Chapter 1

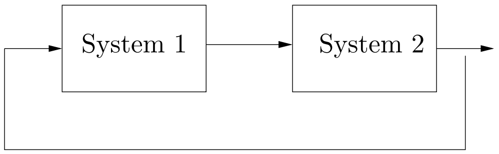
Introduction

> *Feedback is a central feature of life. The process of feedback governs how we grow, respond to stress and challenge, and regulate factors such as body temperature, blood pressure, and cholesterol level. The mechanisms operate at every level, from the interaction of proteins in cells to the interaction of organisms in complex ecologies.*
>
> Mahlon B. Hoagland and B. Dodson, *The Way Life Works*, 1995 [HD95].

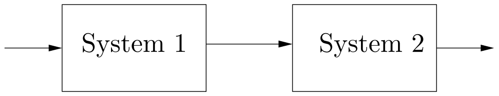
In this chapter we provide an introduction to the basic concept of *feedback* and the related engineering discipline of *control*. We focus on both historical and current examples, with the intention of providing the context for current tools in feedback and control. Much of the material in this chapter is adopted from [Mur03] and the authors gratefully acknowledge the contributions of Roger Brockett and Gunter Stein for portions of this chapter.

1.1 What is Feedback?

The term *feedback* is used to refer to a situation in which two (or more) dynamical systems are connected together such that each system influences the other and their dynamics are thus strongly coupled. By dynamical system, we refer to a system whose behavior changes over time, often in response to external stimulation or forcing. Simple causal reasoning about a feedback system is difficult because the first system influences the second and the second system influences the first, leading to a circular argument. This makes reasoning based on cause and effect tricky and it is necessary to analyze the system as a whole. A consequence of this is that the behavior of feedback systems is often counterintuitive and it is therefore necessary to resort to formal methods to understand them.



(a) Closed loop



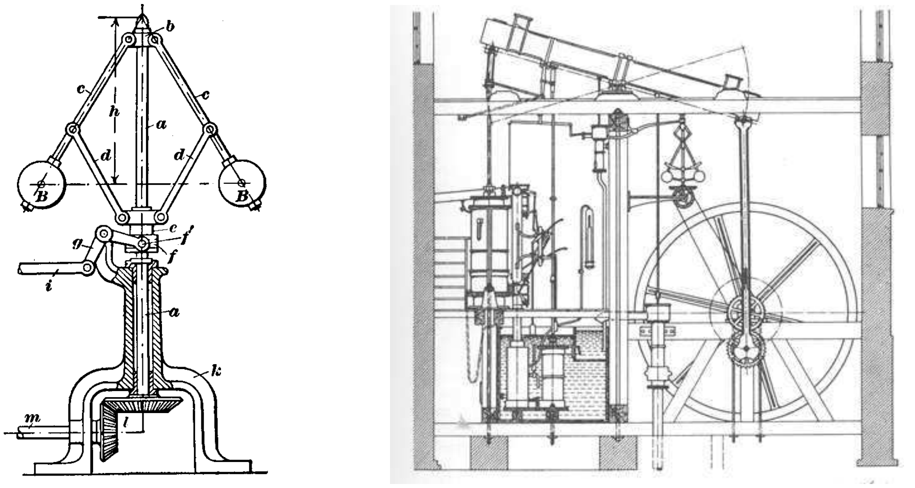
(b) Open loop

Figure 1.1: Open and closed loop systems.

Figure 1.1 illustrates in block diagram form the idea of feedback. We often use the terms *open loop* and *closed loop* when referring to such systems. A system is said to be a closed loop system if the systems are interconnected in a cycle, as shown in Figure 1.1a. If we break the interconnection, we refer to the configuration as an open loop system, as shown in Figure 1.1b.

As the quote at the beginning of this chapter illustrates, a major source of examples for feedback systems is from biology. Biological systems make use of feedback in an extraordinary number of ways, on scales ranging from molecules to microbes to organisms to ecosystems. One example is the regulation of glucose in the bloodstream, through the production of insulin and glucagon by the pancreas. The body attempts to maintain a constant concentration of glucose, which is used by the body's cells to produce energy. When glucose levels rise (after eating a meal, for example), the hormone insulin is released and causes the body to store excess glucose in the liver. When glucose levels are low, the pancreas secretes the hormone glucagon, which has the opposite effect. The interplay between insulin and glucagon secretions throughout the day help to keep the blood-glucose concentration constant, at about 90 mg per 100 ml of blood.

An early engineering example of a feedback system is the centrifugal governor, in which the shaft of a steam engine is connected to a flyball mechanism that is itself connected to the throttle of the steam engine, as illustrated in Figure 1.2. The system is designed so that as the speed of the engine increases (perhaps due to a lessening of the load on the engine), the flyballs spread apart and a linkage causes the throttle on the steam engine to be closed. This in turn slows down the engine, which causes the flyballs to come back together. When properly designed, the flyball governor



(a)

(b)

Figure 1.2: The centrifugal governor (a), developed in the 1780s, was an enabler of the successful Watt steam engine (b), which fueled the industrial revolution. Figures courtesy Richard Adamek (copyright 1999) and Cambridge University.

maintains a constant speed of the engine, roughly independent of the loading conditions.

Feedback has many interesting properties that can be exploited in designing systems. As in the case of glucose regulation or the flyball governor, feedback can make a system very resilient towards external influences. It can also be used to create linear behavior out of nonlinear components, a common approach in electronics. More generally, feedback allows a system to be very insensitive both to external disturbances and to variations in its individual elements.

Feedback has potential disadvantages as well. If applied incorrectly, it can create dynamic instabilities in a system, causing oscillations or even runaway behavior. Another drawback, especially in engineering systems, is that feedback can introduce unwanted sensor noise into the system, requiring careful filtering of signals. It is for these reasons that a substantial portion of the study of feedback systems is devoted to developing an understanding of dynamics and mastery of techniques in dynamical systems.

Feedback systems are ubiquitous in both natural and engineered systems. Control systems maintain the environment, lighting, and power in our buildings and factories, they regulate the operation of our cars, consumer electronics and manufacturing processes, they enable our transportation and communications systems, and they are critical elements in our military and space systems. For the most part, they are hidden from view, buried within the code of embedded microprocessors, executing their functions accurately and reliably. Feedback has also made it possible to increase dramatically the precision of instruments such as atomic force microscopes and telescopes.

In nature, homeostasis in biological systems maintains thermal, chemical, and biological conditions through feedback. At the other end of the size scale, global climate dynamics depend on the feedback interactions between the atmosphere, oceans, land, and the sun. Ecologies are filled with examples of feedback, resulting in complex interactions between animal and plant life. Even the dynamics of economies are based on the feedback between individuals and corporations through markets and the exchange of goods and services.

1.2 What is Control?

The term "control" has many meanings and often varies between communities. In this book, we define control to be the use of algorithms and feedback in engineered systems. Thus, control includes such examples as feedback loops in electronic amplifiers, set point controllers in chemical and materials processing, "fly-by-wire" systems on aircraft, and even router protocols that control traffic flow on the Internet. Emerging applications include high confidence software systems, autonomous vehicles and robots, real-time resource management systems, and biologically engineered systems. At its core, control is an *information* science, and includes the use of information in both analog and digital representations.

A modern controller senses the operation of a system, compares that against the desired behavior, computes corrective actions based on a model of the system's response to external inputs, and actuates the system to effect the desired change. This basic *feedback loop* of sensing, computation, and actuation is the central concept in control. The key issues in designing control logic are ensuring that the dynamics of the closed loop system are stable (bounded disturbances give bounded errors) and that they have the desired behavior (good disturbance rejection, fast responsiveness to changes in operating point, etc). These properties are established using a variety of modeling and analysis techniques that capture the essential physics of the system and permit the exploration of possible behaviors in the presence of uncertainty, noise and component failures.

A typical example of a modern control system is shown in Figure 1.3.

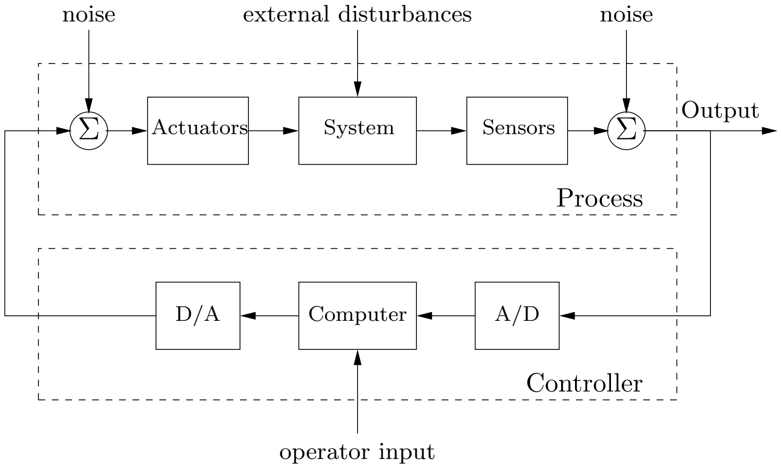
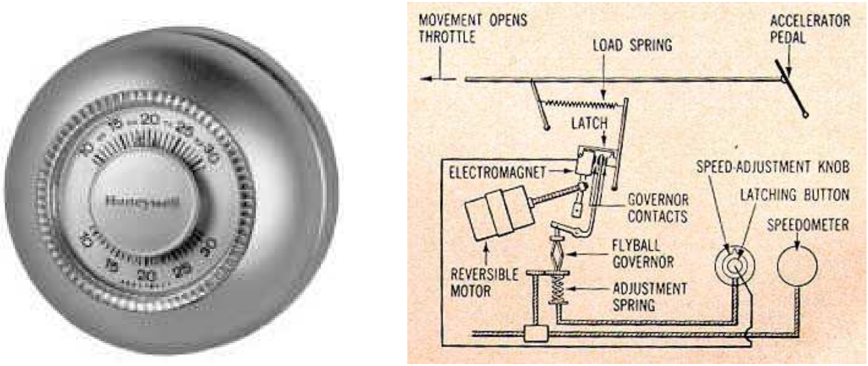


Figure 1.3: Components of a computer controlled system.

The basic elements of of sensing, computation and actuation are clearly seen. In modern control systems, computation is typically implemented on a digital computer, requiring the use of analog-to-digital (A/D) and digital-to-analog (D/A) converters. Uncertainty enters the system through noise in sensing and actuation subsystems, external disturbances that affect the underlying system physics, and uncertain dynamics in the physical system (parameter errors, unmodeled effects, etc). The algorithm that computes the control action as a function of the sensor values is often called a *control law.*

Control engineering relies on and shares tools from physics (dynamics and modeling), computer science (information and software) and operations research (optimization and game theory), but it is also different from these subjects in both insights and approach.

Perhaps the strongest area of overlap between control and other disciplines is in modeling of physical systems, which is common across all areas of engineering and science. One of the fundamental differences between control-oriented modeling and modeling in other disciplines is the way in which interactions between subsystems (components) are represented. Control relies on input/output modeling that allows many new insights into the behavior of systems, such as disturbance rejection and stable interconnection. Model reduction, where a simpler (lower-fidelity) description of the dynamics is derived from a high fidelity model, is also very naturally described in an input/output framework. Perhaps most importantly, modeling in a control context allows the design of *robust* interconnections between



(a) (b)

Figure 1.4: Early control devices: (a) Honeywell T86 thermostat, originally introduced in 1953, (b) Chrysler cruise control system, introduced in the 1958 Chrysler Imperial (note the centrifugal governor) [Row58].

subsystems, a feature that is crucial in the operation of all large engineered systems.

Control is also closely associated with computer science, since virtually all modern control algorithms for engineering systems are implemented in software. However, control algorithms and software are very different from traditional computer software. The physics (dynamics) of the system are paramount in analyzing and designing them and their real-time nature dominates issues of their implementation.

1.3 Feedback Examples

Feedback has many interesting and useful properties. It makes it possible to design precise systems from imprecise components and to make physical variables in a system change in a prescribed fashion. An unstable system can be stabilized using feedback and the effects of external disturbances can be reduced. Feedback also offers new degrees of freedom to a designer by exploiting sensing, actuation and computation. In this section we survey some of the important applications and trends for feedback in the world around us.

Early Technological Examples

The proliferation of control in engineered systems has occurred primarily in the latter half of the 20th century. There are some familiar exceptions, such as the centrifugal governor described earlier and the thermostat (Figure 1.4a), designed at the turn of the century to regulate temperature of buildings.

The thermostat, in particular, is often cited as a simple example of feedback control that everyone can understand. Namely, the device measures the temperature in a building, compares that temperature to a desired set point, and uses the "feedback error" between these two to operate the heating plant, e.g. to turn heating on when the temperature is too low and to turn if off when the temperature is too high. This explanation captures the essence of feedback, but it is a bit too simple even for a basic device such as the thermostat. Actually, because lags and delays exist in the heating plant and sensor, a good thermostat does a bit of anticipation, turning the heater off before the error actually changes sign. This avoids excessive temperature swings and cycling of the heating plant.

This modification illustrates that, even in simple cases, good control system design is not entirely trivial. It must take into account the dynamic behavior of the object being controlled in order to do a good job. The more complex the dynamic behavior, the more elaborate the modifications. In fact, the development of a thorough theoretical understanding of the relationship between dynamic behavior and good controllers constitutes the most significant intellectual accomplishment of the control community, and the codification of this understanding into powerful computer aided engineering design tools makes all modern control systems possible.

There are many other control system examples that have developed over the years with progressively increasing levels of sophistication and impact. An early system with broad public exposure was the "cruise control" option introduced on automobiles in 1958 (see Figure 1.4b). With cruise control, ordinary people experienced the dynamic behavior of closed loop feedback systems in action—the slowdown error as the system climbs a grade, the gradual reduction of that error due to integral action in the controller, the small (but unavoidable) overshoot at the top of the climb, etc. More importantly, by experiencing these systems operating reliably and robustly, the public learned to trust and accept feedback systems, permitting their increasing proliferation all around us. Later control systems on automobiles have had more concrete impact, such as emission controls and fuel metering systems that have achieved major reductions of pollutants and increases in

Figure 1.5: The F-18 aircraft, one of the first production military fighters to use "fly-by-wire" technology, and the X-45 (UCAV) unmanned aerial vehicle. Photographs courtesy of NASA Dryden Flight Research Center.

fuel economy.

In the industrial world, control systems have been a key enabling technology for everything from factory automation (starting with numerically controlled machine tools), to process control in oil refineries and chemical plants, to integrated circuit manufacturing, to power generation and distribution. Early use of regulators for manufacturing systems has evolved to the use of hundreds or even thousands of computer controlled subsystems in major industrial plants.

Aerospace and Transportation

Aerospace and transportation systems encompass a collection of critically important application areas where control is a central technology. These application areas represent a significant part of the modern world's overall technological capability. They are also a major part of its economic strength, and they contribute greatly to the well being of its people.

In aerospace, control has been a key technological capability tracing back to the very beginning of the 20th century. Indeed, the Wright brothers are correctly famous not simply for demonstrating powered flight but *controlled* powered flight. Their early Wright Flyer incorporated moving control surfaces (vertical fins and canards) and warpable wings that allowed the pilot to regulate the aircraft's flight. In fact, the aircraft itself was not stable, so continuous pilot corrections were mandatory. This early example of controlled flight is followed by a fascinating success story of continuous improvements in flight control technology, culminating in the very high performance, highly reliable automatic flight control systems we see on modern commercial and military aircraft today.

Similar success stories for control technology occurred in many other application areas. Early World War II bombsights and fire control servo systems have evolved into today's highly accurate radar-guided guns and precision-guided weapons. Early failure-prone space missions have evolved into routine launch operations, manned landings on the moon, permanently manned space stations, robotic vehicles roving Mars, orbiting vehicles at the outer planets, and a host of commercial and military satellites serving various surveillance, communication, navigation, and earth observation needs. Cars have advanced from manually tuned mechanical/pneumatic technology to computer-controlled operation of all major functions, including fuel injection, emission control, cruise control, braking, and cabin comfort.

Current research in aerospace and transportation systems is investigating the application of feedback to higher levels of decision making, including logical regulation of operating modes, vehicle configurations, payload configurations, and health status. These have historically been performed by human operators, but today that boundary is moving, and control systems are increasingly taking on these functions. Another dramatic trend on the horizon is the use of large collections of distributed entities with local computation, global communication connections, very little regularity imposed by the laws of physics, and no possibility of imposing centralized control actions. Examples of this trend include the national airspace management problem, automated highway and traffic management, and the command and control for future battlefields.

Information and Networks

The rapid growth of communication networks provides several major opportunities and challenges for control. Although there is overlap, we can divide these roughly into two main areas: control *of* networks and control *over* networks.

Control of networks is a large area, spanning many topics, including congestion control, routing, data caching, and power management. Several features of these control problems make them very challenging. The dominant feature is the extremely large scale of the system; the Internet is probably the largest feedback control system man has ever built. Another is the decentralized nature of the control problem: local decisions must be made quickly and based only on local information. Stability is complicated by the presence of varying time lags, as information about the network state

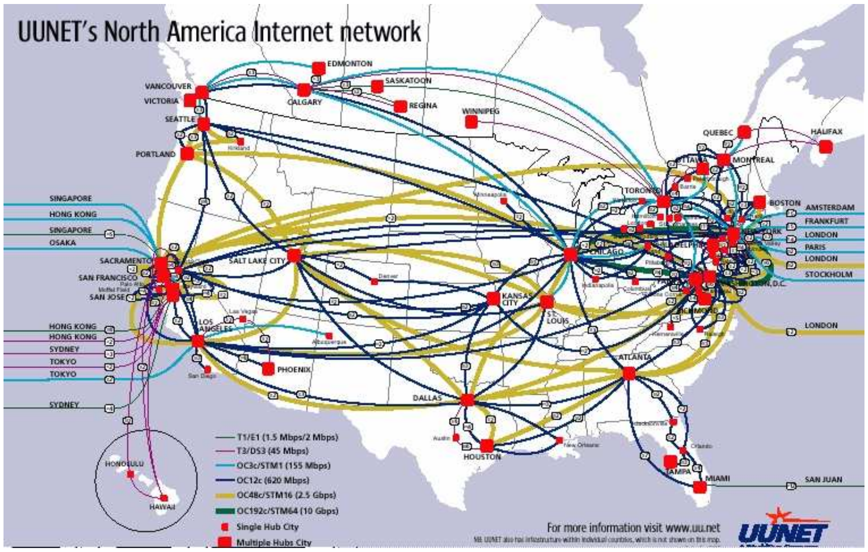


Figure 1.6: UUNET network backbone for North America. Figure courtesy of WorldCom.

can only be observed or relayed to controllers after a delay, and the effect of a local control action can be felt throughout the network only after substantial delay. Uncertainty and variation in the network, through network topology, transmission channel characteristics, traffic demand, available resources, and the like, may change constantly and unpredictably. Other complicating issues are the diverse traffic characteristics—in terms of arrival statistics at both the packet and flow time scales—and the different requirements for quality of service that the network must support.

Resources that must be managed in this environment include computing, storage and transmission capacities at end hosts and routers. Performance of such systems is judged in many ways: throughput, delay, loss rates, fairness, reliability, as well as the speed and quality with which the network adapts to changing traffic patterns, changing resource availability, and changing network congestion. The robustness and performance of the global Internet is a testament to the use of feedback to meet the needs of society in the face of these many uncertainties.

While the advances in information technology to date have led to a global Internet that allows users to exchange information, it is clear that the next phase will involve much more interaction with the physical environment and the increased use of control over networks. Networks of sensor and actuator nodes with computational capabilities, connected wirelessly or by wires,

Figure 1.7: “Spirit”, one of the two Mars Exploratory Rovers, and Sony AIBO Entertainment Robot. Photographs courtesy of Jet Propulsion Laboratory and Sony.

can form an orchestra that controls our physical environment. Examples include automobiles, smart homes, large manufacturing systems, intelligent highways and networked city services, and enterprise-wide supply and logistics chains.

Robotics and Intelligent Machines

Whereas early robots were primarily used for manufacturing, modern robots include wheeled and legged machines capable of competing in robotic competitions and exploring planets, unmanned aerial vehicles for surveillance and combat, and medical devices that provide new capabilities to doctors. Future applications will involve both increased autonomy and increased interaction with humans and with society. Control is a central element in all of these applications and will be even more important as the next generation of intelligent machines are developed.

The goal of cybernetic engineering, already articulated in the 1940s and even before, has been to implement systems capable of exhibiting highly flexible or “intelligent” responses to changing circumstances. In 1948, the MIT mathematician Norbert Wiener gave a widely read account of cybernetics [Wie48]. A more mathematical treatment of the elements of engineering cybernetics was presented by H.S. Tsien in 1954, driven by problems related to control of missiles [Tsi54]. Together, these works and others of that time form much of the intellectual basis for modern work in robotics and control.

Two accomplishments that demonstrate the successes of the field are the Mars Exploratory Rovers and entertainment robots such as the Sony AIBO, shown in Fig. 1.7. The two Mars Exploratory Rovers, launched by the Jet Propulsion Laboratory (JPL), maneuvered on the surface of Mars for over two years starting in January 2004 and sent back pictures and measurements of their environment. The Sony AIBO robot debuted in June of 1999 and was the first "entertainment" robot to be mass marketed by a major international corporation. It was particularly noteworthy because of its use of AI technologies that allowed it to act in response to external stimulation and its own judgment. This "higher level" of feedback is key element of robotics, where issues such as task-based control and learning are prevalent.

Despite the enormous progress in robotics over the last half century, the field is very much in its infancy. Today's robots still exhibit extremely simple behaviors compared with humans, and their ability to locomote, interpret complex sensory inputs, perform higher level reasoning, and cooperate together in teams is limited. Indeed, much of Wiener's vision for robotics and intelligent machines remains unrealized. While advances are needed in many fields to achieve this vision—including advances in sensing, actuation, and energy storage—the opportunity to combine the advances of the AI community in planning, adaptation, and learning with the techniques in the control community for modeling, analysis, and design of feedback systems presents a renewed path for progress.

Materials and Processing

The chemical industry is responsible for the remarkable progress in developing new materials that are key to our modern society. Process manufacturing operations require a continual infusion of advanced information and process control technologies in order for the chemical industry to maintain its global ability to deliver products that best serve the customer reliably and at the lowest cost. In addition, several new technology areas are being explored that will require new approaches to control to be successful. These range from nanotechnology in areas such as electronics, chemistry, and biomaterials to thin film processing and design of integrated microsystems to supply chain management and enterprise resource allocation. The payoffs for new advances in these areas are substantial, and the use of control is critical to future progress in sectors from semiconductors to pharmaceuticals to bulk materials.

There are several common features within materials and processing that pervade many of the applications. Modeling plays a crucial role, and there is a clear need for better solution methods for multidisciplinary systems combining chemistry, fluid mechanics, thermal sciences, and other disciplines at a variety of temporal and spatial scales. Better numerical methods for

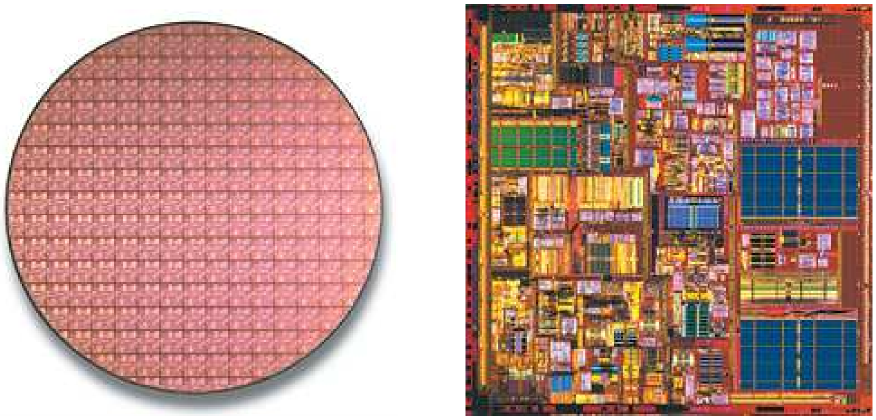


Figure 1.8: Intel Pentium IV wafer and die. Photographs courtesy of Intel.

traversing these scales and designing, controlling, and optimizing under uncertainty are also needed. And control techniques must make use of increased in situ measurements to control increasingly complex phenomena.

In addition to the continuing need to improve product quality, several other factors in the process control industry are drivers for the use of control. Environmental statutes continue to place stricter limitations on the production of pollutants, forcing the use of sophisticated pollution control devices. Environmental safety considerations have led to the design of smaller storage capacities to diminish the risk of major chemical leakage, requiring tighter control on upstream processes and, in some cases, supply chains. And large increases in energy costs have encouraged engineers to design plants that are highly integrated, coupling many processes that used to operate independently. All of these trends increase the complexity of these processes and the performance requirements for the control systems, making the control system design increasingly challenging.

As in many other application areas, new sensor technology is creating new opportunities for control. Online sensors—including laser backscattering, video microscopy, ultraviolet, infrared, and Raman spectroscopy—are becoming more robust and less expensive and are appearing in more manufacturing processes. Many of these sensors are already being used by current process control systems, but more sophisticated signal processing and control techniques are needed to more effectively use the real-time information provided by these sensors. Control engineers can also contribute to the design of even better sensors, which are still needed, for example, in the microelectronics industry. As elsewhere, the challenge is making use of the large amounts of data provided by these new sensors in an effective manner. In addition, a control-oriented approach to modeling the essential physics of the underlying processes is required to understand fundamental limits on observability of the internal state through sensor data.

Instrumentation

Feedback has had a major impact on instrumentation. Consider for example an accelerometer, where early instruments consisted of a mass suspended on a spring with a deflection sensor. The precision of such an instrument depends critically on accurate calibration of spring and the sensor. There is also a design compromise because a weak spring gives high sensitivity but also low bandwidth. An accelerometer based on feedback uses instead a voice coil to keep the mass at a given position and the acceleration is proportional to the current through the voice coil. In such an instrument the precision depends entirely on the calibration of the voice coil and does not depend on the sensor, which is only used as the feedback signal. The sensitivity bandwidth compromise is also avoided. This way of using feedback was applied to practically all engineering fields and it resulted in instruments with drastically improved performance. The development of inertial navigation where position is determined from gyroscopes and accelerometers which permits accurate guidance and control of vehicles is a spectacular example.

There are many other interesting and useful applications of feedback in scientific instruments. The development of the mass spectrometer is an early example. In a paper from 1935 by Nier it is observed that the deflection of the ions depend on both the magnetic and the electric fields. Instead of keeping both fields constant, Nier let the magnetic field fluctuate and the electric field was controlled to keep the ratio of the fields constant. The feedback was implemented using vacuum tube amplifiers. The scheme was crucial for the development of mass spectroscopy.

Another example is the work by the Dutch Engineer van der Meer. He invented a clever way to use feedback to maintain a high density and good quality of the beam of a particle accelerator. The idea is to sense particle displacement at one point in the accelerator and apply a correcting signal at another point. The scheme, called stochastic cooling, was awarded the Nobel Prize in Physics in 1984. The method was essential for the successful experiments in CERN when the existence of the particles W and Z was first demonstrated.

The 1986 Nobel Prize in Physics—awarded to Binnig and Rohrer for their design of the scanning tunneling microscope—is another example of clever use of feedback. The key idea is to move a narrow tip on a cantilever beam across the surface and to register the forces on the tip. The deflection of the tip was measured using tunneling which gave an extreme accuracy so that individual atoms could be registered.

A severe problem in astronomy is that turbulence in the atmosphere blurs images in telescopes because of variations in diffraction of light in the atmosphere. The blur is of the order of an arc-second in a good telescope. One way to eliminate the blur is to move the telescope outside the Earths atmosphere as is done with the Hubble telescope. Another way is to use feedback to eliminate the effects of the variations in a telescope on the Earth which is the idea of "adaptive optics." The reference signal is a bright star or an artificial laser beam projected into the atmosphere. The actuator is a deformable mirror which can have hundreds or thousands of elements. The error signal is formed by analyzing the shape of the distorted wave form from the reference. This signal is sent to the controller which adjusts the deformable mirror. The light from the observed star is compensated because it is also reflected in the deformable mirror before it is sent to the detector. The wave lengths used for observation and control are often different. Since diffraction in the atmosphere changes quite rapidly the response time of the control system must be of the order of milliseconds.

Feedback in Nature

Many cutting edge problems in the natural sciences involve understanding aggregate behavior in complex large-scale systems. This behavior "emerges" from the interaction of a multitude of simpler systems, with intricate patterns of information flow. Representative examples can be found in fields ranging from embryology to seismology. Researchers who specialize in the study of specific complex systems often develop an intuitive emphasis on analyzing the role of feedback (or interconnection) in facilitating and stabilizing aggregate behavior, and it is often noted that one can only have hope of deep understanding if it is somehow possible for theories of collective phenomenology to be robust to inevitable uncertainties in the modeling of fine-scale dynamics and interconnection.

While sophisticated theories have been developed by domain experts for the analysis of various complex systems, the development of rigorous methodology that can discover and exploit common features and essential mathematical structure is just beginning to emerge. Advances in science and technology are creating new understanding of the underlying dynamics and the importance of feedback in a wide variety of natural and technological

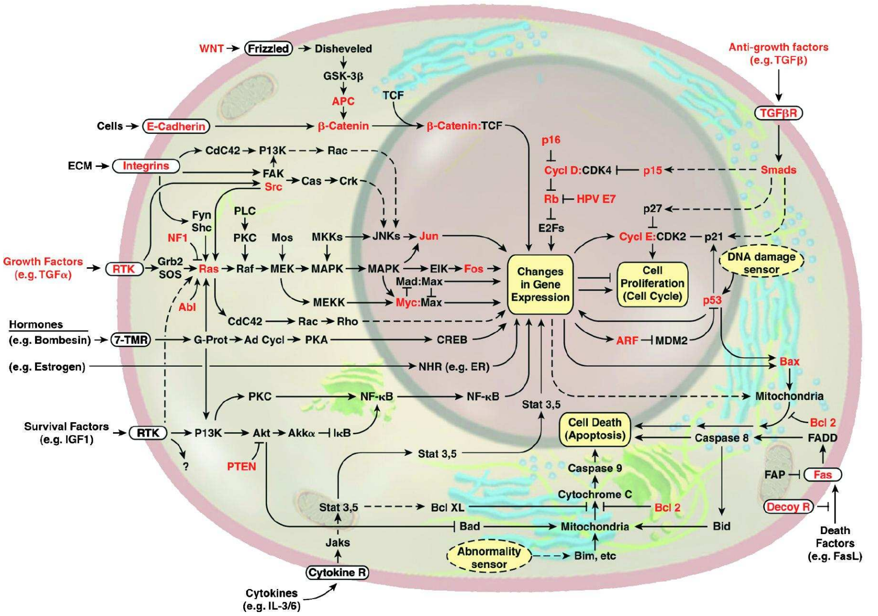


Figure 1.9: The wiring diagram of the growth signaling circuitry of the mammalian cell [HW00].

systems We briefly highlight four application areas here.

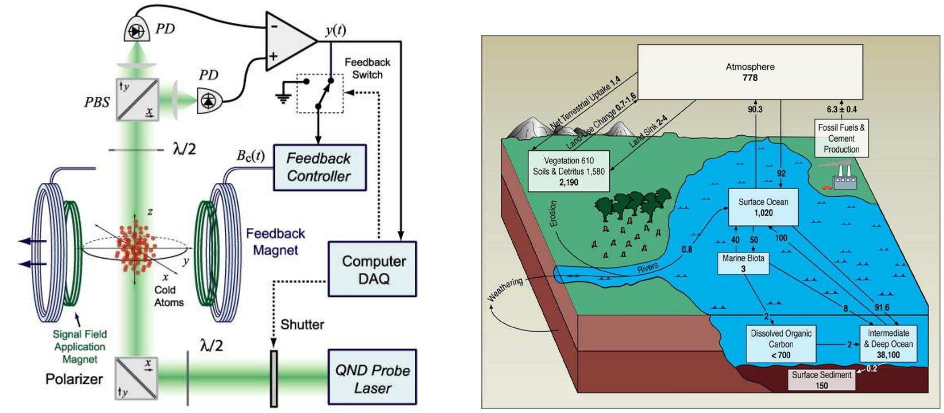
Biological Systems. At a variety of levels of organization—from molecular to cellular to organismal to populational—biology is becoming more accessible to approaches that are commonly used in engineering: mathematical modeling, systems theory, computation, and abstract approaches to synthesis. Conversely, the accelerating pace of discovery in biological science is suggesting new design principles that may have important practical applications in man-made systems. This synergy at the interface of biology and engineering offers unprecedented opportunities to meet challenges in both areas. The principles of feedback and control are central to many of the key questions in biological engineering and will play an enabling role in the future of this field.

A major theme currently underway in the biology community is the science of reverse (and eventually forward) engineering of biological control networks (such as the one shown in Figure 1.9). There are a wide variety of biological phenomena that provide a rich source of examples for control, including gene regulation and signal transduction; hormonal, immunological, and cardiovascular feedback mechanisms; muscular control and locomotion; active sensing, vision, and proprioception; attention and consciousness; and population dynamics and epidemics. Each of these (and many more) provide opportunities to figure out what works, how it works, and what we can do to affect it.

Ecosystems. In contrast to individual cells and organisms, emergent properties of aggregations and ecosystems inherently reflect selection mechanisms which act on multiple levels, and primarily on scales well below that of the system as a whole. Because ecosystems are complex, multiscale dynamical systems, they provide a broad range of new challenges for modeling and analysis of feedback systems. Recent experience in applying tools from control and dynamical systems to bacterial networks suggests that much of the complexity of these networks is due to the presence of multiple layers of feedback loops that provide robust functionality to the individual cell. Yet in other instances, events at the cell level benefit the colony at the expense of the individual. Systems level analysis can be applied to ecosystems with the goal of understanding the robustness of such systems and the extent to which decisions and events affecting individual species contribute to the robustness and/or fragility of the ecosystem as a whole.

Quantum Systems. While organisms and ecosystems have little to do with quantum mechanics in any traditional scientific sense, complexity and robustness issues very similar to those described above can be identified in the modern study of quantum systems. In large part, this sympathy arises from a trend towards wanting to control quantum dynamics and to harness it for the creation of new technological devices. At the same time, physicists are progressing from the study of elementary quantum systems to the study of large aggregates of quantum components, and it has been recognized that dynamical complexity in quantum systems increases exponentially faster with system size than it does in systems describable by classical (macroscopic) physics. Factors such as these are prompting the physics community to search broadly for new tools for understanding robust interconnection and emergent phenomena.

Modern scientific research is rapidly evolving a field of *quantum engineering.* Driven by technology goals in areas such as quantum information processing, nano-electromechanical sensing, chemical synthesis, trace gas detection, and ultrahigh-bandwidth optical communication, researchers are beginning to formulate strategies for achieving robust performance with physical devices or systems in which quantum mechanical effects are prominent. Mathematical tools from control and dynamical systems for analysis and synthesis could have a profound impact on activities of this kind. A schematic diagram of a modern quantum control experiment is show in Figure 1.10a.



(a) (b)

Figure 1.10: Examples of feedback systems in nature: (a) quantum control system and (b) global carbon cycle.

Environmental Science. It is now indisputable that human activities have altered the environment on a global scale. Problems of enormous complexity challenge researchers in this area and first among these is to understand the feedback systems that operate on the global scale. One of the challenges in developing such an understanding is the multiscale nature of the problem, with detailed understanding of the dynamics of microscale phenomena such as microbiological organisms being a necessary component of understanding global phenomena, such as the carbon cycle illustrated Figure 1.10b.

Other Areas

The previous sections have described some of the major application areas for control. However, there are many more areas where ideas from control are being applied or could be applied. Some of these include: economics and finance, including problems such as pricing and hedging options; energy systems, including load distribution and power management for the electrical grid; and manufacturing systems, including supply chains, resource management and scheduling, and factory automation.

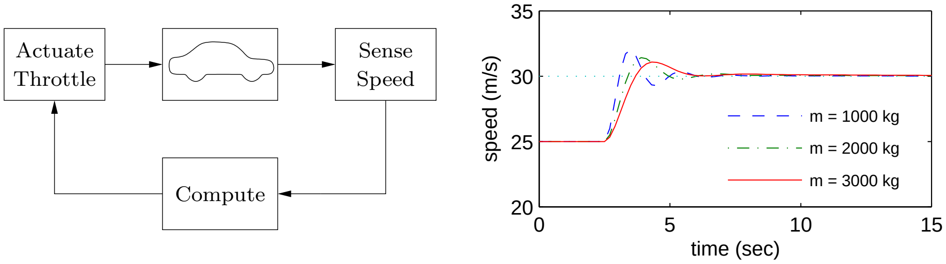


Figure 1.11: A simple feedback system for controlling the speed of a vehicle.

1.4 Feedback Properties

Feedback is a powerful idea which, as we have seen, is used extensively in natural and technological systems. The principle of feedback is very simple: base correcting actions on the difference between desired and actual performance. In engineering, feedback has been rediscovered and patented many times in many different contexts. The use of feedback has often resulted in vast improvements in system capability and these improvements have sometimes been revolutionary, as discussed above. The reason for this is that feedback has some truly remarkable properties. In this section we will discuss some of the properties of feedback that can be understood intuitively. This intuition will be formalized in the subsequent chapters.

Robustness to Uncertainty

One of the key uses of feedback is to provide robustness to uncertainty. By measuring the difference between the sensed value of a regulated signal and its desired value, we can supply a corrective action. If the system undergoes some change that affects the regulated signal, then we sense this change and try to force the system back to the desired operating point. This is precisely the effect that Watt exploited in his use of the centrifugal governor on steam engines.

As an example of this principle, consider the simple feedback system shown in Figure 1.11a. In this system, the speed of a vehicle is controlled by adjusting the amount of gas flowing to the engine. A simple "proportional plus integral" feedback is used to to make the amount of gas depend on both the error between the current and desired speed, and the integral of that error. The plot on the right shows the results of this feedback for a step change in the desired speed and a variety of different masses for the car (which might result from having a different number of passengers or towing a trailer). Notice that independent of the mass (which varies by a factor of 3), the steady state speed of the vehicle always approaches the desired speed and achieves that speed within approximately 5 seconds. Thus the performance of the system is robust with respect to this uncertainty.

Another early example of the use of feedback to provide robustness was the negative feedback amplifier. When telephone communications were developed, amplifiers were used to compensate for signal attenuation in long lines. The vacuum tube was a component that could be used to build amplifiers. Distortion caused by the nonlinear characteristics of the tube amplifier together with amplifier drift were obstacles that prevented development of line amplifiers for a long time. A major breakthrough was invention of the feedback amplifier in 1927 by Harold S. Black, an electrical engineer at the Bell Telephone Laboratories. Black used negative feedback which reduces the gain but makes the amplifier very insensitive to variations in tube characteristics. This invention made it possible to build stable amplifiers with linear characteristics despite nonlinearities of the vacuum tube amplifier.

Design of Dynamics

Another use of feedback is to change the dynamics of a system. Through feedback, we can alter the behavior of a system to meet the needs of an application: systems that are unstable can be stabilized, systems that are sluggish can be made responsive, and systems that have drifting operating points can be held constant. Control theory provides a rich collection of techniques to analyze the stability and dynamic response of complex systems and to place bounds on the behavior of such systems by analyzing the gains of linear and nonlinear operators that describe their components.

An example of the use of control in the design of dynamics comes from the area of flight control. The following quote, from a lecture by Wilbur Wright to the Western Society of Engineers in 1901 [McF53], illustrates the role of control in the development of the airplane:

> “Men already know how to construct wings or airplanes, which when driven through the air at sufficient speed, will not only sustain the weight of the wings themselves, but also that of the engine, and of the engineer as well. Men also know how to build engines and screws of sufficient lightness and power to drive these planes at sustaining speed ... Inability to balance and steer still confronts students of the flying problem. ... When this one

Figure 1.12: The Curtiss-Sperry E in 1912 (left) and a closeup of the Sperry Autopilot (right).

> feature has been worked out, the age of flying will have arrived, for all other difficulties are of minor importance."

The Wright brothers thus realized that control was a key issue to enable flight. They resolved the compromise between stability and maneuverability by building an airplane, Kitty Hawk, that was unstable but maneuverable. Kitty Hawk had a rudder in the front of the airplane, which made the plane very maneuverable. A disadvantage was the necessity for the pilot to keep adjusting the rudder to fly the plane: if the pilot let go of the stick the plane would crash. Other early aviators tried to build stable airplanes. These would have been easier to fly, but because of their poor maneuverability they could not be brought up into the air. By using their insight and skillful experiments the Wright brothers made the first successful flight with Kitty Hawk in 1905.

Since it was quite tiresome to fly an unstable aircraft, there was strong motivation to find a mechanism that would stabilize an aircraft. Such a device, invented by Sperry, was based on the concept of feedback. Sperry used a gyro-stabilized pendulum to provide an indication of the vertical. He then arranged a feedback mechanism that would pull the stick to make the plane go up if it was pointing down and vice versa. The Sperry autopilot is the first use of feedback in aeronautical engineering and Sperry won a prize in a competition for the safest airplane in Paris in 1912. Figure 1.12 shows the Curtiss-Sperry seaplane and the Sperry autopilot. The autopilot is a good example of how feedback can be used to stabilize an unstable system and hence "design the dynamics" of the aircraft.

One of the other advantages of designing the dynamics of a device is that it allows for increased modularity in the overall system design. By us-

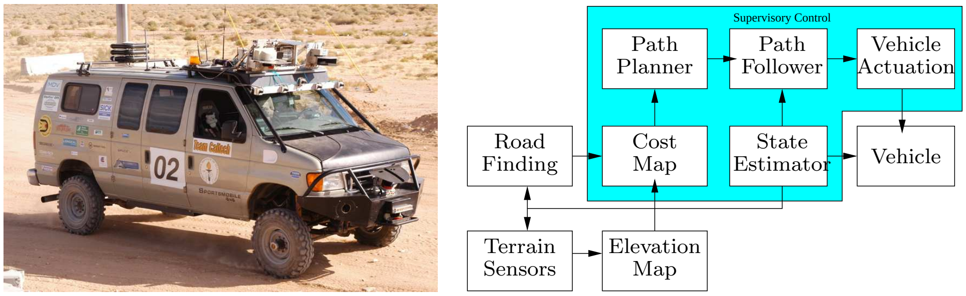


Figure 1.13: DARPA Grand Challenge: (a) "Alice," Team Caltech's entry in the 2005 competition. (b) Networked control architecture used by Alice.

ing feedback to create a system whose response matches a desired profile, we can hide the complexity and variability that may be present inside a subsystem. This allows creation us to create more complex systems by not having to simultaneously tune the response of a large number of interacting components. This was one of the advantages of Black's use of negative feedback in vacuum tube amplifiers: the resulting device had a well-defined linear input/output response that did not depend on the individual characteristics of the vacuum tubes begin used.

Higher Levels of Automation

A major trend in the use of feedback is its use in higher levels of automation and decision making. A good example of this is the DARPA Grand Challenge, a competition sponsored by the US government to build a vehicle that could autonomously drive itself across the desert. Caltech competed in the 2005 Grand Challenge using a modified Ford E-350 offroad van, nicknamed "Alice." It was fully automated, including electronically controlled steering, throttle, brakes, transmission, and ignition. Its sensing systems included 4 black and white cameras sampling at 30 Hz (arranged in two stereo pairs), 1 color camera for finding roads, 5 LADAR (laser radar) units scanning at 10 Hz, and a GPS/IMU package capable of providing full position and orientation at 2.5 millisecond temporal resolution. Computational resources included 7 high speed servers connected together through a 1 Gb/s Ethernet switch. A picture of the vehicle is shown in Figure 1.13a.

Custom software was used to control the vehicle. The control system architecture is shown in Figure 1.13b. This information-rich feedback system fused data from multiple cameras and LADAR units to determine a digital elevation map for the terrain around it and then used this map to compute a speed map that estimated the speed at which the vehicle could drive in the environment. The map was modified to include additional information where roads were identified (through vision-based algorithms) and where no data was present (due to hills or temporary sensor outages). This speed map was then used by an optimization-based planner to determine the path that would allow the vehicle to make the most progress in a fixed period of time. The commands from the planner were sent to a trajectory tracking algorithm that compared the desired vehicle position to its estimated position (from GPS/IMU data) and issued appropriate commands to the steering, throttle and brake actuators. Finally, a supervisor control module performed higher level tasks, including implementing strategies for making continued forward progress if one of the hardware or software components failed temporarily (either due to external or internal conditions).

The software and hardware infrastructure that was developed enabled the vehicle to traverse long distances at substantial speeds. In testing, Alice drove itself over 500 kilometers in the Mojave Desert of California, with the ability to follow dirt roads and trails (if present) and avoid obstacles along the path. Speeds of over 50 km/hr were obtained in fully autonomous mode. Substantial tuning of the algorithms was done during desert testing, in part due to the lack of systems-level design tools for systems of this level of complexity. Other competitors in the race (including Stanford, which one the competition) used algorithms for adaptive control and learning, increasing the capabilities of their systems in unknown environments. Together, the competitors in the grand challenge demonstrated some of the capabilities for the next generation of control systems and highlighted many research directions in control at higher levels of decision making.

Drawbacks of Feedback

While feedback has many advantages, it also has some drawbacks. Chief among these is the possibility for instability if the system is not designed properly. We are all familiar with the effects of "positive feedback" when the amplification on a microphone is turned up too high in a room. This is an example of a feedback instability, something that we obviously want to avoid. This is tricky because of the uncertainty that feedback was introduced to compensate for: not only must we design the system to be stable with the nominal system we are designing for, but it must remain stable under all possible perturbations of the dynamics.

In addition to the potential for instability, feedback inherently couples different parts of a system. One common problem that feedback inherently injects measurement noise into the system. In engineering systems, measurements must be carefully filtered so that the actuation and process dynamics do not respond to it, while at the same time insuring that the measurement signal from the sensor is properly coupled into the closed loop dynamics (so that the proper levels of performance are achieved).

Another potential drawback of control is the complexity of embedding a control system into a product. While the cost of sensing, computation, and (to a lesser extent) actuation has decreased dramatically in the past few decades, the fact remains that control systems are often very complicated and hence one must carefully balance the costs and benefits. An early engineering example of this is the use of microprocessor-based feedback systems in automobiles. The use of microprocessors in automotive applications began in the early 1970s and was driven by increasingly strict emissions standards, which could only be met through electronic controls. Early systems were expensive and failed more often than desired, leading to frequent customer dissatisfaction. It was only through aggressive improvements in technology that the performance, reliability and cost of these systems allowed them to be used in a transparent fashion. Even today, the complexity of these systems is such that it is difficult for an individual car owner to fix problems. There have also been spectacular failures due to unexpected interactions.

Feedforward

When using feedback that there must be an error before corrective actions are taken. Feedback is thus reactive. In some circumstances it is possible to measure a disturbance before it enters the system and this information can be used to take corrective action before the disturbance has influenced the system. The effect of the disturbance is thus reduced by measuring it and generating a control signal that counteracts it. This way of controlling a system is called *feedforward*. Feedforward is particularly useful to shape the response to command signals because command signals are always available. Since feedforward attempts to match two signals, it requires good process models otherwise the corrections may have the wrong size or it may be badly timed.

The ideas of feedback and feedforward are very general and appear in many different fields. In economics, feedback and feedforward are analogous to a market-based economy versus a planned economy. In business a pure feedforward strategy corresponds to running a company based on extensive strategic planning while a feedback strategy corresponds to a pure reactive approach. The experience in control indicates that it is often advantageous to combine feedback and feedforward. Feedforward is particularly useful when disturbances can be measured or predicted. A typical example is in chemical process control where disturbances in one process may be due to processes upstream. The correct balance of the approaches requires insight and understanding of their properties.

Positive Feedback

In most of this text, we will consider the role of negative feedback, in which we attempt to regulate the system by reacting to disturbances in a way that decreases the effect of those disturbances. In some systems, particularly biological systems, *positive feedback* can play an important role. In a system with positive feedback, the increase in some variable or signal leads to a situation in which that quantify is further through its dynamics. This has a destabilizing effect and is usually accompanied by a saturation that limits the growth of the quantity. Although often considered undesirable, this behavior is used in biological (and engineering) systems to obtain a very fast response to a condition or signal.

1.5 Simple Forms of Feedback

The idea of feedback to make corrective actions based on the difference between the desired and the actual value can be implemented in many different ways. The benefits of feedback can be obtained by very simple feedback laws such as on-off control, proportional control and PID control. In this section we provide a brief preview of some of the topics that will be studied more formally in the remainder of the text.

A simple feedback mechanism can be described as follows:

$$u = \begin{cases} u_{\max} & \text{if } e > 0 \\ u_{\min} & \text{if } e < 0 \end{cases} \tag{1.1}$$

where $e = r - y$ is the difference between the reference signal and the output of the system. Figure 1.14a shows the relation between error and control. This control law implies that maximum corrective action is always used.

The feedback in equation (1.1) is called *on-off control.* One of its chief advantages is that it is simple and there are no parameters to choose. On-off control often succeeds in keeping the process variable close to the reference, such as the use of a simple thermostat to maintain the temperature of a

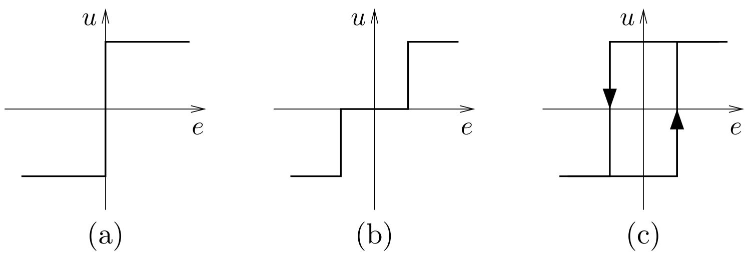


Figure 1.14: Controller characteristics for ideal on-off control (a), and modifications with dead zone (b) and hysteresis (c).

room. It typically results in a system where the controlled variables oscillate, which is often acceptable if the oscillation is sufficiently small.

Notice that in equation (1.1) the control variable is not defined when the error is zero. It is common to have some modifications either by introducing hysteresis or a dead zone (see Figure 1.14b and 1.14c).

The reason why on-off control often gives rise to oscillations is that the system over reacts since a small change in the error will make the actuated variable change over the full range. This effect is avoided in *proportional control*, where the characteristic of the controller is proportional to the control error for small errors. This can be achieved by making the control signal proportional to the error, which gives the control law

$$u = \begin{cases} u_{\max} & \text{if } e > e_{max} \\ ke & \text{if } e_{min} \le e \le e_{max} \\ u_{\min} & \text{if } e < e_{min}, \end{cases} \tag{1.2}$$

where where k is the controller gain, $e_{min} = u_{min}/k$, and $e_{max} = u_{max}/k$. The interval (e_{min}, e_{max}) is called the *proportional band* because the behavior of the controller is linear when the error is in this interval:

$$u = k(r - y) = ke \qquad \text{if } e_{min} \le e \le e_{max}. \tag{1.3}$$

While a vast improvement over on-off control, proportional control has the drawback that the process variable often deviates from its reference value. In particular, if some level of control signal is required for the system to maintain a desired value, then we must have $e \neq 0$ in order to generate the requisite input.

This can be avoided by making the control action proportional to the integral of the error:

$$u(t) = k_i \int_0^t e(\tau) d\tau. \tag{1.4}$$

This control form is called *integral control* and k_i is the integral gain. It can be shown through simple arguments that a controller with integral action will have zero "steady state" error (Exercise 5). The catch is that there may not always be a steady state because the system may be oscillating. This property has been rediscovered many times and is one of the properties that have strongly contributed to the wide applicability of integral controllers.

An additional refinement is to provide the controller with an anticipative ability by using a prediction of the error. A simple prediction is given by the linear extrapolation

$$e(t + T_d) \approx e(t) + T_d \frac{de(t)}{dt},$$

which predicts the error T_d time units ahead. Combining proportional, integral and derivative control we obtain a controller that can be expressed mathematically as follows:

$$\begin{aligned} u(t) &= ke(t) + k_i \int_0^t e(\tau)\, d\tau + k_d \frac{de(t)}{dt} \\ &= k\Big(e(t) + \frac{1}{T_i} \int_0^t e(\tau)\, d\tau + T_d \frac{de(t)}{dt}\Big) \end{aligned} \tag{1.5}$$

The control action is thus a sum of three terms: the past as represented by the integral of the error, the present as represented by the proportional term, and the future as represented by a linear extrapolation of the error (the derivative term). This form of feedback is called a *PID controller* and its action is illustrated in Figure 1.15.

The PID controller is very useful and is capable of solving a wide range of control problems. The PI controller, in which the derivative term is omitted, is one of the most common forms of the controller. It is quoted that about 90% of all control problems can be solved by PID control, although many of these controllers are actually PI controllers because derivative action is often not included. There are more advanced controllers which differ from the PID controller by using more sophisticated methods for prediction.

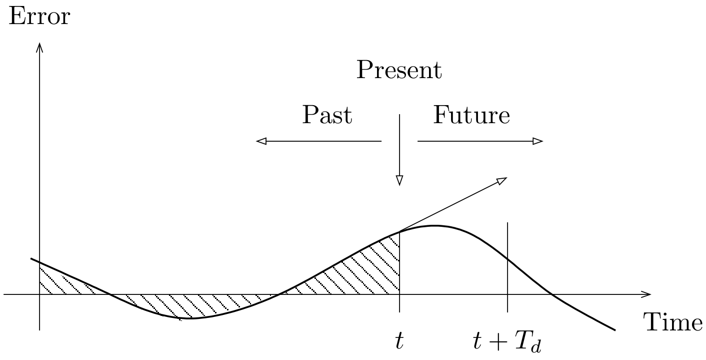


Figure 1.15: A PID controller takes control action based on past, present and future control errors.

1.6 Control Tools

The development of a control system consists of the tasks modeling, analysis, simulation, architectural design, design of control algorithms, implementation, commissioning and operation. Because of the wide use of feedback in a variety of applications, there has been substantial mathematical development in the field of control theory. In many cases the results have also been packaged in software tools that simplifies the design process. We briefly describe some of the tools and concepts here.

Modeling

Models play an essential role in analysis and design of feedback systems. Several sophisticated tools have been developed to build models that are suited for control.

Models can often be obtained from first principles and there are several modeling tools in special domains such as electric circuits and multibody systems. Since control applications cover such a wide range of domains it is also desirable to have modeling tools that cut across traditional discipline boundaries. Such modeling tools are now emerging, with the models obtained by cutting a system into subsystems and writing equations for balance of mass, energy and momentum, and constitutive equations that describe material properties for each subsystem. Object oriented programming can be used very effectively to organize the work and extensive symbolic computation can be used to simplify the equations. Models and components can then be organized in libraries for efficient reuse. Modelica [Til01] is an example of a modeling tool of this type.

Modeling from input/output data or system identification is another approach to modeling that has been developed in control. Direct measurement of the response to step input is commonly used in the process industry to tune proportional-integral (PI) controllers. More accurate models can be obtained by measuring the response to sinusoidal signals, which is particularly useful for systems with fast response time. Control theory has also developed new techniques for modeling dynamics and disturbances, including input/output representations of systems (how disturbances propagate through the system) and data-driven system identification techniques. The use of "forced response" experiments to build models of systems is well developed in the control field and these tools find application in many disciplines, independent of the use of feedback.

Finally, one of the hallmarks of modern control engineering is the development of model reduction techniques that allow a hierarchy of models to be constructed at different levels of fidelity. This was originally motivated by the need to produce low complexity controllers that could be implemented with limited computation. The theory is well developed for linear systems and is now used to produce models of varying levels of complexity with bounds on the input/output errors corresponding to different approximations. Model reduction for general classes of nonlinear systems is an important unsolved problem.

The impact of modeling on engineering and control over the past 50 years has been profound. Today, entire products are designed using only models, with the first prototype being a fully working system. Doing this requires an enormous amount of infrastructure in simulation tools, detailed physical models, and experience in using models. Moving forward, this infrastructure becomes even more important as suppliers of components compete and sell their products based on detailed models of their systems which implement the specifications of their products sometimes before the system has even been built.

Analysis

Control theory has developed an extensive collection of theory and software tools for analysis of feedback systems. These tools include stability analysis for both linear and nonlinear systems, performance measures for input/output systems, and evaluation of robustness. For robustness analysis, the tools are particularly sophisticated in the case of linear dynamical systems, where it is possible to analyze the stability and performance of the system in the presence of external disturbances, parametric uncertainty (e.g., unknown values of parameters such as mass or reaction rates), and unmodeled dynamics. In the case of unmodeled dynamics, one can even reason about the performance of a system without knowing the precise behavior of every subsystem, a very powerful concept.

In almost all cases, the theory used for analysis of feedback systems is implemented in software tools that can be used in computer aided design environments. Two very common environments for this analysis are MATLAB and Mathematica. In both cases, toolboxes are available that implement the common calculations described in this text (and many more) and these have become indispensable aides for modern design. More sophisticated tools, capable of constructing and analyzing very complex hierarchical models, are also available for more discipline-specific environments.

An important point to remember about systems analysis is that it relies on models to describe the behavior of the system. In the simplest case, these models are simulated to provide information about how the system will respond to a given set of initial conditions, disturbances, and environment. But modern computational tools can do much more than just simulate the system, and can provide very sophisticated analyses that answer questions about the parametric behavior of systems, tradeoffs between different design factors, and even provide certificates (proofs) of performance. These tools are very well developed for linear systems, but recent advances in nonlinear analysis are quickly extending these results to larger and larger classes of systems.

Synthesis

In addition to tools for analysis of feedback systems, theory and software has also been developed for synthesizing feedback systems. As an example, consider the control system depicted in Figure 1.3 on page 5. Given models of the process to be controlled, it is today possible to automatically synthesize a control algorithm that satisfies a given performance specification. A typical approach to doing this would involve first obtaining a nonlinear model for the process that captured the key behavior that one was interested in. This model would then be "linearized" around a desired operating point (this is described in Chapter 5) and a performance condition specified (usually as a function that one wishes to minimize). Given the linear model and the control specification, a feedback law can be computed that is the optimal law for the given specification.

Modern implementation environments continue by allowing this control algorithm to be "autocoded", so that programming code implementing the control logic is automatically generated. This has the advantage that the code generator can be carefully verified so that the resulting algorithm is correct (as opposed to hand coding the algorithm, which can lead to errors in translation). This autocoded control logic is then downloaded to a dedicated computing platform with the proper interfaces to the hardware and implemented. In addition to simple feedback algorithms, most modern control environments allow complex decision-making logic to be constructed via block diagrams and this is also automatically compiled and downloaded to a computer that is connected to the hardware.

This mechanism for generating control algorithms directly from specifications has vastly improved the productivity of control engineers and is now standard practice in many application areas. It also provides a clear framework in which new advances, such as real-time, optimization-based control, can be transitioned to applications quickly and efficiently through the generation of standard toolboxes.

1.7 Further Reading

The material in this section draws heavily from the report of the Panel on Future Directions on Control, Dynamics, and Systems [Mur03]. Several recent papers and reports highlighted successes of control [NS99] and new vistas in control [Bro00, Kum01]. A fascinating examination of some of the early history of control in the United States has been written by Mindell [Min02]. Additional historical overviews of the field have been prepared by Bennett [Ben86a, Ben86b] and Mayr[May70], which go back as far as the 1800s. A popular book that describes many control concepts across a wide range of disciplines is “Out of Control” by Kelly [Kel94].

There are many textbooks available that describe control systems in the context of specific disciplines. For engineers, the textbooks by Franklin, Powell and Emami-Naeni [FPEN05], Dorf and Bishop [DB04], Kuo and Golnaraghi [KG02], and Seborg, Edgar and Mellichamp [SEM03] are widely used. A number of books look at the role of dynamics and feedback in biological systems, including Milhorn [Mil66] (now out of print), Murray [Mur04] and Ellner and Guckenheimer [EG05]. There is not yet a textbook targeted specifically at the physics community, although a recent tutorial article by Bechhoefer [Bec05] covers many specific topics of interest to that community.

1.8 Exercises

1. Identify 5 feedback systems that you encounter in your everyday environment. For each system, identify the sensing mechanism, actuation mechanism, and control law. Describe the uncertainty that the feedback system provides robustness with respect to and/or the dynamics that are changed through the use of feedback.

2. Perform the following experiment and explain your results: Holding your head still, move your right or left hand back and forth in front of your face, following it with your eyes. Record how quickly you can move your hand before you begin to lose track of your hand. Now hold your hand still and move your head back and forth, once again recording how quickly you can move before loosing track.

3. Balance yourself on one foot with your eyes closed for 15 seconds. Using Figure 1.3 as a guide, describe the control system responsible for keeping you from falling down. Note that the "controller" will differ from the diagram (unless you are an android reading this in the far future).

4. Download the MATLAB code used to produce Figure 1.11 from the companion web site. Using trial and error, change the parameters of the control law so that the overshoot in the speed is not more than 1 m/s for a vehicle with mass $m = 1000 kg$.

5. We say that an system with a constant input reaches "steady state" if the output of the system approaches a constant value as time increases. Show that a controller with integral action, such as those given in equations (1.4) and (1.5), gives zero error if the closed loop system reaches steady state.