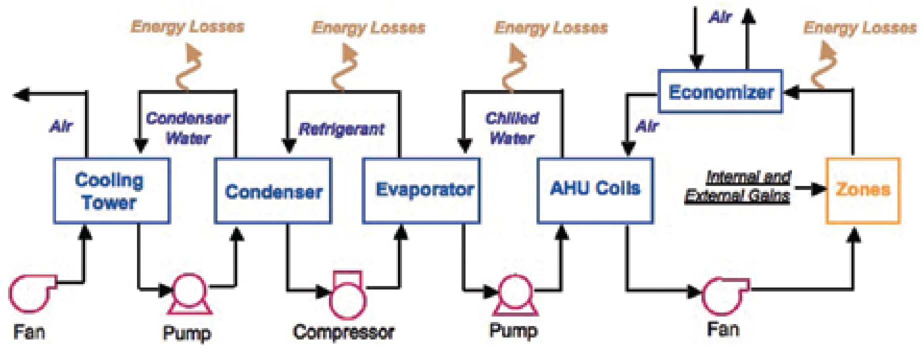
Buildings as Systems of Systems

Buildings are complex systems composed of sub-systems that were traditionally deployed independently of each other. This applies not only to building energy systems, but also to other systems, such as access control, fire protection, and video surveillance.

Heterogeneity. Building components do not necessarily have mathematically similar structures and may involve different scales in time or space. For instance, building control is a complex hybrid problem that includes both continuous and discrete components; major building systems operate in discrete modes (air handlers) or cycle on/off (chiller compressors).

Complex relationships. Building components can be connected in a variety of ways, which may cause complicated dependencies between local and systemwide phenomena. HVAC systems are potentially very large networked systems with complex relationships between the comfort in rooms or zones and operation of individual HVAC components.

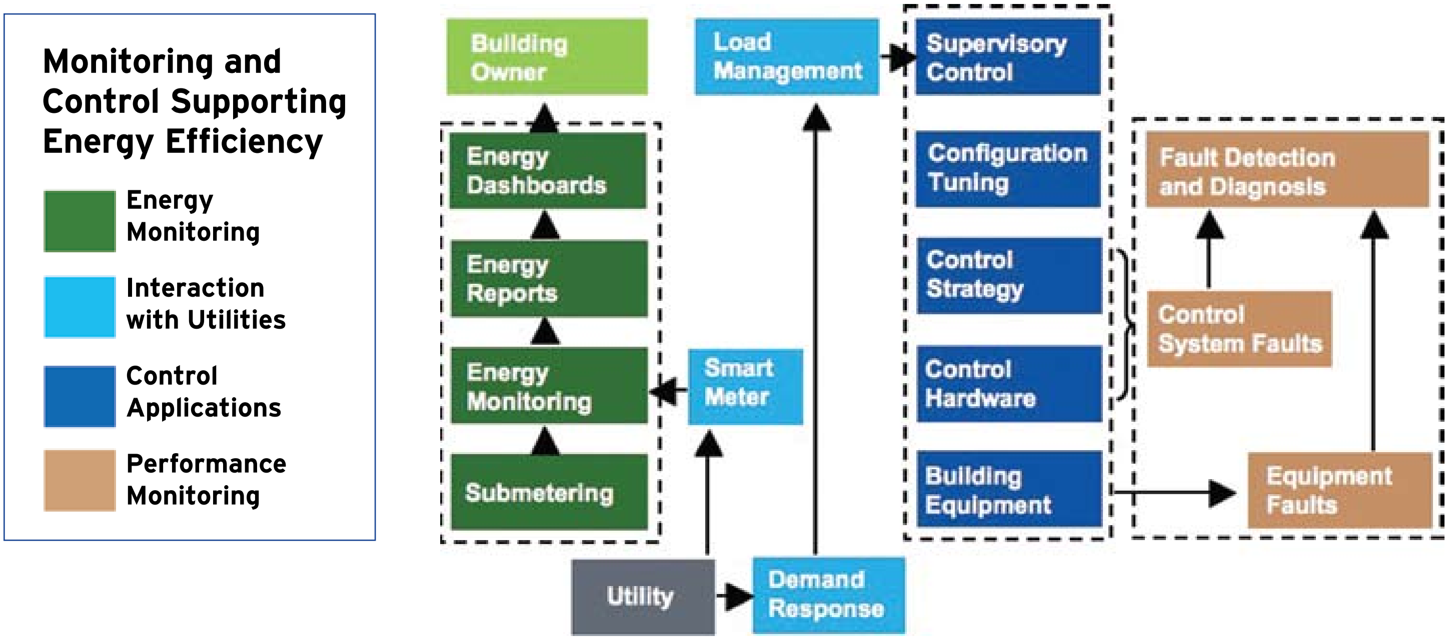
Disturbances. Building operation follows a regular daily cycle. For most office buildings, the two major disturbances are weather and occupancy, and efficient control strategies should take both into consideration. The main challenge is with occupancy that cannot be measured. Other categories of buildings may have additional disturbances caused by their specific principles of operation.



System view of HVAC

***Contributors:** Petr Stluka, Honeywell, Czech Republic and Wendy Foslien, Honeywell, USA*

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.



Challenges and Opportunities in Energy-Efficient Buildings

Entropy is the perpetual enemy of building operations. Building systems degrade over time, and today little can be done about it. Historical operational data is difficult to gather from building management systems, accurate models are generally not available and are hard to develop, and the availability of skilled staff is limited. Recent developments are helping overcome this difficulty: with building information models (BIMs), a standardized way to describe aspects of buildings has become available; most medium-to-large buildings now routinely capture large quantities of operational data; and energy efficiency and load management have become high-priority imperatives. These developments are opening up new opportunities for advanced control in building energy management.

Multivariable supervisory control. The primary goal of building control is to run the HVAC systems to maintain occupants' thermal comfort and system energy efficiency. HVAC control requires adjustments of multiple setpoints, including chilled water temperatures, supply air temperatures, and room air temperatures. Robust multivariable supervisory control strategies need to be developed to enable optimal HVAC operation as well as other building control and optimization applications, such as dynamic load management and dispatching of energy generation, consumption, and storage devices.

Whole-building optimization. Optimization of building energy consumption can be formulated at the whole-building level to cover subsystems such as HVAC, lighting, on-site generation, and storage. The optimization is complicated by disturbances, including weather conditions, occupant behaviors, and prices of electricity and other primary sources of energy. Solving the problem means having to make dynamic decisions on optimal operation of all building energy subsystems. Today this problem is handled by human operators or simplified rule-based logic, but a holistic whole-building optimization approach is needed to address this problem appropriately and achieve maximal cost savings.

Performance monitoring and health management. Physical and control system faults in HVAC systems cause inefficient operation, increased energy use, and reduced equipment life. Many of these problems could be prevented with widespread adoption of automated performance monitoring and equipment health management. The uniqueness of most HVAC implementations, lack of design information, and limited sensors for monitoring have been obstacles that are in the process of being overcome.

Control for Wind Power

Recent experimental assessments show that wind power has the potential to satisfy the global primary energy need. Today, wind power accounts for the largest share of renewable energy generation after hydropower, with 30% global annual growth. Moreover, in the last few years, innovative technologies are being investigated to tap the astonishing power of high-altitude wind using tethered controlled aircraft/kites, with the promise of wind energy at lower cost than fossil sources. Common to both traditional and innovative wind energy concepts are challenging modeling and control design problems. Advances in wind power control will be essential for increasing the penetration of renewable generation and thereby reducing the planet's dependence on fossil sources—a global imperative.

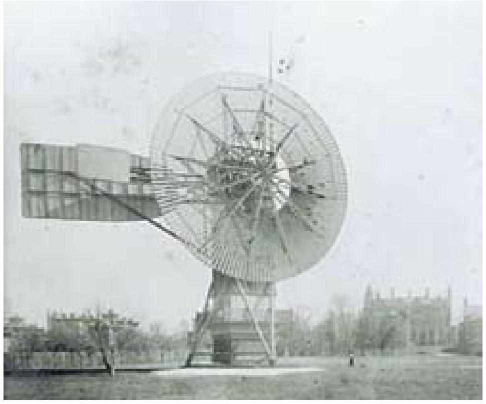
Efficiency Improvement with Control Theory

Today's wind power technology is the result of decades of incremental innovations of wind turbine systems, yet some aspects still require enhanced solutions.

Wind turbines are complex nonlinear systems operating in strong noisy environments with severe constraints on admissible loads. Recent advances developed by the control community in nonlinear modeling, filtering, and control can help realize significant, cost-effective, and safe energy generation improvements.

Adaptive and fast model-predictive control techniques appear to be well suited for the two most critical control problems for wind turbines: blade pitch control and generator torque control. However, the modeling codes available are of questionable accuracy for use in such control design. Data-driven nonlinear modeling and identification approaches could be employed to increase the accuracy of wind tower models. Furthermore, more complex algorithms for blade pitch control based on accurate short-term prediction of wind speed will be fundamental to reducing loads, thus improving reliability and saving material. Finally, distributed control on rotors, using a series of actuators and sensors along each blade, could further reduce loads on the blade, the drive train, and the tower. Such improvements will be even more important for the development of larger offshore wind turbines.

Left: High-altitude wind technology
Right: Wind turbine technology



Wind Energy and Control—The Early Days

The world's first automatically operating wind turbine for electricity generation is attributed to Charles F. Brush, who designed and erected a turbine in Cleveland, Ohio, in 1887. Its peak power production was 12 kW, and it operated for 20 years. Control was critical even then—an automatic control system ensured that the turbine achieved effective action at 330 rpm and that the dc voltage was kept between 70 and 90 V.

Another milestone in wind energy was the 1.25-MW wind turbine developed by Palmer Putnam, also in the U.S., in 1939-1945. This was a giant wind turbine, 53 m in diameter. A hydraulic pitch control system was used for its two blades.

Left: Charles F. Brush's wind turbine, c. 1887

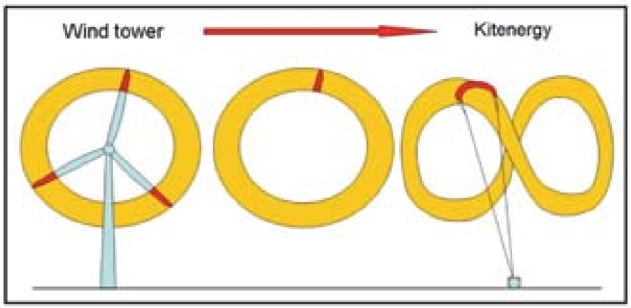
Contributors: Lorenzo Fagiano and Mario Milanese, Politecnico di Torino, Italy

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

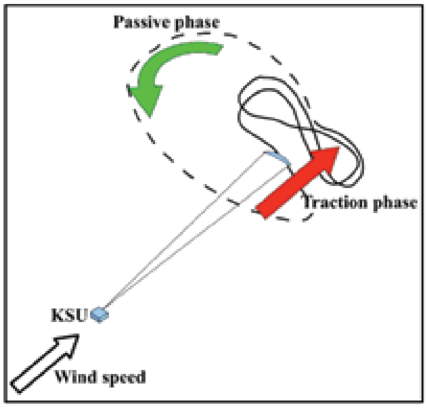
Radical Innovation with Control Theory: High-Altitude Wind Power Systems

In recent years, several university researchers and high-technology companies have been actively working to develop innovative high-altitude wind (HAW) generators. HAW generators aim to harvest wind power at higher altitudes than those reachable by wind turbines, that is, at over 200 m. At these altitudes, stronger and more constant wind can be found almost everywhere.

In all proposed HAW generator concepts, whether rigid wings, kites, rotorcraft, balloons, or other aircraft, control is a key technology. The control system has to maximize the generated energy and satisfy operational constraints while coping with the nonlinear dynamics of the system, the presence of turbulence, and changes in wind speed and direction. Advanced model-predictive control, nonlinear experimental modeling methods, filtering, and sensor fusion techniques will all be important. Furthermore, distributed control strategies will be highly important for the operation of high-altitude wind energy farms, composed of several HAW generators operating in the same site, to avoid interference among the aircraft and to maximize the overall power output.



Example: Kitenergy



One example is a technology developed at Politecnico di Torino called the "Kitenergy" system. The concept is based on wings (for example, power kites like the ones used for surfing or sailing) linked to a kite steering unit (KSU) on the ground. Two lines serve both to control the kite's flight and to convert the aerodynamic forces into electrical power by using suitable rotating mechanisms and electric drives kept on the ground. In one possible configuration, energy is generated by continuously repeating a two-phase cycle. In the traction phase, the kite exploits wind power to unroll the lines, and the electric drives act as generators, driven by the rotation of the drums. The kite is controlled so as to fly fast in the crosswind direction and to generate the maximum amount of power. When the maximal line length is reached, the passive phase begins and the kite is driven in such a way that its aerodynamic lift force collapses; this way, the energy spent to rewind the cables is a fraction (less than 10%) of the amount generated in the traction phase. Theoretical studies, numerical analyses, and experimental results obtained from a KSU prototype show that Kitenergy could bring forth a radical innovation in wind energy, providing large quantities of renewable energy at lower cost than fossil energy (see the table below).

Power Source	*Min. Cost ($/MWh)*	*Max. Cost ($/MWh)*	*Avg. Cost ($/MWh)*
Coal	25	50	34
Gas	37	60	47
Wind turbines	35	95	57
Solar	180	500	325
Kitenergy	**15**	**48**	**20**

Sources for data: "Projected Cost of Generating Energy," IEA Publications, 2008; L. Fagiano et al. IEEE Trans. on Energy Conversion, vol. 25, 2010.

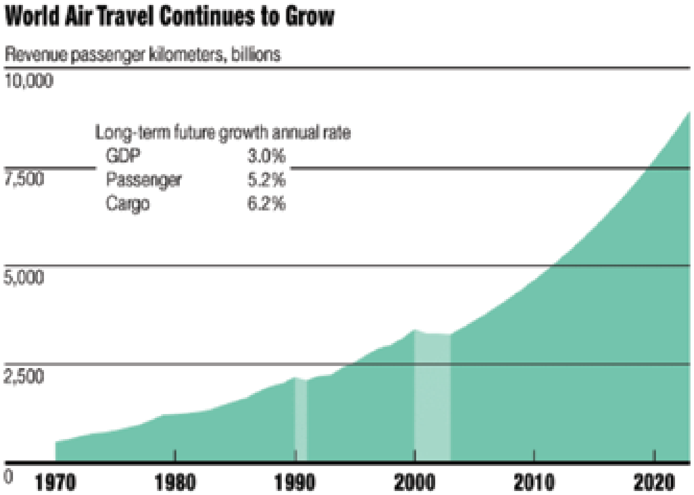
Energy-Efficient Air Transportation

A satellite photo shows the dense coverage of airplane contrails across the Southeast U.S. sky. Contrails may contribute to global warming.

Efficient, robust, safe, and environmentally aware air traffic management is critical to the functioning of the global economy. In the U.S., aviation is responsible for 5% of GDP and an estimated 12 million jobs. The system is already being strained by the current levels of demand, weather disruptions, and volatile fuel prices. Domestic air traffic delays in 2007 cost U.S. airlines an estimated $19 billion in direct operating costs, and the cost to the U.S. economy is estimated at $41 billion. The number of operations in this already congested environment is expected to increase two- to threefold by 2025, posing a new challenge to the effective functioning of the U.S. air traffic system. With similar concerns worldwide, major delays and large economic and environmental impact are inevitable on a global scale unless significant actions are taken. Control and optimization will be key technologies.

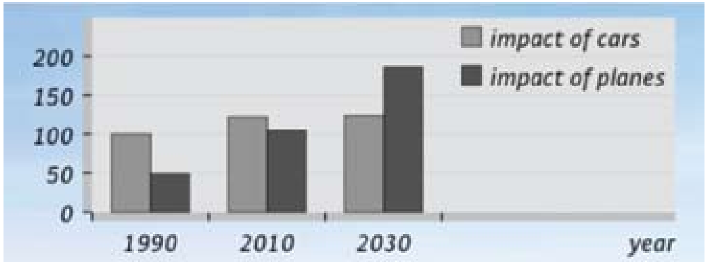
Energy and Environmental Impacts

Despite dramatic increases in aircraft fuel efficiency, the energy requirements of the air transportation system are expected to more than double in the next three decades. Each long-distance flight of a 747 adds about 400 tons of CO_2 to the atmosphere (about the annual per capita emissions in Europe for heating and electricity). Although aviation now consumes only about 13% of transportation-related energy, it is growing rapidly, and the climate impact of emissions at altitude has been estimated to be two to four times greater than reflected by the percentage of carbon emissions. In addition, local noise and emissions and land-use restrictions are limiting the capacity of the transportation system in many areas of the world.



Boeing Current Market Outlook 2004, Demand for Air Travel

"Air travel is the world's fastest growing source of greenhouse gases" (Friends of the Earth). The impact of aircraft may exceed that of cars in the next two decades.



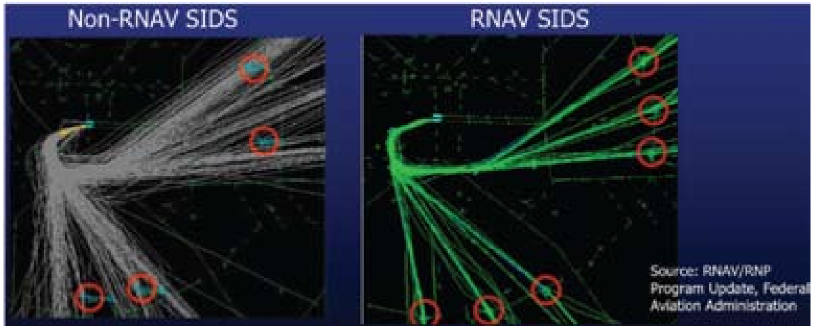
Worldwide comparison cars and planes, impacts on climate change. Scale: cars in 1990 = 100. Source: OECD 2000

Contributors: Juan Alonso, Stanford University, Hamsa Balakrishnan, MIT, Ilan Kroo, Stanford University, and Claire Tomlin, University of California, Berkeley, USA

A Distributed, Large-scale, Multi-objective Control Problem

Active flight management of the entire air transportation system will allow for more reliable operations while respecting constraints on noise and emissions. This concept will involve real-time, massively distributed sensing and modeling to enable path replanning and trajectory optimization. By using local sensing data from other aircraft in a given region of airspace, an aircraft could potentially decrease its fuel burn and emissions by dynamically optimizing for its flight altitude and trajectory. It may also be possible to reduce the formation of contrails by dynamically rerouting to avoid regions of the atmosphere that are saturated with ice. Potential tradeoffs exist between objectives such as fuel burn, operating costs, delays, and system throughput. Multi-objective control techniques for routing are needed that use distributed sensing on board aircraft to simultaneously optimize these various objectives while ensuring safety in the airspace. Market-based schemes are also being envisioned that depend on the impact each airline's operations have on the environment, both individually and through information sharing, to help achieve greener air traffic operations.

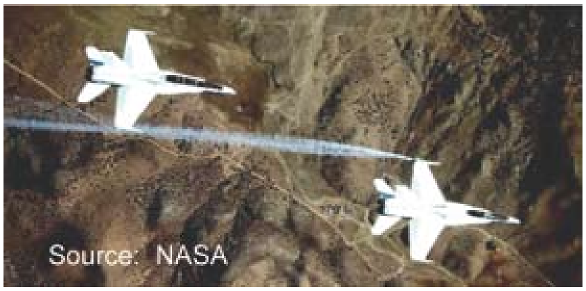
Airports form the critical nodes of the air traffic network, and airport capacity drives overall system capacity. The main constraining factor on airport capacity is the minimum separation mandated between aircraft takeoffs and landings to avoid wake turbulence. Current separation requirements are based on maximum takeoff weight and tend to be overly conservative. Integrating onboard sensors and local weather factors with arrival/departure scheduling algorithms will result in improved runway throughput and decreased delays, fuel burn, and emissions. The same can hold true for optimized airport surface movements using real-time information sharing.



Area navigation (RNAV) standard instrument departures (SIDS) provide structured "lanes" to en route airspace. Benefits include more departures per hour per runway, reduced delays during peak demand, and reduced air/ground voice communication. Annualized benefits are estimated at $39 million for Atlanta International Airport.

Innovations:

- Typical aircraft approaches into airports require a series of short descent and level flight portions. Continuous descent approaches (CDA), which remove the level flight portions, are much more fuel efficient, require less thrust, and reduce engine noise. Noise levels are typically reduced by 5 dB on the ground, and over 4 minutes and almost 400 liters of fuel are saved per approach. Although CDAs are now being implemented in the U.S. and exist in much of Europe, they currently are not used during arrival rushes; more advanced automation is needed to enable conflict-free CDAs in dense traffic.
- Closely spaced parallel approaches could be realized with automated collision avoidance algorithms based on onboard sensing and automatic control. Such approaches would allow parallel runways to be used under more conditions than today, thereby increasing airport capacity.
- Birds achieve significant energy savings by flying in close formation. Formation flight is also being explored for aircraft with similar motivations. Large induced drag, emissions, and noise reduction may be possible.



For further information: www.jpdo.gov; http://www.faa.gov/about/initiatives/nextgen/benefits/environment; http://www1.nasa.gov/centers/ames/greenspace/sustainable-systems.html; http://soe.stanford.edu/research/profiles/energy_alonso_kroo.html; http://web.mit.edu/aeroastro/news/magazine/aeroastro-no3/2006aviationandenvironment.html

Verification, Validation, and Certification of Aerospace Control Systems

Today's Control Systems

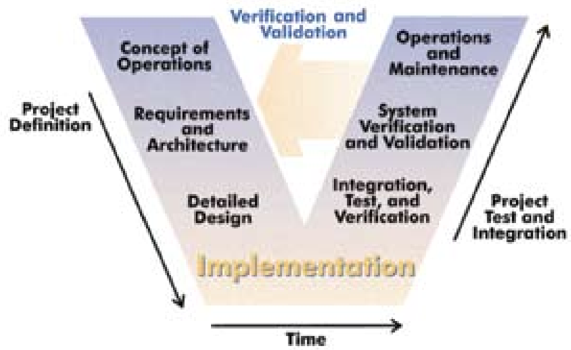
The development of control systems from concept to validation is a complex, multidisciplinary activity. For applications where certification is required prior to operation, specifically aerospace systems, the control laws must go through a rigorous verification and validation process. This process subjects the control laws to a wide variety of analysis and tests to ensure that they will function properly under both nominal and failure conditions. The development process, characterized by numerous iterative design and analysis activities, is lengthy and costly.

Today's safety-critical flight control systems, control laws, software implementation, and tests account for over 60% of the total development cost. (Source: Lockheed Martin Aeronautics Company)

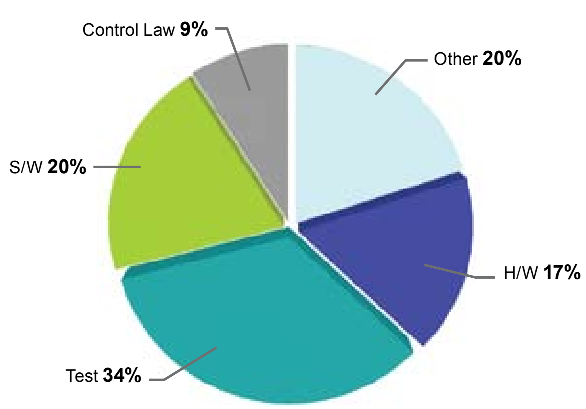
Future Control Systems . . .

For future aeronautics and space applications, the design of the control systems is predicted to include autonomous features such as automated mission planning, target selection and mission replanning, multivehicle cooperative control, situation awareness and automatic collision avoidance, and adaptative and reconfigurable control with diagnostic and prognostic health functionalities.

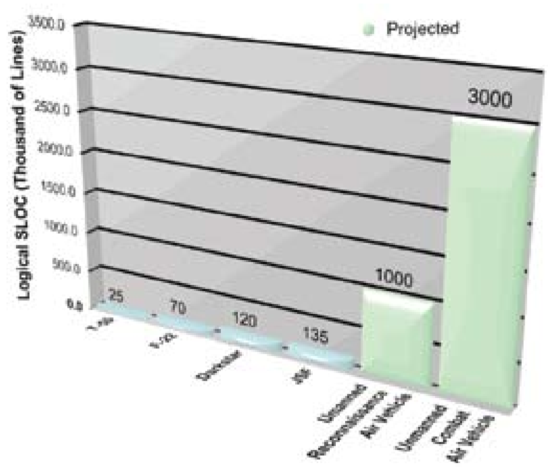
Although the traditional verification, validation, and certification process is producing sufficiently safe and reliable control systems today, it will not be technically adequate and cost-effective for managing the design complexity and safety requirements of future aerospace control systems and for certifying their embedded software. The cost, schedule, and risk impacts are likely to increase exponentially.



A common view of the development timeline
Source: NASA Aviation Safety Program



Typical development cost breakdown for safety-critical flight control systems



Complexity growth, as measured by logical software lines of code, of future safety-critical flight control systems (autonomy)
Source: Lockheed Martin Aeronautics Company

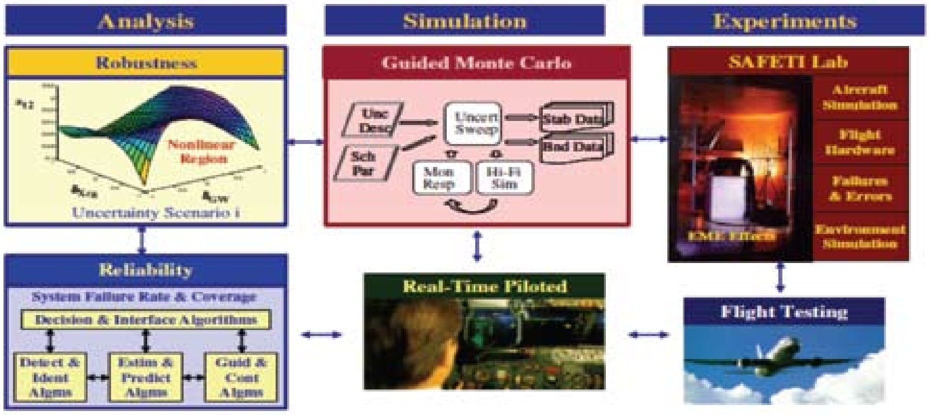
Contributor: Christian Philippe, European Space Agency, The Netherlands

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Closing the Gap Between Future Needs and Certification

The need for engineering tools, methods, and techniques that make V&V cost- and time-effective while applicable to a broad range of aerospace vehicle and airspace systems is widely recognized. An integrated validation process involving analytical, simulation-based, and experimental methods appears to be the most promising approach for coping with the complexity of future control systems. In this process, the analysis results (for example, robustness, worst-case analysis) are used in guided Monte Carlo simulation evaluations and in defining test scenarios for closed-loop real-time simulation with hardware-in-the-loop and flight tests whenever possible.

With regard to worst-case analysis, the use of mixed global/local optimization techniques such as the differential approach augmented with local optimization methods could significantly improve the reliability and efficiency of the current flight clearance process. The use of randomized algorithms and probabilistic methods for robustness design and analysis of control systems affected by random uncertainty could also contribute to the improvement of the traditional V&V process. For the verification and validation of intelligent and autonomous control systems, NASA has proposed a method based on mixing local linear analytical techniques with global random search algorithms.



Integration of analytical and simulation methods for validating vehicle health management and control upset prevention and recovery technologies. Source: NASA

Challenges and Benefits

Control will play an important role in the successful deployment of advanced verification and validation technologies in the aerospace industries. Numerous challenging (and potentially rewarding) problems will need to be addressed by control scientists and engineers.

The main industrial benefits for advanced verification and validation technologies (including theory, methods, and engineering tools) are a reduction of the time-to-market and associated development cost while achieving sufficiently reliable results. Alternatively, V&V technologies could provide increased reliability of the analysis results for reasonable additional effort.

Control of Combustion Instability

Pratt & Whitney FT8 gas turbine

Advances in combustion control are essential for developing engines for propulsion and power generation with high efficiency, increased performance, and low emissions. Lean-burning combustors are needed to meet stringent low-emission requirements and other design criteria, but such combustors are prone to instability.

The potential benefits of lean-burning combustors have been known for decades, as have the challenges. Significant progress was made with passive designs, but achieving the next level of performance improvement and emissions reduction requires integrating active control in the combustor.

From Passive to Active Control of Combustors

Combustion instability results from complex dynamic interactions between acoustics, heat-release, and vortex dynamics. Traditional approaches to controlling instability in combustion turbines have focused on passive mechanisms—no feedback control. One of the earliest examples was the F1 rocket engine, whose initial design, used in the Saturn V rocket, exhibited severe instabilities. Subsequently, a careful, model-based design of baffles helped eliminate the instability.

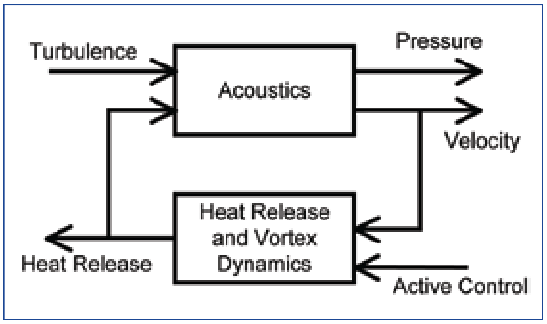
Another example of passive control is the Pratt & Whitney FT8 gas turbine, which features 13 Helmholtz resonators. These passive devices act as vibration absorbers—the geometry of the resonators is chosen so that the absorption is tuned to a particular frequency.

The inherent symmetry of the combustion chamber lends itself to certain natural acoustic modes that are driven into resonance because of the coupling with heat release. Several methods of breaking the symmetry were proven effective in quenching the combustion instabilities.

F1 rocket engine

In active control, incipient instability is dynamically detected and corrected. Active control requires:

- Measurements or estimates of key variables such as pressure and velocity
- Models of the heat release and acoustic phenomena that, in approximated form, can be implemented online
- Actuation capabilities
- Advanced control algorithms that can produce optimized actuation signals based on measurements and models



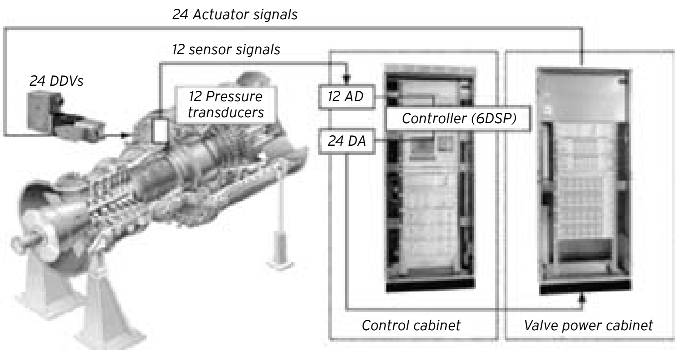
Contributors: Andrzej Banaszuk, United Technologies, Anuradha Annaswamy, MIT, and Sanjay Garg, NASA, USA

The Reality of Combustion Instability

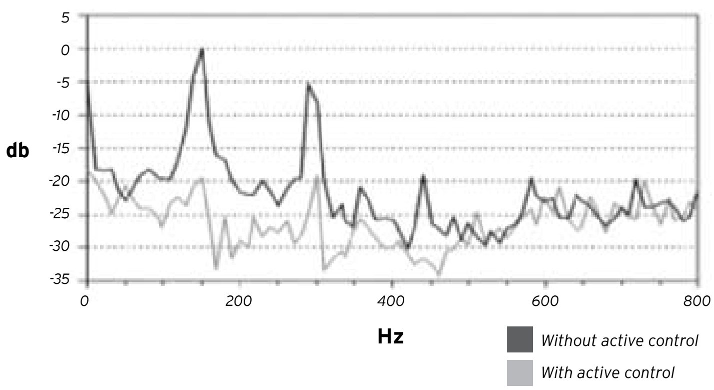
Combustion instability has been a serious problem for gas turbine manufacturers. *The Wall Street Journal* (February 23, 1998) reported the following issues:

- New turbines from a major U.S. and a major European manufacturer, both based on the same design, had to be temporarily shut down after bolts were found to have cracked inside the spinning engines.
- Turbines from another major European manufacturer experienced dangerous levels of vibration. Humming was reported in ring burners with associated flickering gas flames, as well as serious vibration and shaking loose parts. Fixing the faults was a long, sensitive, and costly problem for the manufacturer.
- Yet another major engine manufacturer had bits of the heat shield in their gas turbines break loose, snarling the engine blades.

Similar instabilities were reported in rockets and missiles in the 1960s.



Siemens V94.3A 267-MW gas turbine with control instrumentation



The reduction in acoustic pressure versus acoustic frequency

Active Combustion Control—Progress and Challenges

Fundamental research contributions from numerous academic and industrial organizations have led to a systematic model-based control methodology for instability suppression. Several proof-of-concept demonstrations have been carried out in scaled rigs and fields tests.

- Researchers at NASA GRC and UTRC demonstrated suppression of instabilities in the 300- to 500-Hz range in a single-nozzle combustor rig.
- Research at MIT and the University of Cambridge led to demonstration of instability control in a Rolls Royce RB199 afterburner.
- Instability suppression was demonstrated in the Siemens V94.3A 267-MW gas turbine (see figure above). This design was field-installed and demonstrated successful operation for more than 18,000 hours.
- A range of control design approaches, including advanced linear controllers and adaptive control methods, have been demonstrated successfully.

Major research challenges remain before active combustion control can be commercialized for aerospace propulsion engines:

- Reduced-order models that accurately capture the interactions between acoustics, heat release, vortex dynamics, inlet dynamics, and fuel delivery for a variety of geometries and flame-anchoring mechanisms.
- Controllers that accommodate the intrinsic unstable, infinite-dimensional, uncertain dynamics with limited sensing.
- Actuators such as high-speed valves that can modulate fuel flows in the range of 300 Hz to 1 kHz.

Advanced Driver Assistance Through Massive Sensor Fusion

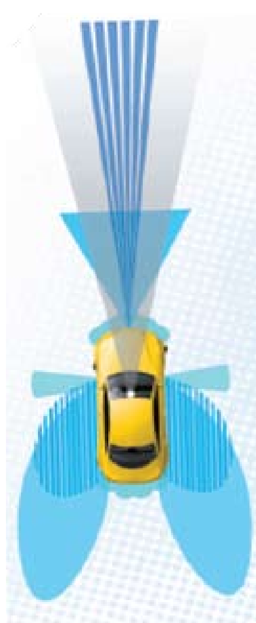
Advanced driver assistance systems (ADASs), as the name suggests, are automotive systems designed to assist in all aspects of driving, including safety, drivability, and fuel economy. The availability of a wide range of sensors, including those in the chassis that measure lateral and longitudinal acceleration, steering wheel angle, yaw rate, and wheel speed, and sensors of the vehicle environment (GPS, infrared sensors, ultrasonic sensors, cameras, and rain, light, solar, and humidity sensors) allow the integration of several sources of information to aid the driver in decision making.

Suitable sensor fusion algorithms must be developed to design a global security system that ensures normal driving commensurate with traffic and weather, and optimizing time, distance, and/or fuel, and driving under emergency situations by commanding the braking system, steering wheel angle, and/or powertrain under unexpected conditions necessitating a fast response.

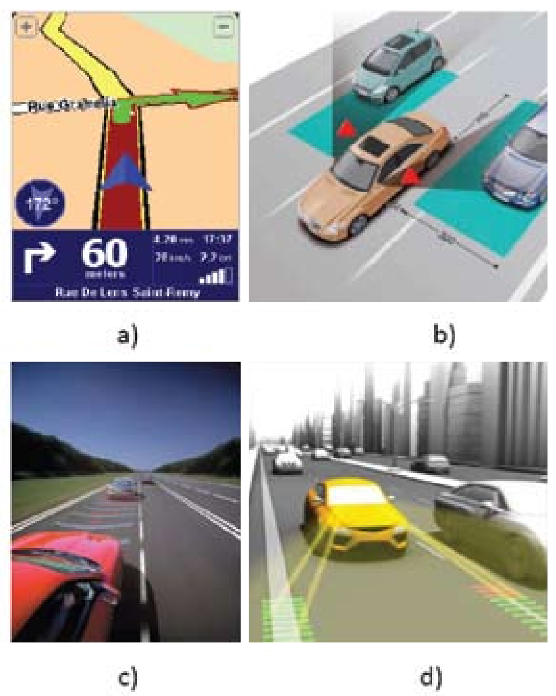
Limited ADAS Functionality Available Today

The low cost and wide availability of these sensors have led to new control strategies such as the navigator (figure a), blind spot detection (figure b), adaptive cruise control (ACC) (figure c), lane keeping (figure d), electronic parking assistance, and the advanced front-lighting system. The antilock braking system and electronic stabilization systems developed over the past 20 years may be viewed as precursors of the ADAS vision.

The majority of ADAS systems currently produce acoustic alarms or suggestions on a display so that the driver closes the loop; in other words, they are decision support systems. In the future, we can expect them to be integrated into the vehicle as active systems. For example, systems are currently being designed to include the information from cameras (such as currently used in ACC or parking systems) in the lateral dynamics controller, for example, to compute the nominal reference trajectory.



Car sensor information equipment with ADAS system (Source: Hella KGaA, Germany)



ADAS examples: a) Tomtom navigator, b) blind spot detection (Mercedes), c) adaptive cruise control (Delphi), d) lane-keeping system (Continental).

Contributor: Luigi Glielmo, Università del Sannio, Italy

From: ***The Impact of Control Technology,*** **T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.**

ADAS systems will also take advantage of new wireless self-powered accelerometers and strain gauges in tires, which will be integrated with pressure/temperature sensors (Figure 1). In fusion with other sensors, these sensors will be able to estimate tire-road friction coefficients and tire forces, the most important variables in active safety systems. Integration with in-vehicle cameras and human body sensors will enable monitoring of driver-vehicle interaction aspects, such as drowsiness of the driver, on the basis of body position, facial posture, eye movements, and the ability to control the vehicle.

Figure 1: Intelligent tire system proposed by Continental

A Role for Control

The main idea of ADAS is to introduce "prediction," a merged "look-ahead" of powertrain, chassis, vehicle environment, and driver's condition obtained through all available information, aimed at semiautonomous behavior of the vehicle in dangerous situations. Control technologies will play a central role in ADAS design.

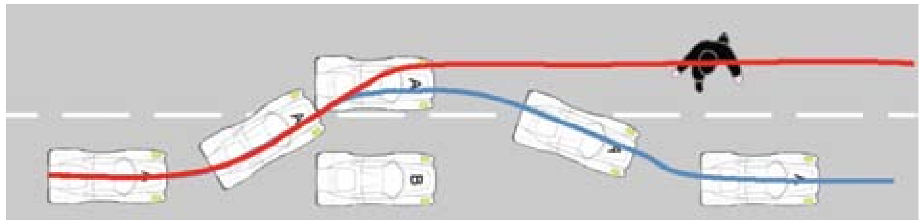
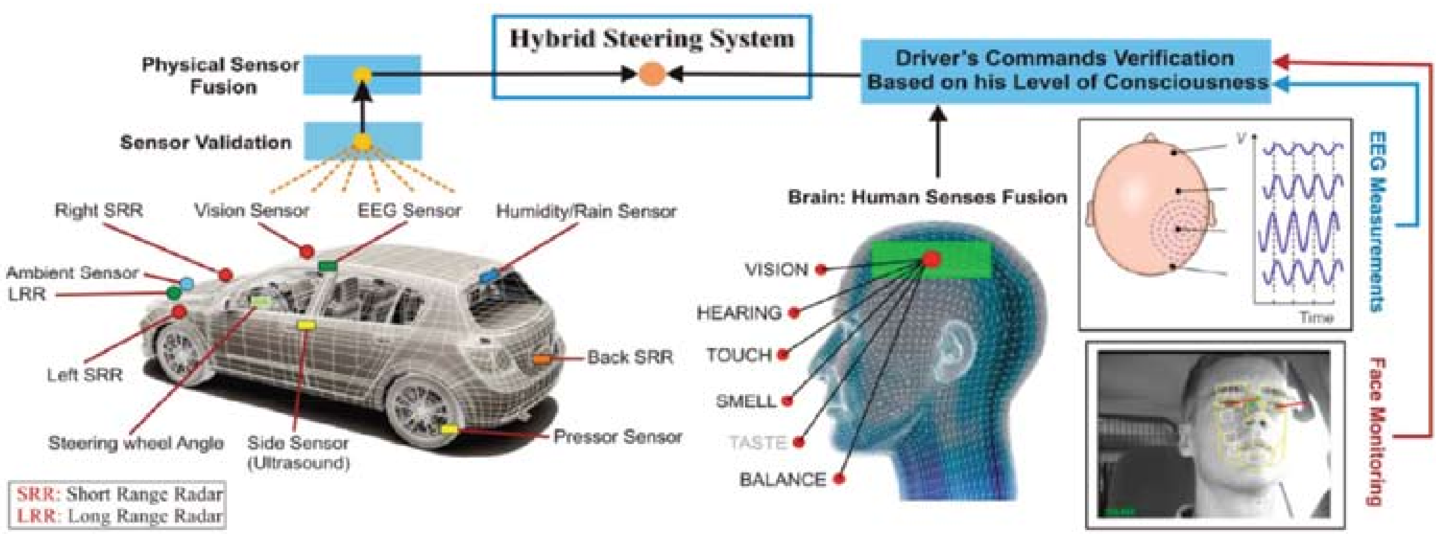


Figure 2

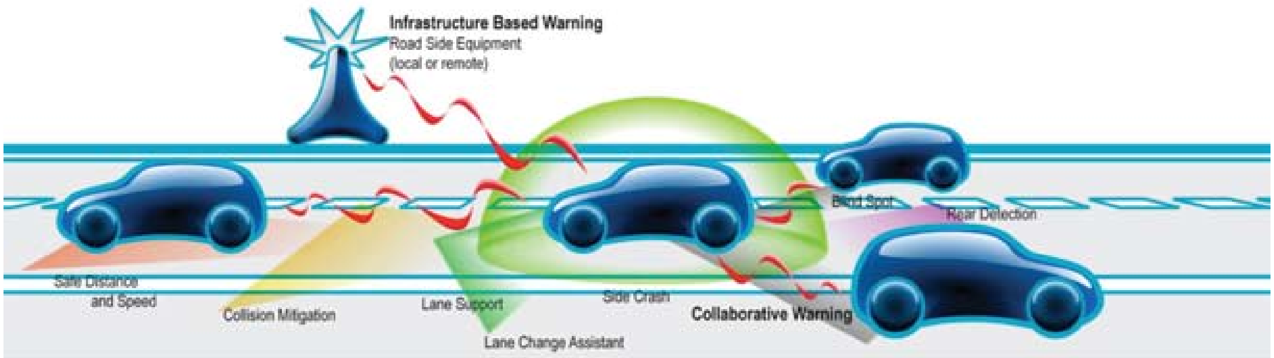
Figure 2 illustrates a possible scenario where vehicle A is performing a lane change and the steering wheel angle imposed by the driver determines a reference trajectory, shown with a red line. The electronic stability program module computes this trajectory, but it cannot take into account the possible presence of an obstacle on the trajectory, such as the pedestrian crossing the street. Cameras and infrared sensors could improve the reference generator module, which could generate a different trajectory that is both feasible and able to avoid the pedestrian, such as the blue one depicted above.

The figure below shows a possible future control loop scheme in which sensor information is merged with driver behavior. To this end, ADAS also includes improved human-machine interface (HMI) systems. The HMI will have bidirectional function in that the driver will take suggestions and alarms while driving and, conversely, will choose among different working options depending on his/her desires (sport driving, family cruise, minimum fuel consumption).



For further information: M. Rezaei, M. Sarshar, and M.M. Sanaatiyan, Toward next generation of driver assistance systems: A multimodal sensor-based platform, International Conference on Computer and Automation Engineering, 2010

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 (see the grand challenge on Advanced Driver Assistance Systems), 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

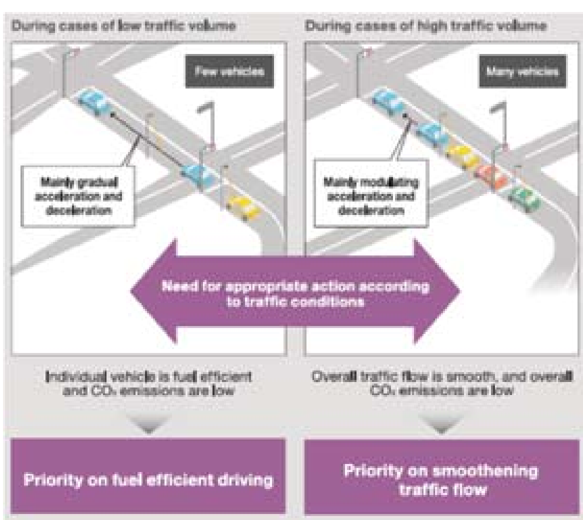
V2I and V2X promise revolutionary improvements in transportation—greater energy efficiency, less road construction, reduced collisions, and safety of vehicle occupants as well as pedestrians and bicyclists. Control is a key contributing discipline for both topics.

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 (measurements 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 onto account the restrictions on automatic vehicle driving imposed by the Vienna Convention on Road Traffic, discussed later).

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



Source: Toyota USA

The figure on the 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

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

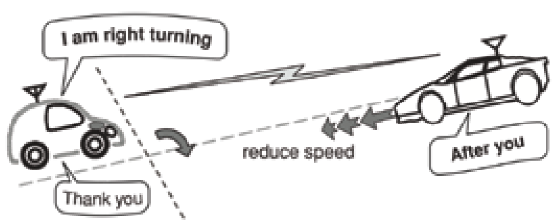
The Vienna Convention on Road Traffic

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 antilock 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 (that is, 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)

Control for Offshore Oil and Gas Platforms

Recent years have seen the formation and growth of the global deepwater offshore industry, which has been driven by increased demand for oil and gas stemming from years of economic growth, reduction in production of existing hydrocarbon fields, and depleting shallow-water reserves. These factors have encouraged operators to invest billions annually chasing 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.

Subsea Production Systems

Subsea systems have to 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.

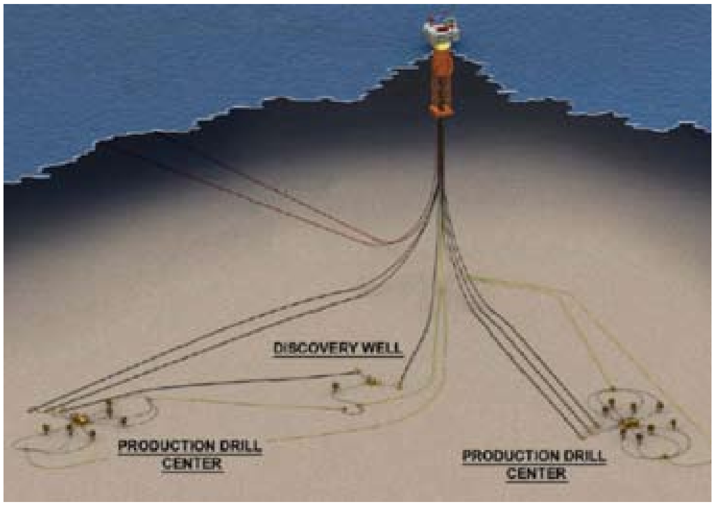
Traditional methods in subsea installation include the use of guidelines or the use of ship dynamic positioning and crane manipulation to obtain the desired position and heading for the payload. Such methods become difficult in deeper waters due to the longer cable between the surface vessel and subsea 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. The control for the dynamic positioning of the subsea payload is challenging due to unpredictable disturbances such as fluctuating currents and transmission of motions from the surface vessel through the lift cable.

A Critical Need for Technology

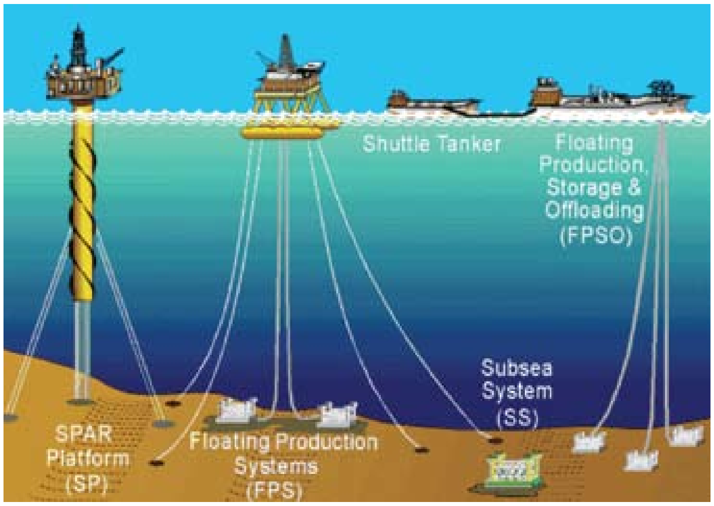
The April 2010 Deepwater Horizon accident in the Gulf of Mexico serves as a reminder of the risks and challenges in offshore operations. In the push toward exploration and production in deeper waters and harsher environments, control theorists and engineers working with colleagues in different disciplines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

A View of the Commercial Subsea System (Wells, Manifold, and Umbilical) on the Seabed



Source: MMS Ocean Science, Nov. 2005

Floating Production and Subsea Systems



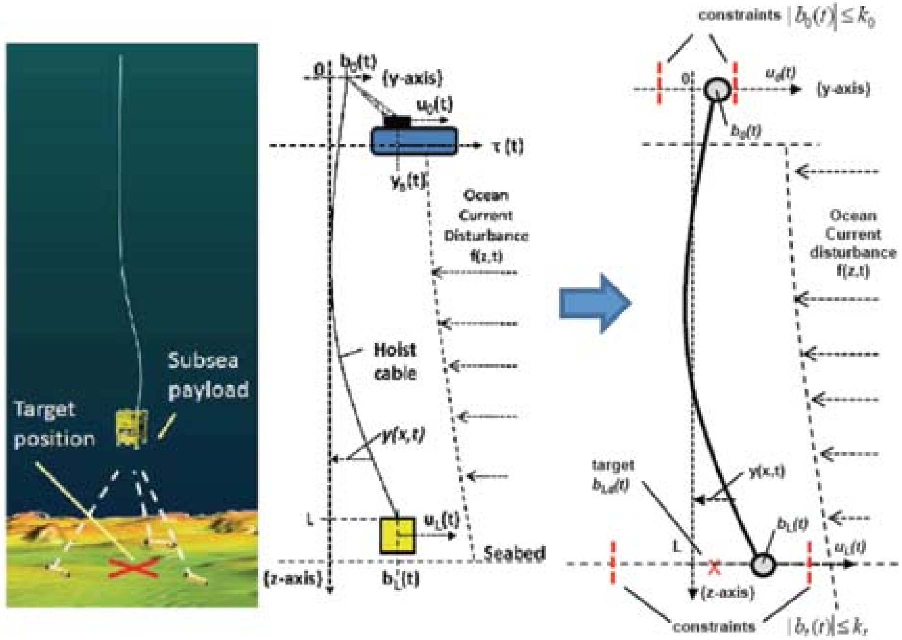
Source: Minerals Management Service, U.S. Department of the Interior

Contributors: Shuzhi Sam Ge, Choo Yoo Sang, and Bernard Voon Ee How, National University of Singapore, Singapore

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

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 control design and operation planning is desirable and challenging.

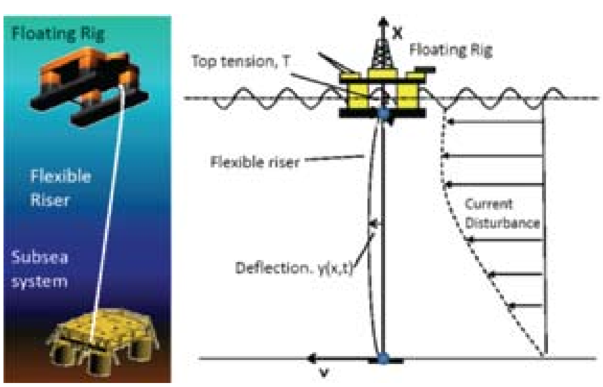


Positioning of subsea hardware using thrusters (left), illustration of subsea positioning (center), and schematic of the installation operation (right)

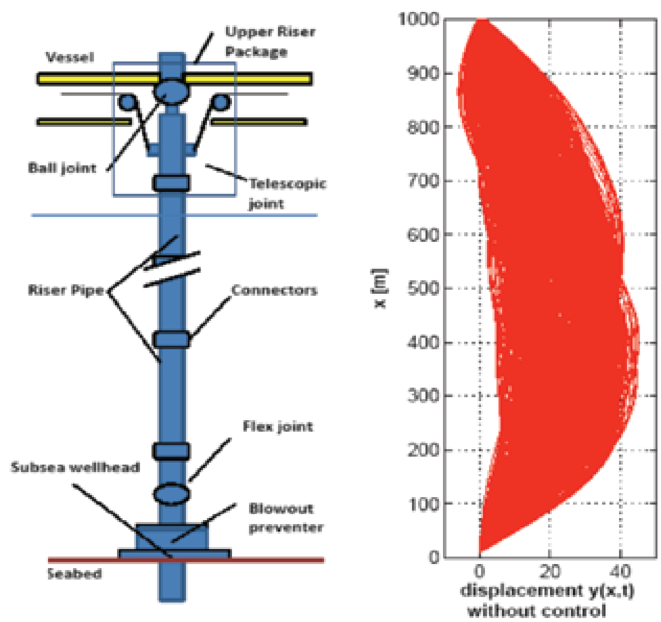
Riser and Drill String Vibration Control

The riser plays a crucial role in offshore oil drilling and production. A marine riser is the connection between a platform on the water surface and the wellhead on the sea floor. A production riser is a pipe used for oil transportation. A drilling riser is used for drilling pipe protection and transportation of the drilling mud. Tension is applied at the top of the riser, which allows it to resist lateral loads, and its effects on natural frequencies, mode shapes, and forced vibration have been studied.

For drilling and workover operations, one objective is to minimize the bending stresses along the riser and the riser angle magnitudes at the platform and wellhead. Hence, vibration reduction to reduce bending stresses and control of the riser angle magnitude are desirable for preventing damage and improving life span.



Schematic of a marine riser



The riser package (left) and an overlay of riser dynamics exposed to current (right)

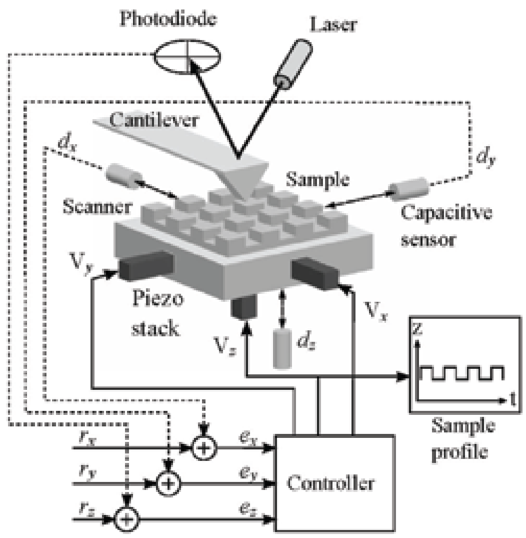
For further information: Social Robotics Laboratory, National University of Singapore, http://robotics.nus.edu.sg; Centre for Offshore Research and Engineering, NUS, http://www.eng.nus.edu.sg/core

Control Challenges in High-Speed Atomic Force Microscopy

The atomic force microscope (AFM) is one of the most versatile methods of imaging structures at nanometer scale. Its 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), the scanning tunneling microscope (STM), and the scanning electron microscope (SEM). Consequently, the AFM has brought about significant progress in a multitude of scientific fields ranging from nanotechnology to life sciences and medicine.

Furthermore, being a "mechanical microscope," the AFM can be used to manipulate matter at nanometer scale as well. Thus, it has emerged as the driving technology in nanomanipulation and nanoassembly and has generated much excitement in nanorobotics.

The AFM's ability to image and manipulate matter at the nanometer scale is entirely dependent on the use of feedback loops. Thus, there are numerous opportunities and a significant need to apply advanced feedback control methods, especially for high-speed AFM.



The figure to the left is a schematic representation of a modern AFM, with a nanopositioning stage and multivariable feedback control. Each axis is driven by a piezoelectric stack actuator. Capacitive displacement sensors measure the scanner's position in three dimensions. The sample topography is measured directly by the interferometer in the vertical direction. In early AFMs, the scanners operated in an open loop. Today, most commercial AFMs are instrumented with displacement sensors, allowing for feedback; however, the feedback controllers used are rudimentary.

The Atomic Force Microscope

When used for imaging, the purpose of an AFM is to characterize a sample by bringing a sharp probe very close to the sample surface and then moving it, relative to the sample, in a raster pattern. This movement is achieved using a nanopositioner such as a piezoelectric tube scanner or a piezoelectrically driven flexure-guided stage.

The probe tip is affected by the forces on the surface, some of which are attractive and some repulsive. These forces cause a deflection of the micro-cantilever on which the tip resides. This deflection is detected using a laser beam that is bounced off the cantilever and back onto a photodetector.

The AFM can be used in various operating modes, broadly classified as "static" or "dynamic." In the static mode, the probe is dragged on the sample surface and a constant force is maintained by the z-axis controller, a PI controller in almost all commercial AFMs. In dynamic modes, the micro-cantilever is oscillated sinusoidally at or close to its resonance frequency, and variations in its oscillations due to the interactions with the sample are monitored to infer sample properties.

The Need for High-Speed AFM

Conventional AFMs are slow, operating at scan frequencies of several Hertz. Consequently, it can take the microscope several minutes to develop an image. Distortion in the AFM image can occur if the surface features being interrogated change rapidly compared to the AFM's operating speed. The image distortion occurs because the measurements at the initial and final pixels of an image are taken at significantly different time instants. Thus, a high-speed AFM is needed to minimize image distortions when the surface or the process being studied, manipulated, or controlled has fast dynamics. For example, AFM imaging of living cells currently takes in excess of 1 minute per image frame. This is clearly too slow to investigate biological processes that occur in seconds. Significant challenges are associated with operating an AFM at high speeds, most of which lead back to feedback control problems.

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

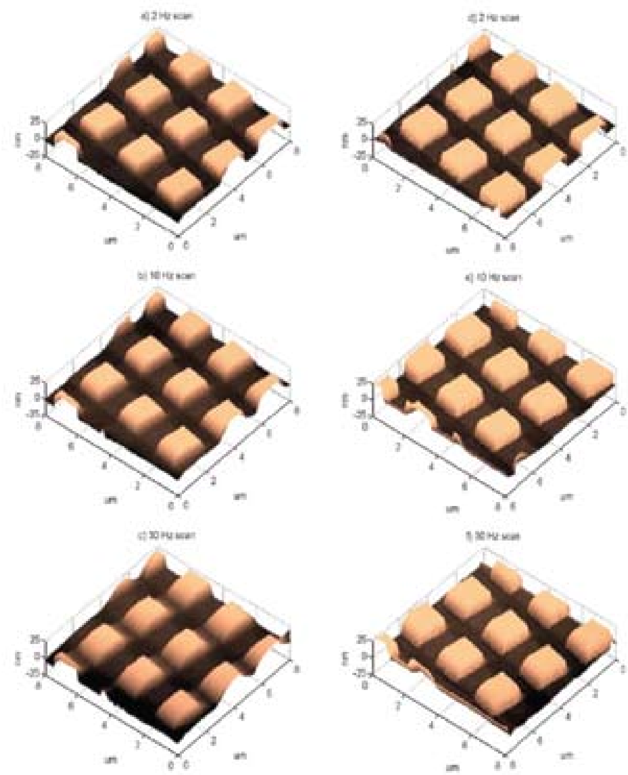
From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Control Challenges...
and Opportunities!

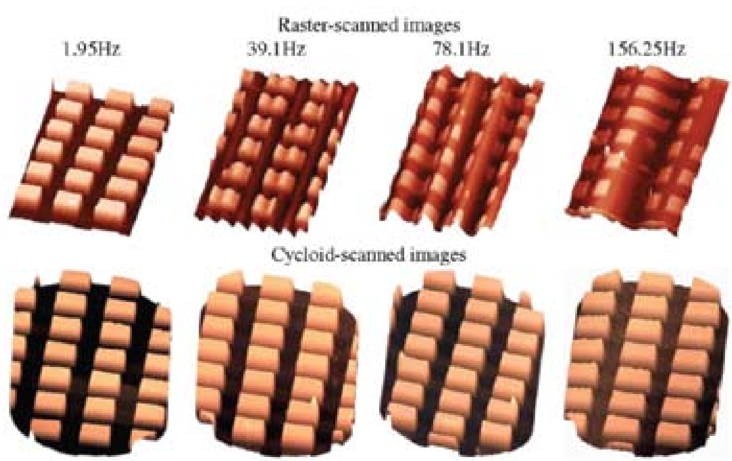
Advanced control is a key technology for high-speed atomic force microscopy, but control designs will need to address several challenges:

- AFM scanners are highly resonant systems. Control design must be informed by properties and parameters of the resonance modes.
- The performance of piezoelectric actuators can degrade over time. Furthermore, they are prone to hysteresis and creep effects. Control designs must be robust to such changes.
- High-bandwidth control is required for positioning accuracy in the AFM scanner. Sensor noise complicates controller realization.
- For high-speed AFM in particular, optimal non-raster-scan methods will be required. Such methods will require further advances in control design.
- The AFM scanner is a multivariable system. Significant cross-couplings exist that cannot be adequately managed with today's PID controllers.
- The vertical axis control loop is especially nonlinear. Conventional linear control methods are inadequate.
- The AFM microcantilever is a highly resonant system, but when operated in a fluid environment, it is prone to significant damping. Feedback can be used to mitigate this problem.

Images of a calibration grating developed on a commercial AFM are shown below. The features are 3 mm apart and have a height of 20 nm. Images (a)-(c) were developed using the AFM's standard control loops at scan frequencies of 2 Hz, 10 Hz, and 30 Hz, respectively; (d)-(f) were developed with an advanced controller that combines positive position feedback to flatten the frequency response of the scanner together with integral action to improve tracking. Both sets of experiments were conducted at the same scan frequencies and under similar conditions.



As illustrated below, significant improvement in tracking can be achieved by using a properly chosen non-raster-scan method. If a scanner is required to follow a cycloid-like trajectory, its lateral and transversal axes must track sinusoidal signals. This is a far less stringent requirement on the controller than tracking triangular signals, as needed in a conventional raster-scanned AFM. An alternative non-raster-scan method is based on the spiral of Archimedes. Control problems associated with closed-loop implementations of both methods are exciting and challenging.



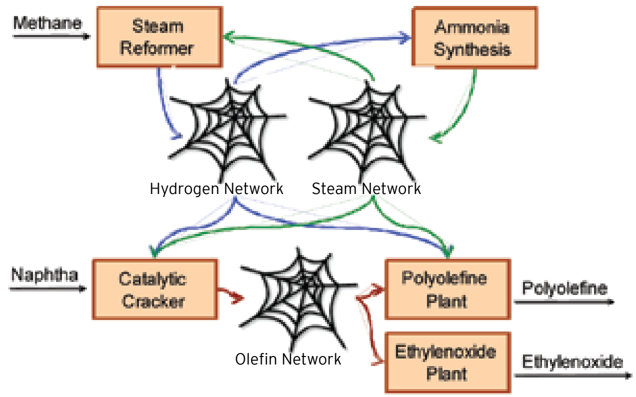
Process Manufacturing Networks

Growing competition on the global scale, the transition from supply-driven to demand-driven markets, and tightening of process safety and environmental regulations are all placing increasing pressure on process manufacturing and operations. Leveraging the full economical 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

Figure 1 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.

Figure 1



Integrated Production

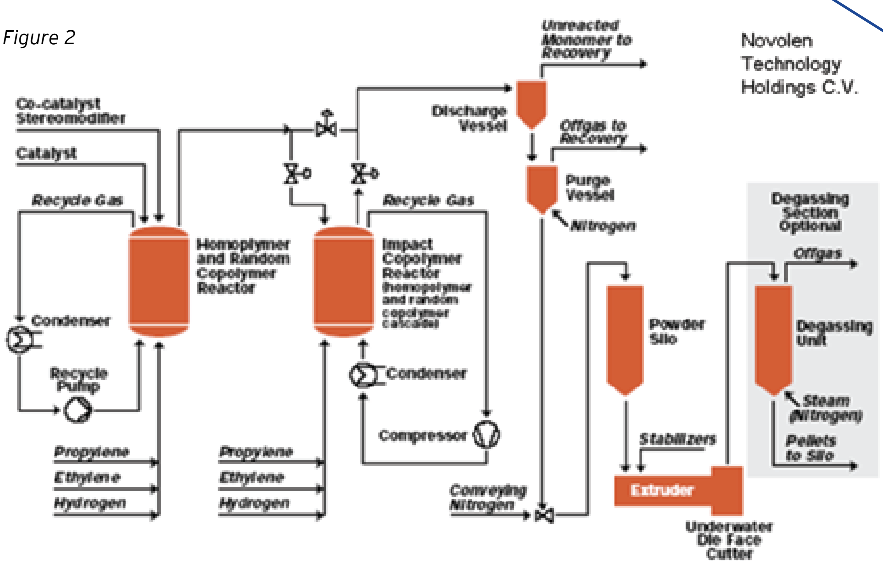
The interconnection between different process plants and between the units of a single plant account for efficient energy integration and for a largely complete recycling of materials. BASF's Verbund concept, for example, implements a tight integration of all chemical plants at one site.

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

Figure 2 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.

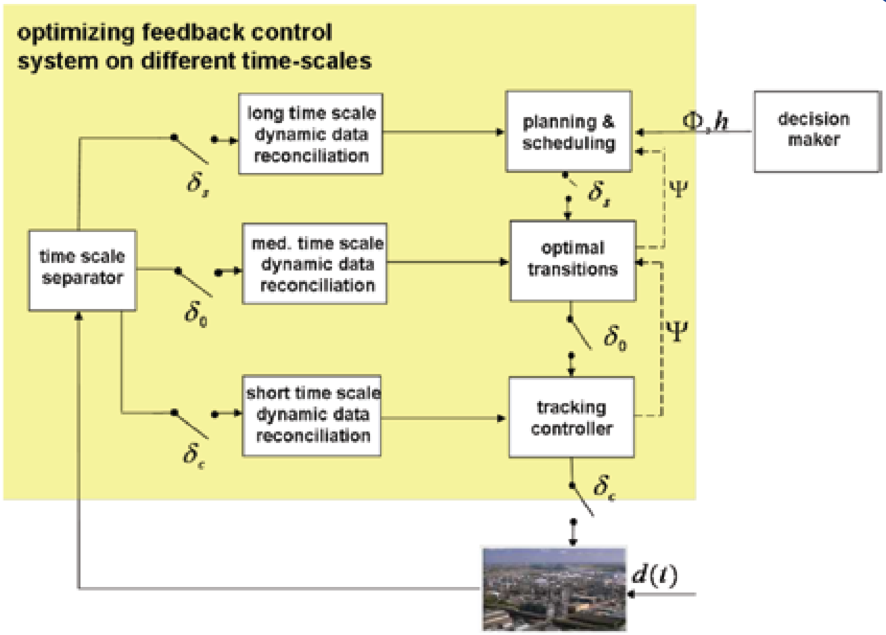
Figure 2



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.

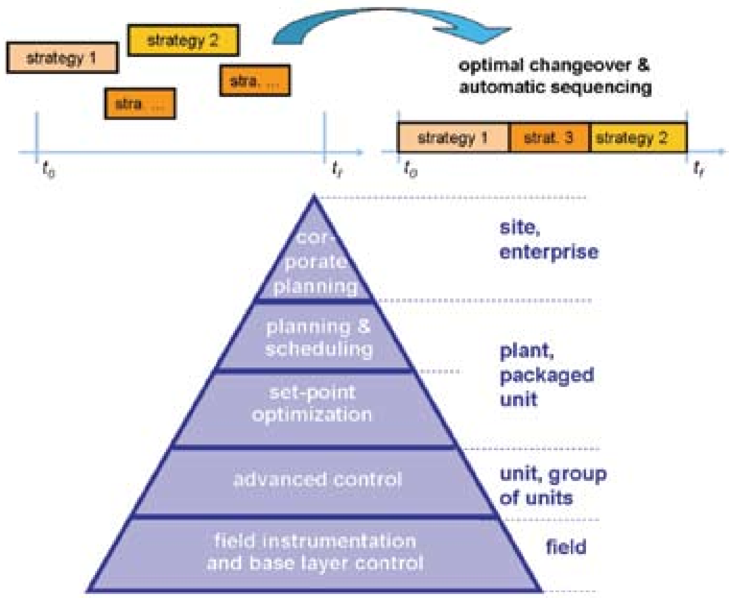


Challenges—Planning and Scheduling

Real-time business decisions relate to operational strategies such as the start-up 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 electrical power networks, to mention only a few.

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.

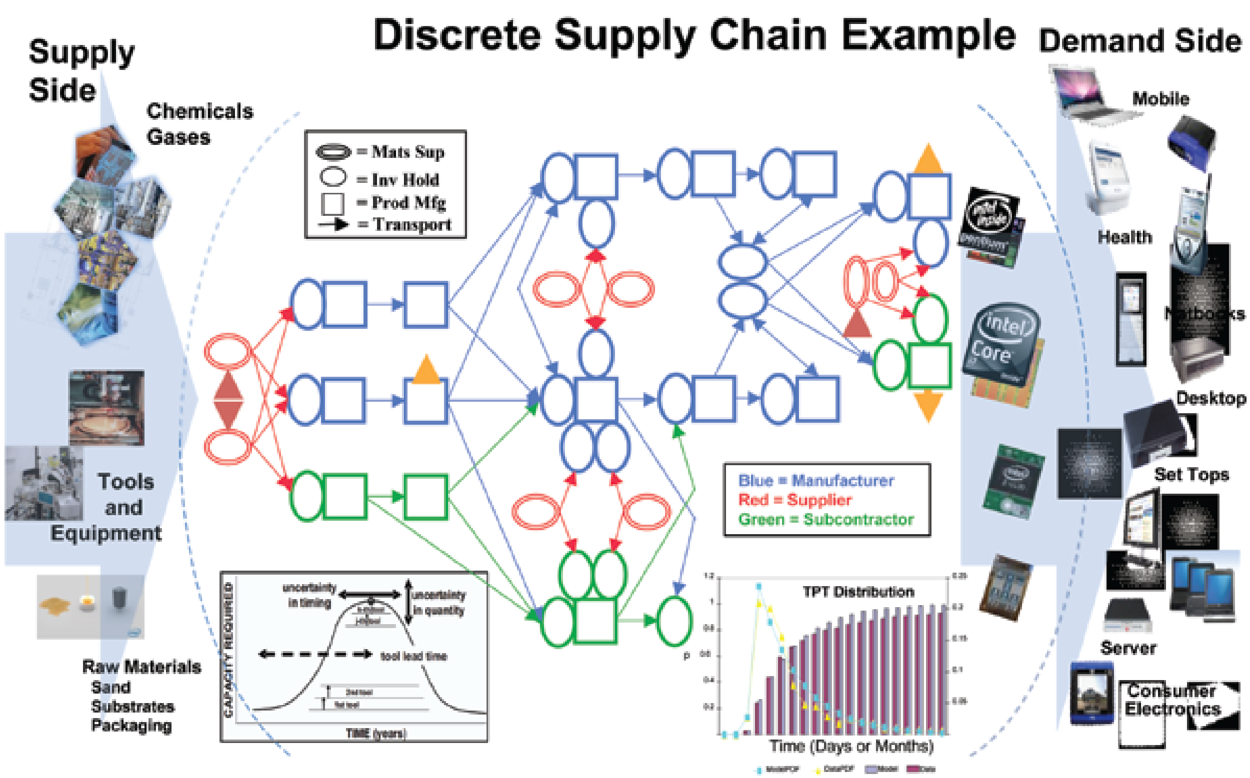
Grand Control Challenge

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.



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 1).

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, share information, and create 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 achieve 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?

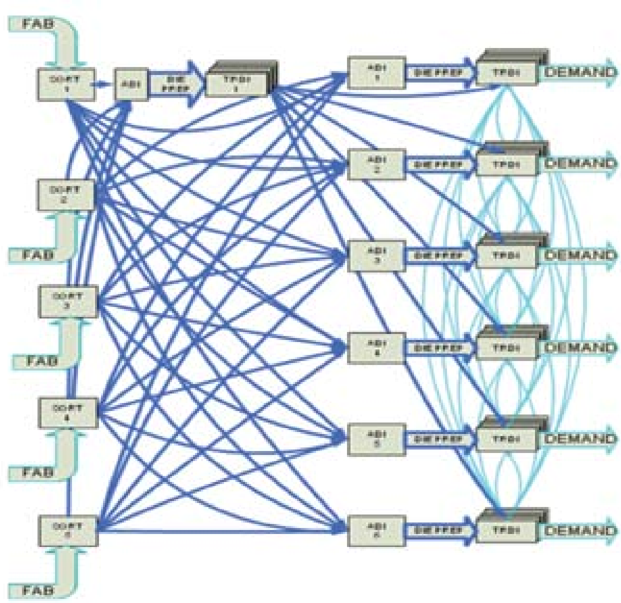
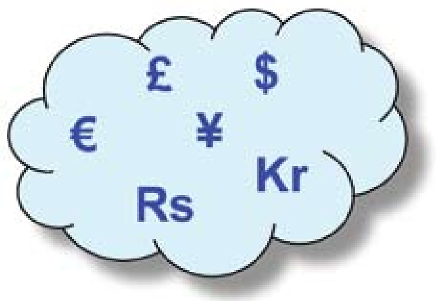


Figure 1: A model predictive control application for inventory and production management.

Financial Engineering Done Right!



A Control Systems Perspective on Financial Crises

Financial engineering is about risk assessment and risk management. At an individual (including institutional) level, the key issues are to assess the risk/reward trade-off of one's investment portfolio and to minimize if not eliminate risk due to factors beyond one's control. At an aggregate (national) level, the challenge for policy makers is to assess the collective risk of the entire economy. Both tasks are closely related and require the development and calibration of suitable statistical models, as well as methodologies based on these models. Armed with such strategies, individuals, retirement fund managers, investment bankers, and policy makers can make well-informed decisions or, in some cases, offer well-informed recommendations. Actually implementing the recommendations, however, will require political will.

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

What Went Wrong?

The Basel II norms for risk assessment and management permitted each institution (such as banks) to use its own internally developed model provided the model afforded adequate risk protection over historical data. Unfortunately, the validation of these in-house models was for the most part based on short-term data, from the benign period from 2002 to 2007. Most of these risk models proved inadequate in assessing the level of risk during the events of late 2007. Even if the models had gone as far back as just 15 or 20 years, it would have been obvious that risk levels were being seriously underestimated.

Certain basic notions such as correlation were ignored in the recent financial crisis. To cite just one instance, there are 50 states in the U.S., about 15 of which are large states. When the risk of mortgage-backed instruments defaulting was estimated, the probability of default in each state was assumed to be an independent random variable! The simple idea that if there is a recession in Michigan, there is likely to be one Florida too seems to have escaped everyone! Many other such simple modeling errors can be pointed out post facto.

Most fundamentally, those tasked with assessing risk must recognize that financial models, no matter how well supported by data and intuition, are not physical laws! There is no "F = ma" in mathematical finance!

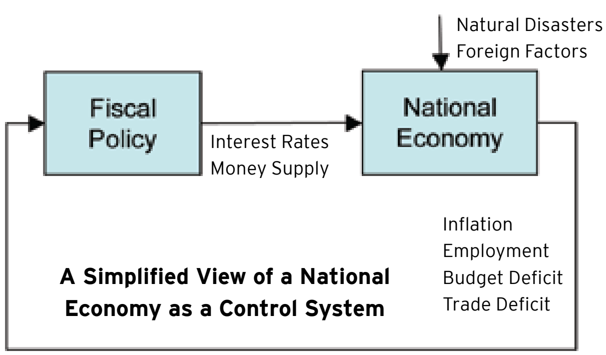
Value at Risk

The financial industry uses the concept of VaR (Value at Risk) as a metric. VaR is the 99th percentile of the probability distribution function (of an individual or institutional portfolio), but can also be applied at a national level. The philosophy is that, using whatever statistical methods we have at our disposal, we can estimate an amount of loss that is likely to be exceeded with a probability of only 1%. This is a useful metric and offers an excellent approach to regulation.

First Targets

A controls perspective has the greatest potential for impact at societal, national, or global levels, rather than in terms of assisting individuals or institutions. In particular, helping governments make informed policy decisions provides the greatest opportunities for advanced control.

For example, accurately assessing the risk of default on household or sovereign debt (which led to problems in the U.S. and Europe) is an area where a controls perspective can have an immediate impact. On the flip side, banning naked credit default swaps would probably do more to stabilize the bond market than any highbrow controls strategy, but that requires political courage.



A Simplified View of a National Economy as a Control System

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Heavy-Tailed Distributions

Another problem is that of “heavy tails.” When we try to fit probability distributions on values that are far beyond those usually observed, the standard Gaussian distribution seriously underestimates the risk of extreme events. So we need different kinds of laws of large numbers to study heavy tails. Such theories already exist in the probability theory community, but they have not been applied in the financial community to estimating risk or to crafting suitable regulations based on the risk assessment. To give one specific example, the Black-Scholes formula for valuing options assumes a geometric Brownian motion model. This is unrealistic—data from around the world over several decades have shown that actual asset prices have heavy tails.

The controls community can act as a bridge between the worlds of pure probability theory and the financial sector (including regulators). We can help regulators assess the level of risk in the finance system. We can warn that the credit markets are getting overheated or that the risk of certain banks defaulting is beyond acceptable levels.

Figure 1 illustrates the difference between a heavy-tailed distribution (in red) and a Gaussian distribution. Figure 2 shows a histogram of the S&P 500 fund index (daily logarithmic returns over a 17-year period) along with a normal distribution fit.

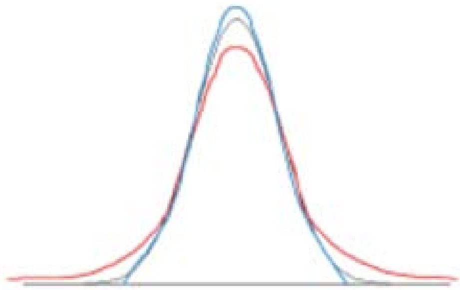


Figure 1

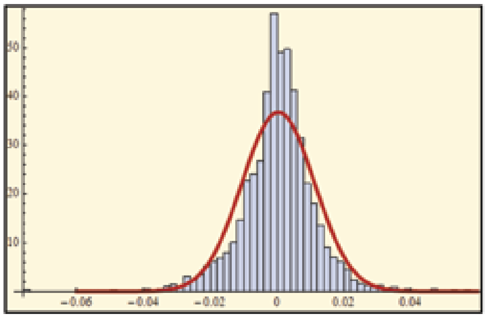


Figure 2

http://www.mathestate.com/tools/Financial/wp1/MarketTheory.html

The Importance of System Identification

If we view financial systems as dynamical systems under feedback, we can clearly see that control technologies have a substantial role to play in financial engineering.

The relevance of system identification especially bears emphasis. “First-principles” models of economic systems are generally unavailable or unreliable; models must be identified from data. However, the levels of uncertainty, the presence of delays, even the possibility that the system we’re trying to model is nonstationary . . . such factors create challenges that are often not encountered in usual control engineering domains.

Control Challenges in Mobile Telecommunications

Mobile telecommunications technology is having an unprecedented impact on human society. Currently, there are more than 4 billion cellular subscribers worldwide; some 2 million new phones are sold each day and 80 billion email messages are sent! Global revenue exceeds $4 trillion annually. Also, new services are 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 many challenges. For example, the control problems in telecommunications have their own distinctive characteristics, including different demands on data rate and delay latency (Figure 1). Also, the control is necessarily carried out over the telecommunication channel itself, giving rise to networked control issues.

Control in Telecommunications

Control appears at various levels in mobile telecommunications (Figure 2):

Inner-loop power control: Used to adjust the signal-to-interference ratios of users so that they are maintained at an appropriate level at the base station. This loop operates at 1500 Hz and is quantized to 1 bit with delays of several samples.

Outer-loop power control: Used to adjust the signal-to-interference target so that the block error rate reaches a desired value. This loop operates at a slower rate (approximately 40 ms).

Scheduling: 3G and 4G systems allow for high uplink (between user and base station) data rates, which is achieved by giving users the opportunity to use increased transmit power. This loop operates at a relatively slow rate (2 to 10 ms with delays to 40 ms). There is also a scheduler in the downlink.

Future systems: More complex control problems will arise in future systems, such as multicarrier scheduling for LTE and advanced scheduling in cognitive radio. In the latter case, uncertainty is expected to become even more significant as spectrum availability will also be uncertain.

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.

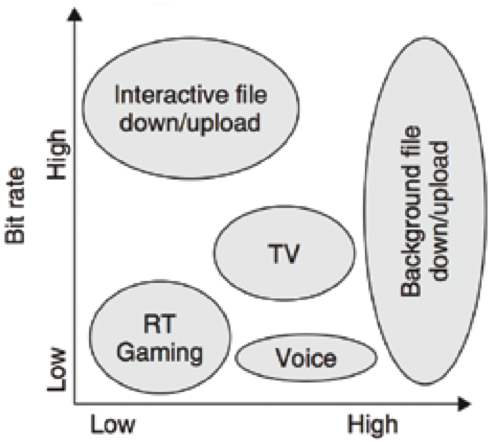


Figure 1: Bit rate/delay issues for mobile services

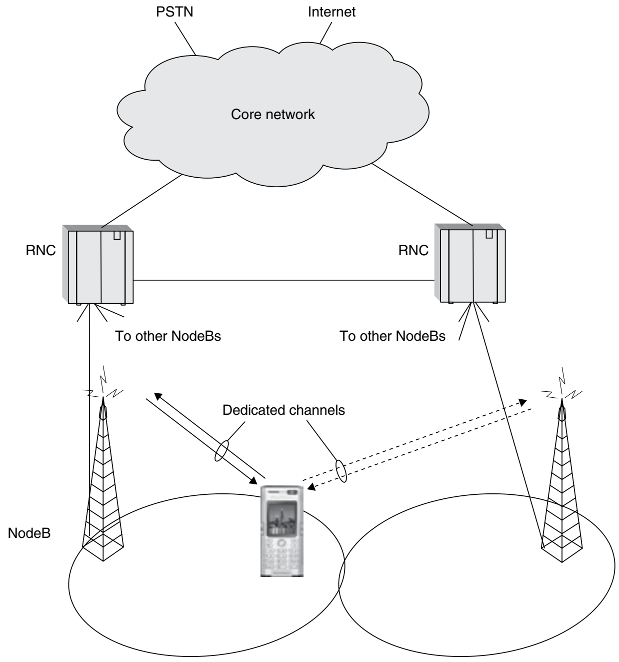


Figure 2: Typical 3G topology

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

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

Challenges

Some of the many challenges associated with these telecommunication control problems include:

Inner power control loop

- Heavily quantized control (1 bit)
- Delays
- Lost control signals
- Highly variable channel fading
- Significant nonlinearities
- Multivariable interactions (each user is an interference source on every other user)
- Decentralized information pattern

Scheduling

- Large and variable delays
- High uncertainty in channel gains
- Unused grants (users may already be at maximum power or may have exhausted their data)
- Interactions with neighboring cells (a neighboring cell can issue a "relative grant" to tell a user to turn down its power; this is not known by the "serving cell")

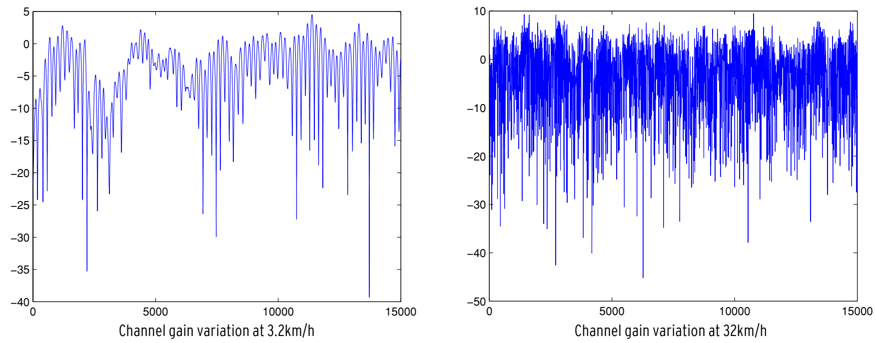
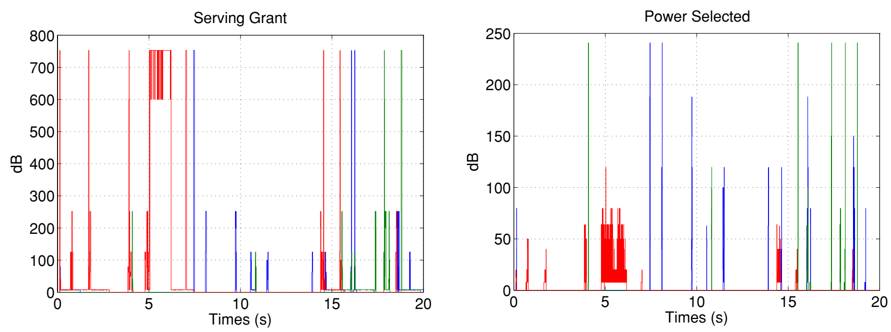
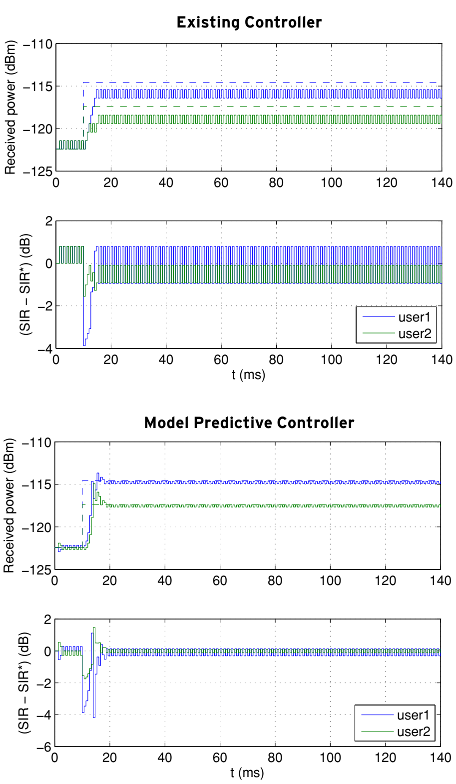


Illustration of fading at different user speeds



Simulation of allocated grants (serving grants) and used grants (selected power)

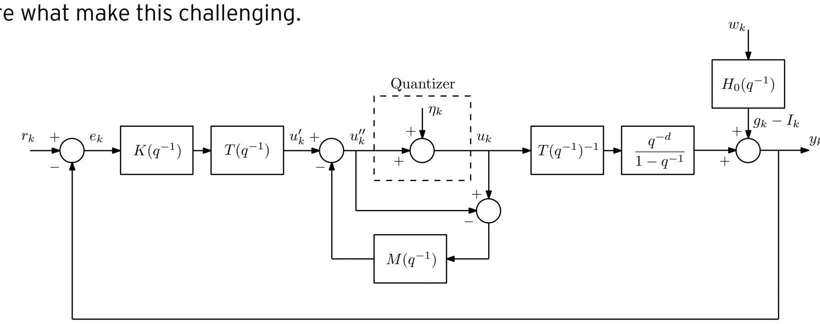
Opportunities for Advanced Control



Improved inner-loop power control achieved via use of model predictive control

The control challenges in telecommunications suggest that gains can be achieved using sophisticated control tools. However, the application of these tools in the telecommunications context raises new, and as yet unsolved, challenges.

- New ideas in networked control are needed for the inner power loop. This is challenging because only 1 bit can be sent per sample and bits can 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 (up to 50 potential users), and short sampling periods (40 ms) 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. Here the challenge is due to high state dimension, severe nonlinearities, and fast sampling rates.
- New insights into decentralized control are needed to implement the solution. The stochastic nature of the problem and high demands on quality of service for users are what make this challenging.



Three-degree-of-freedom inner power control loop

Appendices

Appendix A. "Impact of Control" Berchtesgaden Workshop

Anuradha Annaswamy

The idea for the workshop came about during my sabbatical at the Institute for Advanced Study (IAS), Technische Universität München (TUM), during 2008-2009. As part of the Hans Fischer Senior Fellowship offered me by IAS, funds were available to hold a workshop in my area of research. Motivated primarily by a desire to showcase the activities in control theory and applications to a research group dominated by physicists and biologists, the illustrious and enterprising Martin Buss and I began to brainstorm. Rather than limit the topic to specific control areas or applications, Martin and I developed the idea of covering control at large. Serendipitously, at the same time, the then president of the Control Systems Society, Tariq Samad, together with Gary Balas, had initiated a Task Force on Outreach, one recommendation of which was to hold a workshop where success stories and grand challenges and the overall impact of advanced control would be presented. The four of us joined forces, and with multiple sponsors from both Europe and the United States, and strong support from Patrick Dewilde, Director of IAS at TUM, we put together an international workshop on the "Impact of Control: Past, Present, and Future," held October 18-20, 2009, at the InterContinental Berchtesgaden Resort, Berchtesgaden, Germany. The sponsors included TUM-IAS, Cognitive Technical Systems (CoTeSyS), Deutsche Forschungs-Gemeinschaft, and the FeedNetBack and DISC projects from Europe and IEEE-CSS, the National Science Foundation, and the Institute for Systems Research from the U.S.

Seventy leading experts from all over the world representing academia, government, and industry attended the workshop. A range of topics related to the broad impact of control were discussed: the successes of advanced control in practice, new and emerging control technologies, grand challenges for the future, research opportunities, and barriers to technology transition.

The workshop, held over two and a half days, explored the impact of control from two distinct viewpoints. The first was applications, on the basis of which the participants were grouped into seven sessions: Aerospace, Automotive, Biological Systems, Manufacturing Automation & Robotics, Networked Systems, Process industries, and Renewable Energy & Smart Grid. After a day of deliberations, the groups made their presentations summarizing the control achievements, grand challenges, and research opportunities in their particular domain of application. The second approach addressed the workshop topic with a thematic flavor. Related breakout sessions were organized on the following topics: Application & Market Requirements, Cognitive Control, Controls Education, Implications for Research Communities, Outreach & Visibility, and Tools & Platforms. Following extended deliberations, session chairs presented key issues and recommendations related to their topic.

The workshop also featured plenary lectures by Peter Terwiesch, Chief Technology Officer, ABB; Karl Åström, IEEE Field Medal Winner, Lund Institute of Technology; and Alkis Konstantellos, Deputy Head, Embedded Systems and Control, European Commission, on industrial, academic, and government perspectives, respectively. A panel discussion was held at the close of the workshop addressing final thoughts and comments of the participants. The workshop agenda is included below.

Significant preparation was undertaken prior to the workshop to help accomplish the ambitious agenda. Given the broad scope of topics and content, care was taken to ensure several aspects: selecting participants who have played a leadership role in their domain, communicating guidelines to these

From: *The Impact of Control Technology*, T. Samad and A.M. Annaswamy (eds.), 2011. Available at www.ieeecss.org.

participants in terms of questions that needed to be addressed in session deliberations, and identifying session chairs to assemble and engage each group in an extensive dialog that addressed these questions. These preparations helped the workshop participants to “hit the ground running” and arrive at consensus on the impact of control, key achievements, opportunities, and recommendations.

The Berchtesgaden Workshop Participants, Intercontinental Hotel, October 2009

International Workshop on the Impact of Control: Past, Present, and Future

InterContinental Berchtesgaden Resort, Berchtesgaden, Germany

October 18-20, 2009

Sunday, October 18, 2009

14.00	Welcoming Remarks
14.30	Hike
17.30	Panel Discussions
18.30	Reception
19.30	Dinner

Monday, October 19, 2009

9:00	Welcoming Remarks
9:10	Plenary 1: Peter Terwiesch
9:50	Breakouts I Aerospace Automotive Biodevices Manufacturing automation & robotics Networked systems Process industries Smart grid & renewables
12:00	Lunch
13:30	Breakouts I: Interim presentations
15:30	Breakouts I (contd.)
18:00	Plenary 2: Karl J. Åström
19:30	Dinner

Tuesday, October 20, 2009

OUTREACH DAY

9:00	Plenary 3: Alkis Konstantellos
9:30	Breakouts I: Final Presentations
11:00	Breakouts II (topics to be finalized)
13:00	Lunch
14:30	Breakouts II: (contd.)
17:00	Breakouts II: Final Presentations
19:00	Panel Discussions/Wrap-up
20:00	Dinner (optional)

Berchtesgaden Workshop Agenda

Appendix B. Contributors to the Report

Chris Aldrich
University of Stellenbosch, South Africa
ca1@ing.sun.ac.za

Juan Alonso
Stanford University, U.S.A.
jjalonso@stanford.edu

Anuradha Annaswamy
Massachusetts Institute of Technology, U.S.A.
aanna@mit.edu

Karl Johan Åström
Lund University, Sweden
kja@control.lth.se

Radhakishan Baheti
National Science Foundation, U.S.A.
rbaheti@nsf.gov

Gary Balas
University of Minnesota, U.S.A.
balas@aem.umn.edu

Hamsa Balakrishnan
Massachusetts Institute of Technology, U.S.A.
hamsa@mit.edu

Johann Bals
German Aerospace Center
johann.bals@dlr.de

Andrzej Banaszuk
United Technologies, U.S.A.
banasza@utrc.utc.com

Wayne Bequette
Rensselaer Polytechnic Institute, U.S.A.
bequette@rpi.edu

Bo Bernhardsson
Lund University, Sweden,
bo.bernhardsson@control.lth.se

John Bosworth
NASA, U.S.A.
john.t.bosworth@nasa.gov

Richard Braatz
Massachusetts Institute of Technology, U.S.A.
braatz@mit.edu

Martin Braun
Intel, U.S.A.
martin.w.braun@intel.com

Martin Buss
Technical University of Munich, Germany
mb@tum.de

Ken Butts
Toyota, U.S.A.
ken.butts@tema.toyota.com

Eduardo Camacho
University of Seville, Spain
eduardo@esi.us.es

Carlos Canudas-de-Wit
CNRS, France
carlos.canudas-de-wit@gipsa-lab.inpg.fr

Mauricio Cea
University of Newcastle, Australia
mauricio.cea.g@gmail.com

Lorenzo Ciani
Danieli Automation, Italy
l.ciani@dca.it

Ian Craig
University of Pretoria, South Africa
icraig@postino.up.ac.za

Francesco Cuzzola
Danieli Automation, Italy
f.cuzzola@dca.it

Munther Dahleh
Massachusetts Institute of Technology, U.S.A.
dahleh@mit.edu

Raffaelo d'Andrea
Kiva Systems, Switzerland
rdandrea@ethz.ch

Elom Domlan
University of Alberta, Canada
elom.domlan@ualberta.ca

Francis J. Doyle III
University of California–Santa Barbara, U.S.A.
doyle@engineering.ucsb.edu

David Eldredge
General Electric, U.S.A.
david.eldredge@ge.com

Sebatian Engell
Technical Universityof Dortmund, Germany
sebastian.engell@bci.tu-dortmund.de

Dale Enns
Honeywell, U.S.A.
dale.enns@honeywell.com

Heinz Erzberger
NASA, U.S.A.
heinz.erzberger@nasa.gov

Lorenzo Fagiano
Polytechnic University of Turin, Italy
lorenzo.fagiano@polilto.it

Riccardo M.G. Ferrari
Danieli Automation, Italy
r.ferrari@dca.it

Wendy Foslien
Honeywell, U.S.A.
wendy.foslien@honeywell.com

Kathrin Frankl
RWTH Aachen University, Germany
kathrin.frankl@avt.rwth-aachen.de

Masayuki Fujita
Tokyo Institute of Technology, Japan
fujita@ctrl.titech.ac.jp

Eduardo Gallestey
ABB, Switzerland
eduardo.gallestey@ch.abb.com

Mario Garcia-Sanz
Case Western Reserve University, U.S.A.
mario@case.edu

Sanjay Garg
NASA, U.S.A.
sanjay.garg@nasa.gov

Sam Ge
National University of Singapore, Singapore
samge@nus.edu.sg

Helen Gill
National Science Foundation, U.S.A.
hgill@nsf.gov

Luigi Glielmo
University of Sannio, Italy
glielmo@unisannio.it

Graham Goodwin
University of Newcastle, Australia
graham.goodwin@newcastle.edu.au

Tore Hägglund
Lund University, Sweden
tore@control.lth.se

Jurgen Hahn
Texas A&M University, U.S.A.
hahn@tamu.edu

Vladimir Havlena
Honeywell, Czech Republic
vladimir.havlena@honeywell.com

Sandra Hirche
Technical University of Munich, Germany
hirche@lsr.ei.tum.de

Ian Hiskens
University of Michigan, U.S.A.
hiskens@umich.edu

Alexander Horch
ABB, Germany
alexander.horch@de.abb.com

Paul Houpt
General Electric, U.S.A.
houpt@research.ge.com

Bernard Voon Ee How
National University of Singapore, Singapore
bernardhow@singapore.com

Biao Huang
University of Alberta, Canada
biao.huang@ualberta.ca

Luigi Ianelli
University of Sannio, Italy
luigi.iannelli@unisannio.it

Karl Kempf
Intel, U.S.A.
karl.g.kempf@intel.com

Maryam Khanbaghi
Corning, U.S.A.
khanbaghm@Corning.com

Alexander Knoll
EADS, Germany
alexander.knoll@eads.com

Ilya Kolamnovsky
University of Michigan, U.S.A.
ilya@umich.edu

Alkis Konstantellos
European Commission, Belgium
alkis.konstantellos@ec.europa.eu

Kalmanje Krishnakumar
NASA, U.S.A.
kalmanje.krishnakumar@nasa.gov

Ilan Kroo
Stanford university, U.S.A.
kroo@stanford.edu

Katrina Lau
University of Newcastle, Australia
k.lau@newcatle.edu.au

Eugene Lavretsky
Boeing, U.S.A.
eugene.lavretsky@boeing.com

Boris Lohmann
Technical University of Munich, Germany
lohmann@tum.de

Joseph Lu
Honeywell, U.S.A.
joseph.lu@honeywell.com

Massimo Maroni
Alenia Aermacchi, Italy
massimo.maroni@tiscali.it

Wolfgang Marquardt
RWTH Aachen University, Germany
wolfgang.marquardt@avt-rwth-aachen.de

Tom McAvoy
University of Maryland, U.S.A.
mcavoy@umd.edu

Lalit K. Mestha
Xerox, U.S.A.
lalit.mestha@xerox.com

Richard Middleton
National University of Ireland, Maynooth, Ireland
richard.middleton@nuim.ie

Mario Milanese
Polytechnic University of Turin, Italy
mario.milanese@polito.it

Duane Morningred
Honeywell, U.S.A.
duane.morningred@honeywell.com

Richard Murray
California Institute of Technology, U.S.A.
murray@caltech.edu

Ravi Nath
Honeywell, U.S.A.
ravi.nath@honeywell.com

Babatunde Ogunnaike
University of Delaware, U.S.A.
ogunnaike@udel.edu

Robert Osterhuber
EADS, Germany
robert.osterhuber@eads.com

Brad Paden
University of California–Santa Barbara, U.S.A.
paden@engineering.ucsb.edu

Markos Papageorgiou
Technical University of Crete, Greece
markos@dssl.tuc.gr

Ioannis Papamichail
Technical University of Crete, Greece
ipapa@dssl.tuc.gr

Thomas Parisini
Imperial College London, UK
t.parisini@imperial.ac.uk

Robert Parker
University of Pittsburgh, U.S.A.
rparker@engr.pitt.edu

Christian Philippe
European Space Agency, Netherlands
christian.philippe@esa.int

Eftratos Pistikopoulos
Imperial College London, UK
e.pistikopoulos@imperial.ac.uk

S.O. Reza Moheimani
University of Newcastle, Australia
reza.moheimani@newcastle.edu.au

Michael Rinehart
Massachusetts Institute of Technology, U.S.A.
mdrine@mit.edu

Tariq Samad
Honeywell, U.S.A.
tariq.samad@honeywell.com

Choo Yoo Sang
National University of Singapore, Singapore
cvecys@nus.edu.sg

Sigurd Skogestad
Norwegian University of Science and Technology, Norway
sigurd.skogestad@ntnu.no

Malcolm Smith
Cambridge University, UK
mcs@eng.cam.ac.uk

Kirk D. Smith
Intel, U.S.A.
kirk.d.smith@intel.com

Mark Spong
University of Texas at Dallas, U.S.A.
mspong@utdallas.edu

Gregory Stewart
Honeywell, U.S.A.
greg.stewart@honeywell.com

Petr Stluka
Honeywell, Czech Republic
petr.stluka@honeywell.com

Nina Thornhill
Imperial College London, UK
n.thornhill@imperial.ac.uk

Claire Tomlin
University of California–Berkeley, U.S.A.
tomlin@eecs.berkeley.edu

Andras Varga
German Aerospace Center, Germany
andras.Varga@dir.de

Mathukumalli Vidyasagar
University of Texas at Dallas, U.S.A.
m.vidyasagar@utdallas.edu

Mark Whorton
Teledyne Brown Engineering, U.S.A.
mark.whorton@tbe.com

Torbjörn Wigren
Ericsson, Sweden
torbjorn.wigren@ericsson.com

Kevin Wise
Boeing, U.S.A.
kevin.a.wise@boeing.com

Yutaka Yamamoto
Kyoto University, Japan
yy@i.kyoto-u.ac.jp

Ying C. Yeh
Boeing, U.S.A.
ying.c.yeh@boeing.com

Jie Yu
Shell, U.S.A.
j.yu@shell.com

Aircraft and spacecraft, process plants and factories, homes and buildings, automobiles and trains, cellular telephones and networks . . . these and other complex systems are testament to the ubiquity of control technology. At the same time, decades of successful applications have hardly exhausted the potential or vitality of the field. Fundamental advances in emerging areas such as biomedicine, renewable energy, and critical infrastructures are expected to be enabled by control systems.

The Impact of Control Technology discusses and showcases the role of control systems both for products, solutions, and systems today and for addressing outstanding problems facing society and industry tomorrow. The report consists of four main parts:

Part 1: Overviews of control applications and opportunities in a number of domains

Part 2: Two-page flyers highlighting about 20 "success stories" in control

Part 3: Discussions of emerging research areas for the field

Part 4: An additional set of about 20 two-page flyers summarizing "grand challenges" for control technology

ABOUT THE EDITORS

Tariq Samad is a Corporate Fellow at Honeywell Automation and Control Solutions in Minneapolis, MN, U.S.A. He served as the President of IEEE Control Systems Society in 2009.

Anuradha Annaswamy is a Senior Research Scientist in the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, U.S.A. She is a member of the IEEE Control Systems Society's Board of Governors for 2010 – 2012.

This report is also available electronically at www.ieeecss.org/main/IoCT-report.
Published by IEEE Control Systems Society