

Feedback Systems

An Introduction for Scientists and Engineers

SECOND EDITION

Karl Johan Åström
Richard M. Murray

Version v3.1.5 (2020-07-24)

This is the electronic edition of *Feedback Systems* and is available from <http://fbsbook.org>. Hardcover editions may be purchased from Princeton University Press, <http://press.princeton.edu/titles/8701.html>.

This manuscript is for personal use only and may not be reproduced, in whole or in part, without written consent from the publisher (see <http://press.princeton.edu/permissions.html>).

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.

M. 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.

1.1 What Is Feedback?

A *dynamical system* is a system whose behavior changes over time, often in response to external stimulation or forcing. The term *feedback* refers 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. 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. A consequence of this is that the behavior of feedback systems is often counter-intuitive, and it is therefore necessary to resort to formal methods to understand them.

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. Note that since the system is in a feedback loop, the choice of system 1 versus system 2 is somewhat arbitrary. It just depends where you want to start describing how the system works.

As the quote at the beginning of this chapter illustrates, a major source of examples of feedback systems is biology. Biological systems make use of feedback

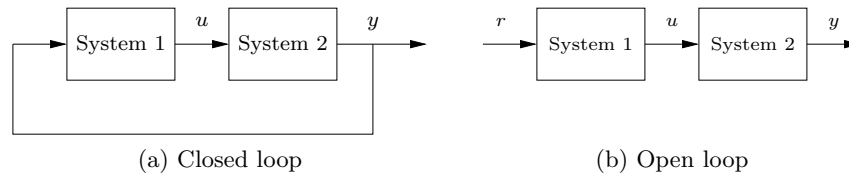


Figure 1.1: Open and closed loop systems. (a) The output of system 1 is used as the input of system 2, and the output of system 2 becomes the input of system 1, creating a closed loop system. (b) The interconnection between system 2 and system 1 is removed, and the system is said to be open loop.

in an extraordinary number of ways, on scales ranging from molecules to cells to organisms to ecosystems. One example is the regulation of glucose in the blood-stream 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. Referring to Figure 1.1, we can view the liver as system 1 and the pancreas as system 2. The output from the liver is the glucose concentration in the blood, and the output from the pancreas is the amount of insulin or glucagon produced. The interplay between insulin and glucagon secretions throughout the day helps 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 a 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 because of 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. We can model this system as a closed loop system by taking system 1 as the steam engine and system 2 as the governor. When properly designed, the flyball governor maintains a constant speed of the engine, roughly independent of the loading conditions. The centrifugal governor was an enabler of the successful Watt steam engine, which fueled the industrial revolution.

The examples given so far all deal with *negative feedback*, in which we attempt to react to disturbances in such a way that their effects decrease. *Positive feedback* is the opposite, where the increase in some variable or signal leads to a situation in which that quantity is further increased through feedback. 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. Encouragement is a type of positive feedback that is very useful in both industry and academia. Another common use of positive feedback is in the design of systems with oscillatory dynamics.

Feedback has many interesting properties that can be exploited in designing

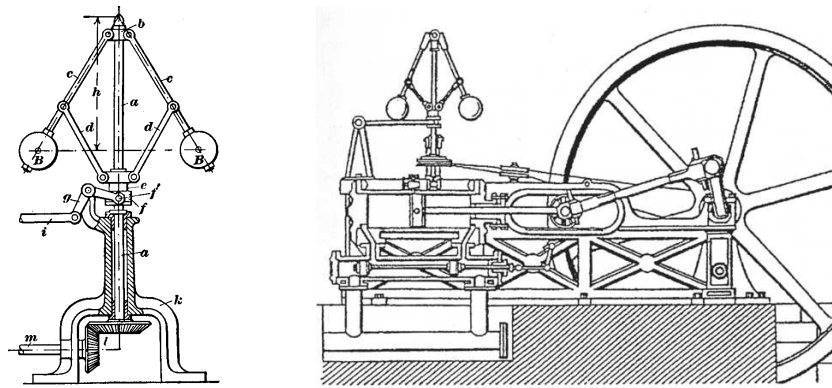


Figure 1.2: The centrifugal governor and the steam engine. The centrifugal governor on the left consists of a set of flyballs that spread apart as the speed of the engine increases. The steam engine on the right uses a centrifugal governor (above and to the left of the flywheel) to regulate its speed. (Credit: *Machine à Vapeur Horizontale* de Philip Taylor [1828].)

systems. As in the case of glucose regulation or the flyball governor, feedback can make a system resilient to 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 insensitive both to external disturbances and to variations in its individual elements.

Feedback has potential disadvantages as well. 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 a 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 (AFMs) 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, the oceans, the land, and the sun. Ecosystems are filled with examples of feedback due to the 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.

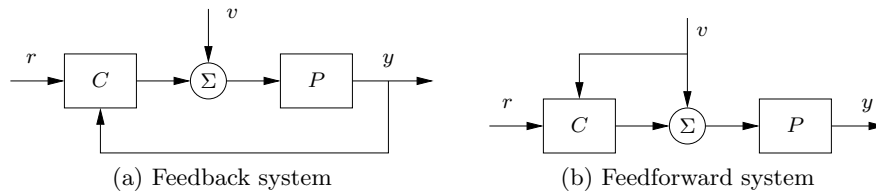


Figure 1.3: Feedback system versus feedforward system. In both systems we have a process P and a controller C . The feedback controller (a) measures the output y to determine the effect of the disturbance v , while the feedforward controller (b) measures the disturbance directly, but does not directly measure the process output.

1.2 What is Feedforward?

Feedback is reactive: there must be an error before corrective actions are taken. However, in some circumstances it is possible to measure a disturbance before it enters the system, and this information can then 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 in shaping the response to command signals, which are used as external inputs to the control system, 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 may be badly timed.

Figure 1.3 illustrates the difference between feedforward and feedback control. In both figures there is a reference signal r that describes the desired output of the process P and a disturbance signal v that represents an external perturbation to the process. In a feedback system, we measure the output y of the system and the controller C attempts to adjust the input to the process in a manner that causes the process output to maintain the desired the reference value. In a feedforward system, we instead measure the reference r and disturbance v and compute an input to the process that will create the desired output. Notice that the feedback controller does not directly measure the disturbance v while the feedforward controller does not measure the actual output y .

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 feedforward strategy corresponds to running a company based on extensive strategic planning, while a feedback strategy corresponds to a reactive approach. In biology, feedforward has been suggested as an essential element for motion control in humans that is tuned during training. Experience indicates that it is often advantageous to combine feedback and feedforward, and the correct balance requires insight and understanding of their respective properties, which are summarized in Table 1.1.

Table 1.1: Properties of feedback and feedforward.

Feedback	Feedforward
Closed loop	Open loop
Acts on deviations	Acts on plans
Robust to model uncertainty	Sensitive to model uncertainty
Risk for instability	No risk for instability

1.3 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, setpoint 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 it 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 additional desired behavior (good disturbance attenuation, fast responsiveness to changes in operating point, etc.). These properties are established using a variety of modeling and analysis techniques that capture the essential dynamics of the system and permit the exploration of possible behaviors in the presence of uncertainty, noise, and component failure.

A typical example of a control system is shown in Figure 1.4. The basic elements 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 operation, and uncertain dynamics in the 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*. The system can be influenced externally by an operator who introduces *command signals* to the system. These command signals can be reference values for the system output or may be more general descriptions of the task that the control system is supposed to implement.

Control engineering relies on and shares tools from physics (dynamics and modeling), computer science (information and software), and operations research (optimization, probability theory, 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

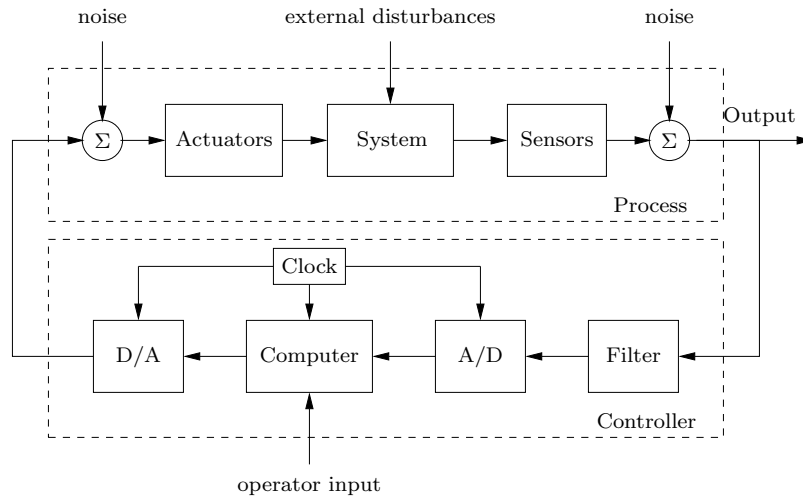


Figure 1.4: Components of a computer-controlled system. The upper dashed box represents the process dynamics, which includes the sensors and actuators in addition to the dynamical system being controlled. Noise and external disturbances can perturb the dynamics of the process. The controller is shown in the lower dashed box. It consists of a filter and analog-to-digital (A/D) and digital-to-analog (D/A) converters, as well as a computer that implements the control algorithm. A system clock controls the operation of the controller, synchronizing the A/D, D/A, and computing processes. The operator input is also fed to the computer as an external input.

the 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 are represented. Control relies on a type of input/output modeling that allows many new insights into the behavior of systems, such as disturbance attenuation and stable interconnection. Model reduction, where a simpler (lower-fidelity) description of the dynamics is derived from a high-fidelity model, is also naturally described in an input/output framework. Perhaps most importantly, modeling in a control context allows the design of *robust* interconnections between 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 can be very different from traditional computer software because of the central role of the dynamics of the system and the real-time nature of the implementation.

1.4 Uses of Feedback and Control

Feedback has many interesting and useful properties. It makes it possible to design precise systems from imprecise components and to make relevant quantities in a system change in a prescribed fashion. An unstable system can be stabilized

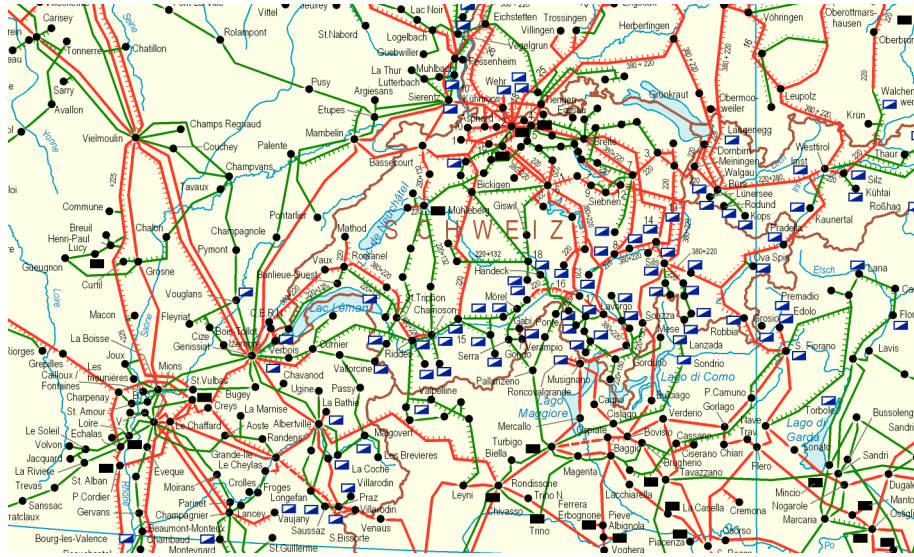


Figure 1.5: A small portion of the European power network. In 2016 European power suppliers operated a single interconnected network covering a region from the Arctic to the Mediterranean and from the Atlantic to the Urals. The installed power was more than 800 GW (8×10^{11} W) serving more than 500 million citizens. (Source: ENTSO-E <http://www.entsoe.eu>)

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 briefly survey some of the important applications and trends for feedback in the world around us. Considerably more detail is available in several reports describing advances and directions in the field of control [LLAE⁺17, Mur03, MAB⁺03, SA14].

Power Generation and Transmission Access to electrical power has been one of the major drivers of technological progress in modern society. Much of the early development of control was driven by the generation and distribution of electrical power. Control is mission critical for power systems, and there are many control loops in individual power stations. Control is also important for the operation of the whole power network since it is difficult to store energy and it is thus necessary to match production to consumption. Power management is a straightforward regulation problem for a system with one generator and one power consumer, but it is more difficult in a highly distributed system with many generators and long distances between consumption and generation. Power demand can change rapidly in an unpredictable manner, and combining generators and consumers into large networks makes it possible to share loads among many suppliers and to average consumption among many customers. Large transcontinental and transnational power systems have therefore been built, such as the one shown in Figure 1.5.

Telecommunications When telecommunications emerged in the early 20th century there was a strong need to amplify signals to enable telephone communication over long distances. The only amplifier available at the time was based on vacuum tubes. Since the properties of vacuum tubes are nonlinear and time varying, the amplifiers created a lot of distortion. A major advance was made when Black invented the negative feedback amplifier [Bla34, Bla77], which made it possible to obtain stable amplifiers with linear characteristics. Research on feedback amplifiers also generated fundamental understanding of feedback in the form of Nyquist's stability criterion [Nyg32] and Bode's methods for design of feedback amplifiers and his theorems on fundamental limits [Bod45]. Feedback is used extensively in cellular phones and networks, and the future 5G communication networks will permit execution of real-time control systems over the networks [TFKH16].

Ⓛ

Aerospace and Transportation In aerospace, control has been a key technological capability tracing back to the beginning of the 20th century. Indeed, the Wright brothers are correctly famous not for demonstrating simply 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 was followed by a fascinating success story of continuous improvements in flight control technology, culminating in the high-performance, highly reliable automatic flight control systems we see in modern commercial and military aircraft today.

Materials and Processing The chemical industry is responsible for the remarkable progress in developing new materials that are key to our modern society. 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 control system design increasingly challenging.

Instrumentation The measurement of physical variables is of prime interest in science and engineering. 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 the spring and the sensor. There is also a design compromise because a weak spring gives high sensitivity but low bandwidth. A different way of measuring acceleration is to use *force feedback*. The spring is replaced by a voice coil that is controlled so that the mass remains at a constant position. The acceleration is proportional to the current through the voice coil. In such an instrument, the precision depends

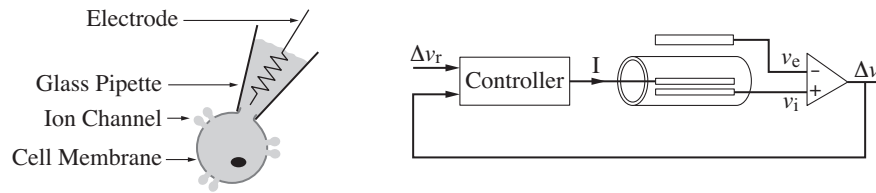


Figure 1.6: The voltage clamp method for measuring ion currents in cells using feedback. A pipette is used to place an electrode in a cell (left) and maintain the potential of the cell at a fixed level. The internal voltage in the cell is v_i , and the voltage of the external fluid is v_e . The feedback system (right) controls the current I into the cell so that the voltage drop across the cell membrane $\Delta v = v_i - v_e$ is equal to its reference value Δv_r . The current I is then equal to the ion current.

entirely on the calibration of the voice coil and does not depend on the sensor, which is used only as the feedback signal. The sensitivity/bandwidth compromise is also avoided.

Another important application of feedback is in instrumentation for biological systems. Feedback is widely used to measure ion currents in cells using a device called a *voltage clamp*, which is illustrated in Figure 1.6. Hodgkin and Huxley used the voltage clamp to investigate propagation of action potentials in the giant axon of the squid. In 1963 they shared the Nobel Prize in Medicine with Eccles for “their discoveries concerning the ionic mechanisms involved in excitation and inhibition in the peripheral and central portions of the nerve cell membrane.” A refinement of the voltage clamp called a *patch clamp* made it possible to measure exactly when a single ion channel is opened or closed. This was developed by Neher and Sakmann, who received the 1991 Nobel Prize in Medicine “for their discoveries concerning the function of single ion channels in cells.”

Robotics and Intelligent Machines 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 [ÅK14]. 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 the 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 recent areas of advancement in robotics and autonomous systems are (consumer) drones and autonomous cars, some examples of which are shown in Figure 1.7. Quadcopters such as the DJI Phantom make use of GPS receivers, accelerometers, magnetometers, and gyros to provide stable flight and also use stabilized camera platforms to provide high quality images and movies. Autonomous vehicles, such as the Google autonomous car project (now Waymo), make use of a variety of laser rangefinders, cameras, and radars to perceive their environment and then use sophisticated decision-making and control algorithms to enable them to safely drive in a variety of traffic conditions, from high-speed freeways to crowded city streets.



Figure 1.7: Autonomous vehicles. The figure on the left is a DJI Phantom 3 drone, which is able to maintain its position using GPS and inertial sensors. The figure on the right is an autonomous car that was developed by nuTonomy and is capable of driving on city streets by using sophisticated sensing and decision-making (control) software (photo courtesy Hyundai-Aptiv Autonomous Driving Joint Venture, LLC).

Networks and Computing Systems Control of networks is a large research 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 humans have ever built. Another is the decentralized nature of the control problem: 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 can be observed or relayed to controllers only after a delay, and the effect of a local control action can be felt throughout the network only after substantial delay.

Related to the control of networks is control of the servers that sit on these networks. Computers are key components of the systems of routers, web servers, and database servers used for communication, electronic commerce, advertising, and information storage. A typical example of a multilayer system for e-commerce is shown in Figure 1.8a. The system has several tiers of servers. The edge server accepts incoming requests and routes them to the HTTP server tier where they

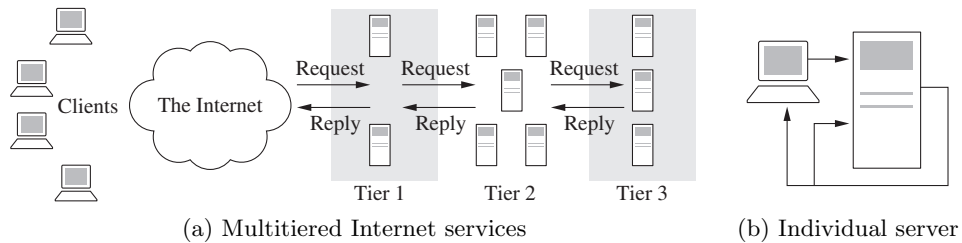


Figure 1.8: A multitier system for services on the Internet. In the complete system shown schematically in (a), users request information from a set of computers (tier 1), which in turn collect information from other computers (tiers 2 and 3). The individual server shown in (b) has a set of reference parameters set by a (human) system operator, with feedback used to maintain the operation of the system in the presence of uncertainty. (Based on Hellerstein et al. [HDPT04].)

are parsed and distributed to the application servers. The processing for different requests can vary widely, and the application servers may also access external servers managed by other organizations. Control of an individual server in a layer is illustrated in Figure 1.8b. A quantity representing the quality of service or cost of operation—such as response time, throughput, service rate, or memory usage—is measured in the computer. The control variables might represent incoming messages accepted, priorities in the operating system, or memory allocation. The feedback loop then attempts to maintain quality-of-service variables within a target range of values.

Economics The economy is a large dynamical system with many actors: governments, organizations, companies, and individuals. Governments control the economy through laws and taxes, the central banks by setting interest rates, and companies by setting prices and making investments. Individuals control the economy through purchases, savings, and investments. Many efforts have been made to model and control the system both at the macro level and at the micro level, but this modeling is difficult because the system is strongly influenced by the behaviors of the different actors in the system.

The financial system can be viewed as a global controller for the economy. Unfortunately this important controller does not always function as desired, as expressed in the following quote by Paul Krugman [Kru09]:

We have magneto trouble, said John Maynard Keynes at the start of the Great Depression: most of the economic engine was in good shape, but a crucial component, the financial system, was not working. He also said this: “We have involved ourselves in a colossal muddle, having blundered in the control of a delicate machine, the working of which we do not understand.” Both statements are as true now as they were then.

One of the reasons why it is difficult to model economic systems is that conservation laws for important variables are missing. A typical example is that the value of a company as expressed by its stock can change rapidly and erratically. There are, however, some areas with conservation laws that permit accurate modeling. One example is the flow of products from a manufacturer to a retailer, as illustrated in Figure 1.9. The products are physical quantities that obey a conservation law, and the system can be modeled by accounting for the number of products in the different inventories. There are considerable economic benefits in controlling supply chains so that products are available to customers while minimizing products that are in storage. Realistic supply chain problems are more complicated than indicated in the figure because there may be many different products, there may be different factories that are geographically distributed, and the factories may require raw material or subassemblies.

Feedback in Nature Many 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

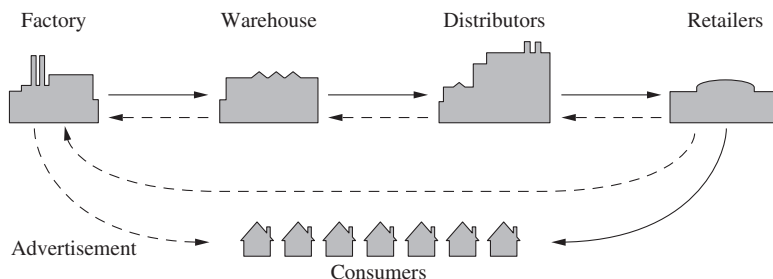


Figure 1.9: Supply chain dynamics (after Forrester [For61]). Products flow from the producer to the customer through distributors and retailers as indicated by the solid lines. There are typically many factories and warehouses and even more distributors and retailers. Dashed lines represent feedback and feedforward information flowing between the various agents in the chain. Multiple feedback loops are present as each agent tries to maintain the proper inventory level.

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. We briefly highlight three application areas here.

A major theme currently of interest to the biology community is the science of reverse (and eventually forward) engineering of biological control networks such as the one shown in Figure 1.10. There are a wide variety of biological phenomena that provide a rich source of examples of 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.

In contrast to individual cells and organisms, emergent properties of aggregations and ecosystems inherently reflect selection mechanisms that 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 the 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 [Kit04, SSS⁺04, YHSD00]. 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.

In nature, development of organisms and their control systems have often developed in synergy. The development of birds is an interesting example, as noted by John Maynard Smith in 1952 [Smi52]:

[T]he earliest birds, pterosaurs, and flying insects were stable. This

Ⓐ

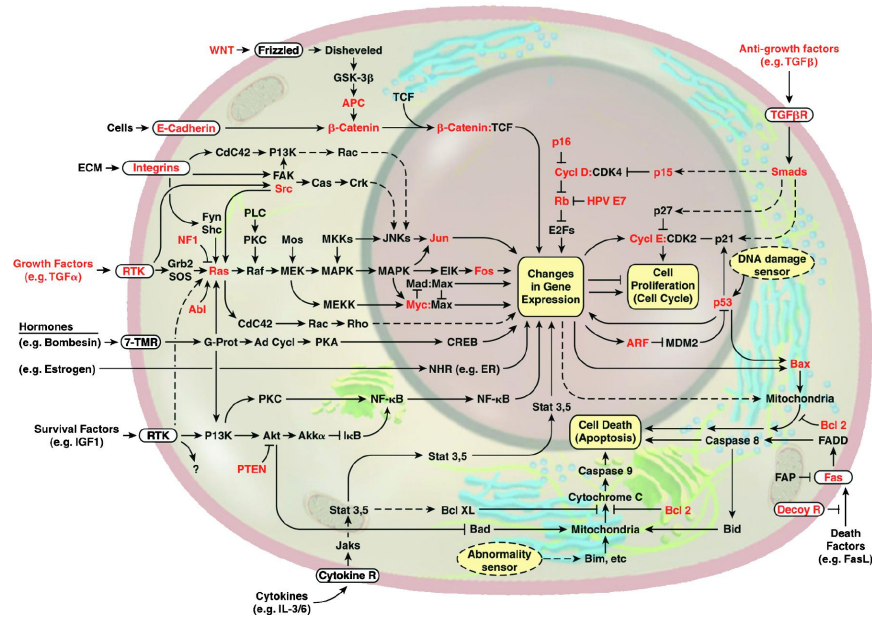


Figure 1.10: The wiring diagram of the growth-signaling circuitry of the mammalian cell [HW00]. The major pathways that are thought to play a role in cancer are indicated in the diagram. Lines represent interactions between genes and proteins in the cell. Lines ending in arrowheads indicate activation of the given gene or pathway; lines ending in a T-shaped head indicate repression. (Used with permission of Elsevier Ltd. and the authors.)

is believed to be because in the absence of a highly evolved sensory and nervous system they would have been unable to fly if they were not. ... To a flying animal there are great advantages to be gained by instability. The greater manoeuvrability is of equal importance to an animal which catches its food in the air and to the animals upon which it preys. ... It appears that in the birds and at least in some insects [...] the evolution of the sensory and nervous systems rendered the stability found in earlier forms no longer necessary.

1.5 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 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 previously. 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 subsequent chapters.

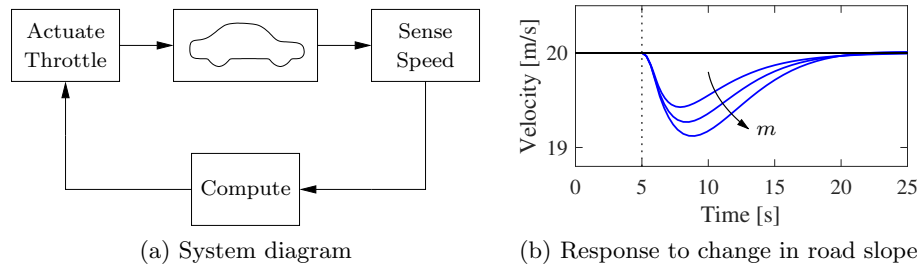


Figure 1.11: A feedback system for controlling the velocity of a vehicle. In the block diagram on the left, the velocity of the vehicle is measured and compared to the desired velocity within the “Compute” block. Based on the difference in the actual and desired velocities, the throttle (or brake) is used to modify the force applied to the vehicle by the engine, drivetrain, and wheels. The figure on the right shows how the velocity changes when the car travels on a horizontal road and the slope of the road changes to a constant uphill slope. The three different curves correspond to differing masses of the vehicle, between 1200 and 2000 kg, demonstrating that feedback can indeed compensate for the changing slope and that the closed loop system is robust to a large change in the vehicle characteristics.

Robustness to Uncertainty

One of the key uses of feedback is to provide robustness to uncertainty. For example, by measuring the difference between the sensed value of a regulated signal and its desired value, we can supply a corrective action to partially compensate for the effect of disturbances. This is precisely the effect that Watt exploited in his use of the centrifugal governor on steam engines. Another use of feedback is to provide robustness to variations in the process dynamics. 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, even if the process parameters are not directly measured. In this way, a feedback system provides robust performance in the presence of uncertain dynamics.

As an example, consider the simple feedback system shown in Figure 1.11. In this system, the velocity of a vehicle is controlled by adjusting the amount of gas flowing to the engine. Simple *proportional-integral* (PI) feedback is used to make the amount of gas depend on both the error between the current and the desired velocity and the integral of that error. The plot on the right shows the effect of this feedback when the vehicle travels on a horizontal road and it encounters an uphill slope. When the slope changes, the car decelerates due to gravity forces and the velocity initially decreases. The velocity error is sensed by the controller, which acts to restore the velocity to the desired value by increasing the throttle. The figure also shows what happens when the same controller is used for different masses of the car, which might result from having a different number of passengers or towing a trailer. Notice that the steady-state velocity of the vehicle always approaches the desired velocity and achieves that velocity within approximately 15 s, independent of the mass (which varies by a factor of $\pm 25\%$). Thus feedback improves both performance and robustness of the system.

Another early example of the use of feedback to provide robustness is the neg-

ative feedback amplifier. When telephone communications were developed, amplifiers were used to compensate for signal attenuation in long lines. A 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 the development of line amplifiers for a long time. A major breakthrough was the invention of the feedback amplifier in 1927 by Harold S. Black, an electrical engineer at Bell Telephone Laboratories. Black used *negative feedback*, which reduces the gain but makes the amplifier insensitive to variations in tube characteristics. This invention made it possible to build stable amplifiers with linear characteristics despite the nonlinearities of the vacuum tube amplifier.

Design of Dynamics

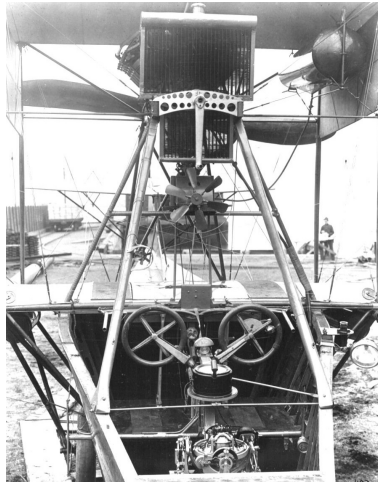
Another use of feedback is to change the dynamics of a system. Through feedback, we can alter the behavior of a system to meet the needs of an application: systems that are unstable can be stabilized, systems that are sluggish can be made responsive, and systems that have drifting operating points can 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 presented 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 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, the Wright Flyer, that was unstable but maneuverable. The Flyer 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. The Wright Brothers were well aware of the compromise between stability and maneuverability when they designed the Wright Flyer [Dra55] and they made the first successful flight at Kitty Hawk in 1903. Modern fighter airplanes are also unstable in certain flight regimes, such as take-off and landing.

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



(a) Sperry autopilot



(b) 1912 Curtiss biplane

Figure 1.12: Aircraft autopilot system. The Sperry autopilot (a) contained a set of four gyros coupled to a set of air valves that controlled the wing surfaces. The 1912 Curtiss used an autopilot to stabilize the roll, pitch, and yaw of the aircraft and was able to maintain level flight as a mechanic walked on the wing (b) [Hug93].

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 was the first use of feedback in aeronautical engineering, and Sperry won a prize in a competition for the safest airplane in Paris in 1914. Figure 1.12 shows the Curtiss 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.

Creating Modularity

Feedback can be used to create modularity and shape well-defined relations between inputs and outputs in a structured hierarchical manner. A modular system is one in which individual components can be replaced without having to modify the entire system. By using feedback, it is possible to allow components to maintain their input/output properties in a manner that is robust to changes in its interconnections. A typical example is the electrical drive system shown in Figure 1.13, which has an architecture with three cascaded loops. The innermost loop is a current loop, where the current controller (CC) drives the amplifier so that the current to the motor tracks a commanded value (often called the “setpoint”). The middle feedback loop uses a velocity controller (VC) to drive the setpoint of the current controller so that velocity follows its commanded value. The outer loop drives the setpoint of the velocity loop to follow the setpoint of the position controller PC.

The control architecture with nested loops shown in Figure 1.13 is common. It simplifies design, commissioning, and operation. Consider for example the design of

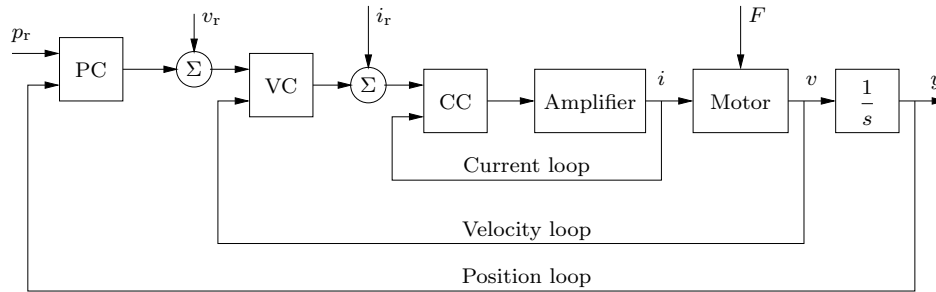


Figure 1.13: Block diagram of a system for position control. The system has three cascaded loops for control of current, velocity, and position. Each loop has an externally supplied reference value (denoted by the subscript ‘r’) that sets the nominal value of the input to the loop, which is added to output from next outermost loop to determine the commanded value for the loop (called the “setpoint”).

the velocity loop. With a well-designed current controller the motor current follows the setpoint of the controller CC. Since the motor torque is proportional to the current, the dynamics relating motor velocity to the input of the current controller is approximately an integrator. This simplified model can be used to design the velocity loop so that effects of friction and other disturbances are reduced. With a well-designed velocity loop, the design of the position loop is also simple. The loops can also be tuned sequentially starting with the inner loop.

This architecture illustrates how feedback can be used to simplify the overall design of the controller by breaking the problem into stages. This architecture also provides a level of modularity since each design stage depends only on the closed loop behavior of the system. If we replace the motor with a new motor, then by redesigning the current controller (CC) to give the same closed loop performance, we can leave the outer level loops unchanged. Similarly, if we need to redesign one of the outer layer controllers for an application with different specifications, we can often make use of an existing inner loop design (as long as the existing design provides enough performance to satisfy the outer loop requirements).

Challenges of Feedback

While feedback has many advantages, it also has some potential drawbacks. Chief among these is the possibility of 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 feedback instability, something that we obviously want to avoid. This is tricky because we must design the system not only to be stable under nominal conditions but also to 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 is that feedback often injects measurement noise into the system. Measurements must be carefully filtered so that the actuation and process dynamics do not respond to them, while at the same time ensuring that the measurement signal from the sensor is properly coupled into the closed loop dynamics (so that the proper levels of performance are achieved).

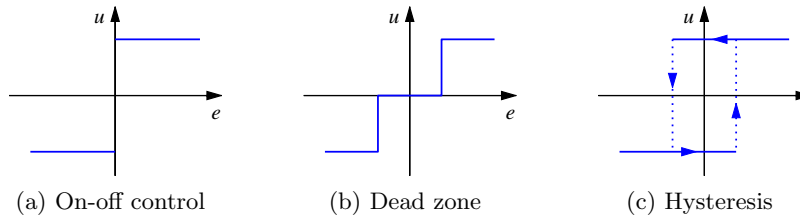


Figure 1.14: Input/output characteristics of on-off controllers. Each plot shows the input on the horizontal axis and the corresponding output on the vertical axis. Ideal on-off control is shown in (a), with modifications for a dead zone (b) or hysteresis (c). Note that for on-off control with hysteresis, the output depends on the value of past inputs.

Another potential drawback of control is the complexity of embedding a control system into a product. While the cost of sensing, computation, and actuation has decreased dramatically in the past few decades, the fact remains that control systems are often 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 be met only 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.

1.6 Simple Forms of Feedback

The idea of feedback to make corrective actions based on the difference between the desired and the actual values of a quantity 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 proportional-integral-derivative 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.

On-Off Control

A simple feedback mechanism can be described as follows:

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

where the *control error* $e = r - y$ is the difference between the reference (or command) signal r and the output of the system y , and u is the actuation command. 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 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 make modifications by introducing either a dead zone or hysteresis (see Figures 1.14b and 1.14c).

PID Control

The reason why on-off control often gives rise to oscillations is that the system overreacts since a small change in the error makes 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 with the control law

$$u = \begin{cases} u_{\max} & \text{if } e \geq e_{\max}, \\ k_p e & \text{if } e_{\min} < e < e_{\max}, \\ u_{\min} & \text{if } e \leq e_{\min}, \end{cases} \quad (1.2)$$

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

$$u = k_p(r - y) = k_p e \quad \text{if } e_{\min} \leq e \leq e_{\max}. \quad (1.3)$$

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:

$$u(t) = k_i \int_0^t e(\tau) d\tau. \quad (1.4)$$

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 has zero steady-state error (Exercise 1.5). The catch is that there may not always be a steady state because the system may be oscillating. In addition, if the control action has magnitude limits, as in equation (1.2), an effect known as “integrator windup” can occur and may result in poor performance unless appropriate “anti-windup” compensation is used. Despite the potential drawbacks, which can be overcome with careful analysis and design, the benefits of integral feedback in providing zero error in the presence of constant disturbances have made it one of the most used forms of feedback.

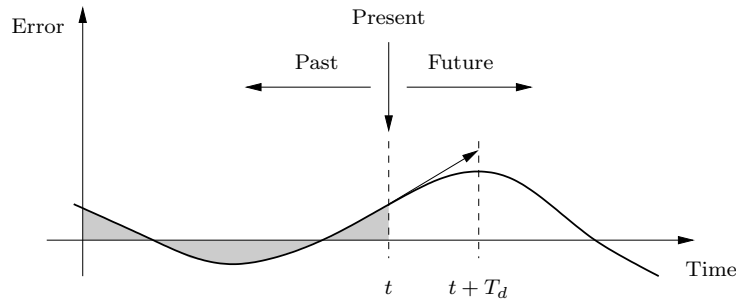


Figure 1.15: Action of a PID controller. At time t , the proportional term depends on the instantaneous value of the error. The integral portion of the feedback is based on the integral of the error up to time t (shaded portion). The derivative term provides an estimate of the growth or decay of the error over time by looking at the rate of change of the error. T_d represents the approximate amount of time in which the error is projected forward (see text).

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

$$u(t) = k_p e(t) + k_i \int_0^t e(\tau) d\tau + k_d \frac{de(t)}{dt}. \quad (1.5)$$

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

A PID controller is very useful and is capable of solving a wide range of control problems. More than 95% of all industrial control problems are solved by PID control, although many of these controllers are actually *proportional-integral (PI) controllers* because derivative action is often not included [DM02a].

1.7 Combining Feedback with Logic

Continuous control is often combined with logic to cope with different operating conditions. Logic is typically related to changes in operating mode, equipment protection, manual interaction, and saturating actuators. One situation is when there is one variable that is of primary interest, but other variables may have to be controlled for equipment protection. For example, when controlling a compressor the outflow is the primary variable but it may be necessary to switch to a different mode to avoid compressor stall, which may damage the compressor. We illustrate some ways in which logic and feedback are combined by a few examples.

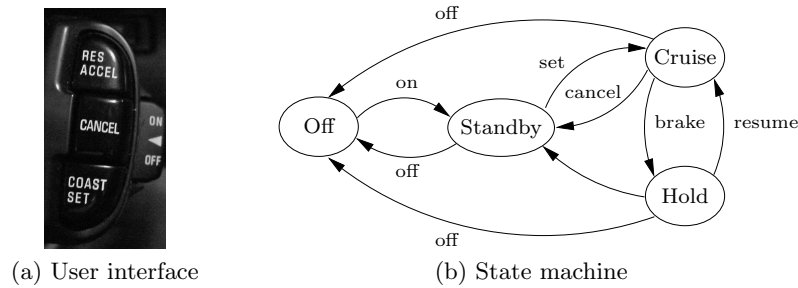


Figure 1.16: Finite state machine for cruise control system. The figure on the left shows some typical buttons used to control the system. The controller can be in one of four modes, corresponding to the nodes in the diagram on the right. Transition between the modes is controlled by pressing one of the four buttons on the cruise control interface: on/off, set, resume, or cancel.

Cruise control

The basic control function in a cruise controller, such as the one shown in Figure 1.11, is to keep the velocity constant. It is typically done with a PI controller. The controller normally operates in automatic mode but it is necessary to switch it off when braking, accelerating, or changing gears. The cruise control system has a human-machine interface that allows the driver to communicate with the system. There are many different ways to implement this system; one version is illustrated in Figure 1.16a. The system has four buttons: on/off, coast/set, resume/accelerate, and cancel. The operation of the system is governed by a finite state machine that controls the modes of the PI controller and the reference generator, as shown in Figure 1.16b.

The finite state machine has four modes: off, standby, cruise, and hold. The state changes depending on actions of the driver who can brake, accelerate, and operate using the buttons. The on/off switch moves the states between off and standby. From standby the system can be moved to cruise by pushing the set/coast button. The velocity reference is set as the velocity of the car when the button is released. In the cruise state the operator can change the velocity reference; it is increased using the resume/accelerate button and decreased using the set/coast button. If the driver accelerates by pushing the gas pedal the speed increases, but it will go back to the set velocity when the gas pedal is released. If the driver brakes then the car slows, and the cruise controller goes into hold but it remembers the setpoint of the controller. It can be brought to the cruise state by pushing the resume/accelerate button. The system also moves from cruise mode to standby if the cancel button is pushed. The reference for the velocity controller is remembered. The system goes into off mode by pushing on/off when the system is engaged.

The PI controller is designed to have good regulation properties and to give good transient performance when switching between resume and control modes.



Figure 1.17: Large computer “server farm.” The National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory. (Figure courtesy U.S. Department of Energy)

Server Farms

Server farms consist of a large number of computers for providing Internet services (cloud computing). Large server farms, such as the one shown in Figure 1.17, may have thousands of processors. Power consumption for driving the servers and for cooling them is a prime concern. The cost for energy can be more than 40% of the operating cost for data centers [EKR03]. The prime task of the server farm is to respond to a strongly varying computing demand. There are constraints given by electricity consumption and the available cooling capacity. The throughput of an individual server depends on the clock rate, which can be changed by adjusting the voltage applied to the system. Increasing the supply voltage increases the energy consumption and more cooling is required.

Control of server farms is often performed using a combination of feedback and logic. Capacity can be increased rapidly if a server is switched on simply by increasing the voltage to a server, but a server that is switched on consumes energy and requires cooling. To save energy it is advantageous to switch off servers that are not required, but it takes some time to switch on a new server. A control system for a server farm requires individual control of the voltage and cooling of each server and a strategy for switching servers on and off. Temperature is also important. Overheating will reduce the life time of the system and may even destroy it. The cooling system is complicated because cooling air goes through the servers in series and parallel. The measured value for the cooling system is therefore the server with the highest temperature. Temperature control is accomplished by a combination of feedforward logic to determine when servers are switched on and off and feedback using PID control.

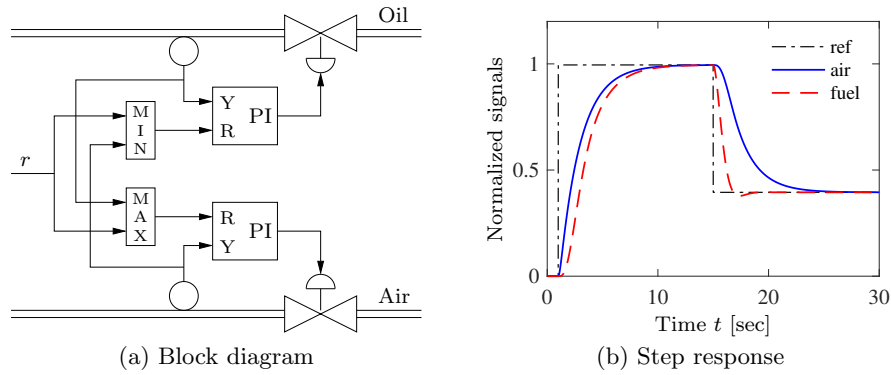


Figure 1.18: Air-fuel controller based on selectors. The left figure shows the system architecture. The letters R and Y in the PI controller denote the input ports for reference and measured signal respectively. The right figure shows a simulation where the power reference r is changed stepwise at $t = 1$ and $t = 15$. Notice that the normalized air flow is larger than the normalized fuel flow both for increasing and decreasing reference steps.

Air-Fuel Control

Air-fuel control is an important problem for ship boilers. The control system consists of two loops for controlling air and oil (fuel) flow and a *supervisory controller* that adjusts the air-fuel ratio. The ratio should be adjusted for optimal efficiency when the ships are on open sea but it is necessary to run the system with air excess when the ships are in the harbor, since generating black smoke will result in heavy penalties.

An elegant solution to the problem can be obtained by combining PI controllers with maximum and minimum selectors, as shown in Figure 1.18a. A *selector* is a static system with several inputs and one output: a maximum selector gives an output that is the largest of the inputs, a minimum selector gives an output that is the smallest of the inputs. Consider the situation when the power demand is increased: the reference r to the air controller is selected as the commanded power level by the maximum selector, and the reference to the oil flow controller is selected as the measured airflow. The oil flow will lag the air flow and there will be air excess. When the commanded power level is decreased, the reference of the oil flow controller is selected as the power demand by the minimum selector and the reference for the air flow controller is selected as the oil flow by the maximum selector. The system then operates with air excess when power is decreased.

The resulting response of the system for step changes in the desired power level is shown in Figure 1.18b, verifying that the system maintains air excess for both power increases and decreases.

Selectors are commonly used to implement logic in engines and power systems. They are also used for systems that require very high reliability: by introducing three sensors and only accepting values where two sensors agree it is possible to guard for the failure of a single sensor.

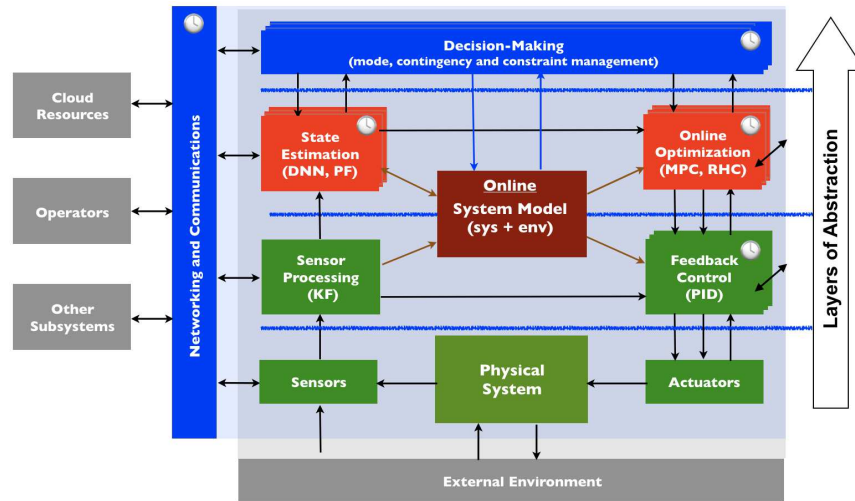


Figure 1.19: Layered decomposition of a control system.

1.8 Control System Architectures

Most of the control systems we are investigating in this book will be relatively simple feedback loops. In this section we will try to give a glimpse of the fact that in reality the simple loops combine to form a complex system that often has a hierarchical structure with controllers, logic, and optimization in different combinations. Figure 1.19 shows one representation of such a hierarchy, exposing different “layers” of the control system. This class of systems is discussed in more detail in Chapter 15. We focus here on a few representative examples to illustrate some basic points.

Freight Train Trip Optimizer

An example of two of the layers represented in Figure 1.19 can be seen in the control of modern locomotives developed by General Electric (GE). Typical requirements for operating a freight train are to arrive on time and to use as little fuel as possible. The key issue is to avoid unnecessary braking. Figure 1.20 illustrates a system developed by GE. At the low layer the train has a speed regulator and simple logic to avoid entering a zone where there is another train. The key disturbance for the speed control is the slope of the ground. The speed controller has a model of the track, a GPS sensor, and an estimator. The setpoint for the speed controller is obtained from a trip optimizer, which computes a driving plan that minimizes the fuel consumption while arriving at the desired arrival time. The arrival time is provided by a dispatch center, which in turn may use its own optimization. These optimizations represent the second layer in Figure 1.19, with the top layer (decision-making) provided by the human operator.

Diesel-electric freight locomotives pull massive loads of freight cars, weighing more than 20,000 tons (US), and may be more than a mile in length. A typical

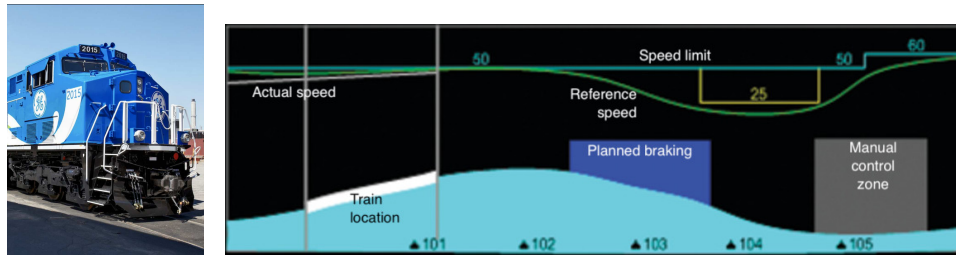


Figure 1.20: Freight train trip optimizer. GE's Trip Optimizer™ takes data about the train, the terrain, and the propulsion system and computes the best speed for the train in order to reach the destination on time while burning the least amount of diesel fuel. (Figure courtesy GE.)

locomotive burns about 35,000 gallons per year and can save an average 10% using the Trip Optimizer autopilot, representing a substantial savings in cost, natural resources, and pollution.

Process Control Systems

Process control systems are used to monitor and regulate the manufacturing of a wide range of chemicals and materials. One example is a paper factory, such as the one depicted in Figure 1.21. The factory produces paper for a variety of purposes from logs of wood. There are multiple fiber lines and paper machines, with a few dozen mechanical and chemical production processes that convert the logs to a slurry of fibers in different steps, and then paper machines that convert the fiber slurry to paper. Each production unit has PI(D) controllers that control flow, temperature, and tank levels. The loops typically operate on time scales from fractions of seconds to minutes. There is logic to make sure that the process is safe and there is sequencing for start, stop, and production changes. The setpoints of the low level control loops are determined from production rates and recipes, sometimes using optimization. The operation of the system is governed by a supervisory system that measures tank levels and sets the production rates of the different production units. This system performs optimization based on demanded production, measurement of tank levels, and flows. The optimization is performed at the time scale of minutes to hours, and it is constrained by the production rates of the different production units. Processes for continuous production in the chemical and pharmaceutical industry are similar to the paper factory but the individual production units may be very different.

One of the features of modern process control systems is that they operate across many time and spatial scales. Modern process control systems are also integrated with supply chains and product distribution chains, leading to the use of production planning systems and enterprise resource management systems. An example of an architecture for a distributed control system (DCS), typical for complex manufacturing systems, is shown in Figure 1.22.

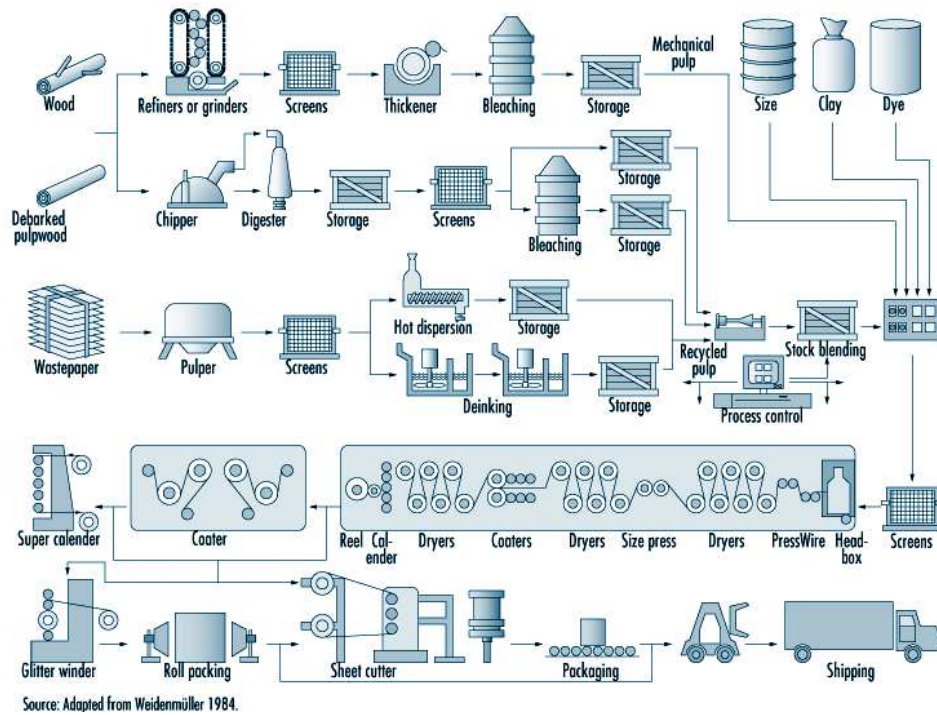


Figure 1.21: Schematic diagram for a pulp and paper manufacturing plant. The input to the plant is wood (upper left), which is then processed through a number of stages to create paper products.

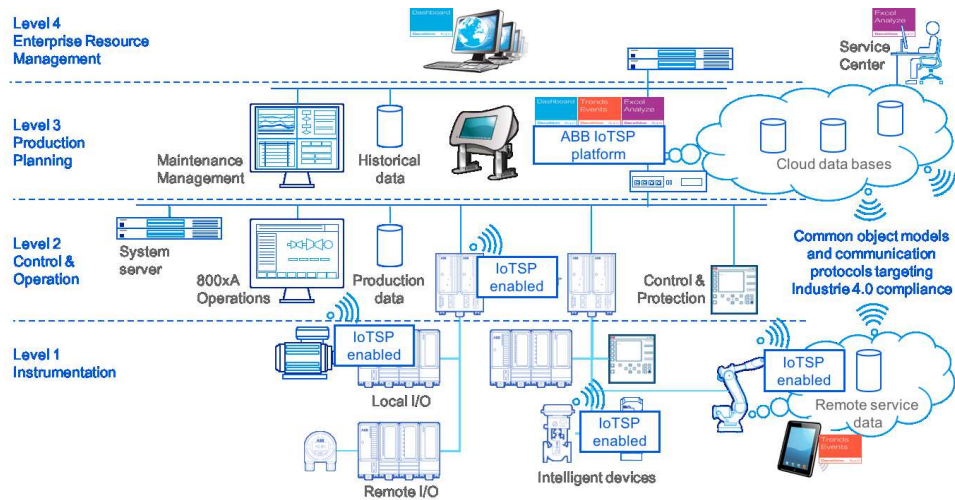


Figure 1.22: Functional architecture of process control system, implemented as a distributed control system (DCS). Figure courtesy of ABB, Inc.

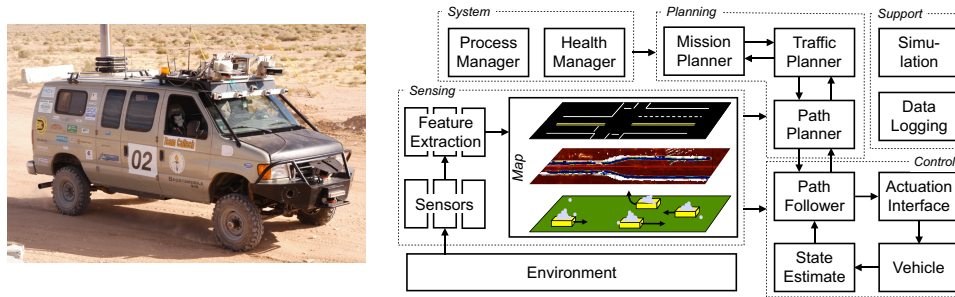


Figure 1.23: DARPA Grand Challenge. “Alice,” Team Caltech’s entry in the 2005 and 2007 competitions and its networked control architecture [CFG⁺06].

Autonomous Driving

The cruise controller in Figure 1.11 relieves the driver of one task, to keep constant speed, but a driver still has many tasks to perform: plan the route, avoid collisions, decide the proper speed, perform lane changes, make turns, and keep proper distance to the car ahead. Car manufacturers are continuously automating more and more of these functions, going as far as automatic driving. An example of a control system for an autonomous vehicle is shown in Figure 1.23. This control system is designed for driving in urban environments. The feedback system fuses data from road and traffic sensors (cameras, laser range finders, and radar) to create a multi-layer “map” of the environment around the vehicle. This map is used to make decisions about actions that the vehicle should take (drive, stop, change lanes) and plan a specific path for the vehicle to follow. An optimization-based planner is used to compute the trajectory for the vehicle to follow, which is passed to a trajectory tracking (path following) module. A supervisory control module performs higher-level tasks such as mission planning and contingency management (if a sensor or actuator fails).

We see that this architecture has the basic features shown in Figure 1.19. The control layers are shown in the planning and control blocks, with the mission planner and traffic planner representing two levels of discrete decision-making logic, the path planner representing a trajectory optimization function, and then the lower layers of control. Similarly, there are multiple layers of sensing, with low level information, such as vehicle speed and position in the lane, being sent to the trajectory tracking controller, while higher level information about other vehicles on the road and their predicted motions is sent to the trajectory, traffic, and mission planners.

1.9 Further Reading

The material in the first half of this chapter draws from the report of the Panel on Future Directions on Control, Dynamics and Systems [Mur03]. Several additional papers and reports have highlighted the successes of control [NS99] and new vistas in control [Bro00, Kum01, LLAE⁺17, SA14, Wis07]. The early development of control is described by Mayr [May70] and in the books by Bennett [Ben79, Ben93], which cover the period 1800–1955. A fascinating examination of some of the early history

of control in the United States has been written by Mindell [Min02]. 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-Naeini [FPEN05], Dorf and Bishop [DB04], Kuo and Golnaraghi [KG02], and Seborg, Edgar, and Mellichamp [SEM04] are widely used. More mathematically oriented treatments of control theory include Sontag [Son98] and Lewis [Lew03]. At the opposite end of the spectrum, the textbook *Feedback Control for Everyone* [AM10] provides a readable introduction with minimal mathematical background required. The books by Hellerstein et al. [HDPT04] and Janert [Jan14] provide descriptions of the use of feedback control in computing systems. A number of books look at the role of dynamics and feedback in biological systems, including Milhorn [Mil66] (now out of print), J. D. Murray [Mur04], and Ellner and Guckenheimer [EG05]. The book by Fradkov [Fra07] and the tutorial article by Bechhoefer [Bec05] cover many specific topics of interest to the physics community.

Systems that combine continuous feedback with logic and sequencing are called *hybrid systems*. The theory required to properly model and analyze such systems is outside the scope of this text, but a comprehensive description is given by Goebel, Sanfelice, and Teele [GST12]. It is very common that practical control systems combine feedback control with logic sequencing and selectors; many examples are given by Åström and T. Hägglund [ÅH06].

Exercises

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

1.2 (Balance systems) Balance yourself on one foot with your eyes closed for 15 s. Using Figure 1.4 as a guide, describe the control system responsible for keeping you from falling down. Note that the “controller” will differ from that in the diagram (unless you are an android reading this in the far future).

1.3 (Eye motion) Perform the following experiment and explain your results: Holding your head still, move one of your hands left and right 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 it. Now hold your hand still and shake your head left to right, once again recording how quickly you can move before losing track of your hand. Explain any difference in performance by comparing the control systems used to implement these behaviors.



1.4 (Cruise control) Download the MATLAB code used to produce simulations for the cruise control system in Figure 1.11 from the companion web site. Using trial and error, change the parameters of the control law so that the overshoot in speed is not more than 1 m/s for a vehicle with mass $m = 1200$ kg. Does the same controller work if we set $m = 2000$ kg?

1.5 (Integral action) We say that a system with a constant input reaches steady state if all system variables approach constant values 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. Notice that there is no saturation in the controller.

1.6 (Combining feedback with logic) Consider a system for cruise control where the overall function is governed by the state machine in Figure 1.16. Assume that the system has a continuous input for vehicle velocity, discrete inputs indicating braking and gear changes, and a PI controller with inputs for the reference and measured velocities and an output for the control signal. Sketch the actions that have to be taken in the states of the finite state machine to handle the system properly. Think about if you have to store some extra variables, and if the PI controller has to be modified.

1.7 Search the web and pick an article in the popular press about a feedback and control system. Describe the feedback system using the terminology given in the article. In particular, identify the control system and describe (a) the underlying process or system being controlled, along with the (b) sensor, (c) actuator, and (d) computational element. If the some of the information is not available in the article, indicate this and take a guess at what might have been used.

