1

Introduction

1.1 Introduction

Control systems are ubiquitous. They appear in our homes, in cars, in industry and in systems for communication and transport, just to give a few examples. Control is increasingly becoming mission critical, processes will fail if the control does not work. Control has been important for design of experimental equipment and instrumentation used in basic sciences and will be even more so in the future. Principles of control also have an impact on such diverse fields as economics, biology, and medicine.

Control, like many other branches of engineering science, has developed in the same pattern as natural science. Although there are strong similarities between natural science and engineering science it is important to realize that there are some fundamental differences. The inspiration for natural science is to understand phenomena in nature. This has led to a strong emphasis on analysis and isolation of specific phenomena, so called reductionism. A key goal of natural science is to find basic laws that describe nature. The inspiration of engineering science is to understand, invent, and design man-made technical systems. This places much more emphasis on interaction and design. Interaction is a key feature of practically all man made systems. It is therefore essential to replace reductionism with a holistic systems approach. The technical systems are now becoming so complex that they pose challenges comparable to the natural systems. A fundamental goal of engineering science is to find system principles that make it possible to effectively deal with complex systems. Feedback, which is at the heart of automatic control, is an example of such a principle.

A simple form of feedback consists of two dynamical systems connected in a closed loop which creates an interaction between the systems. Simple causal reasoning about such a 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 difficult and it is necessary to analyze the system as a whole. A consequence of this is that the behavior of a feedback system is often counterintuitive. To understand feedback systems it is therefore necessary to resort to formal methods based on mathematics.

Feedback has many advantages. It is possible to create linear behavior out of nonlinear components. Feedback can make a system very resilient towards external influences. The total system can be made very insensitive to external disturbances and to variations in its individual components. Feedback has one major disadvantage, it may create instability, which is intrinsically a dynamic phenomenon. To understand feedback systems it is therefore necessary to have a good insight into dynamics.

The wide applicability of control has many advantages. Since control can be used in so many different fields, it is a very good vehicle for technology transfer. Ideas invented in one field can be applied to another technical field.

Control is inherently multidisciplinary. A typical control system contains sensors, actuators, computers and software. Analysis of design of control systems require domain knowledge about the particular process to be controlled, knowledge of the techniques of control and specific technology used in sensors and actuators. Controllers are typically implemented using digital computers. Knowledge about real time computing and software is therefore also essential. Sensors and actuators are often often connected by communication networks. This implies that knowledge about communication is also important. In the future we can see a convergence of the technologies of control, computing and communication.

Team work is essential in control because of the wide range of technologies and techniques involved. Education in control has proven to be an excellent background when working with complex engineering systems. The interdisciplinary nature of control has created some difficulties for educators. Education and research in engineering grew out of specific technologies such as mining, building of roads and dams, construction of machines, generation and transmission of electricity, and industrial use of chemistry. This led to an organization of engineering schools based on departments of mining, civil engineering, mechanical engineering, electrical engineering, and chemical engineering etc. This served very well in the end of the 19th century and the beginning of the 20th century. The situation changed significantly with the advent of fields like control, that cut cross traditional department boundaries. Industry has adapted quickly to the new demands but academia has not.

There are many reasons why an engineer should know control. First of all because practically all engineers will use control, and some will design control systems. But the most important reason is that control is an essential element of practically all engineering systems. It happens too often that systems perform poorly because they are designed from purely static analysis with no consideration of dynamics and control. This can be avoided by engineers being aware of control even if they are not specialists. Control can also give designers extra degrees of freedom. It is in fact a very powerful tool for designers of all systems. Cars are typical examples. The stringent requirements on emission were solved by controlling the combustion engines. Other examples are anti-lock braking systems (ABS) and systems for traction control. Other reasons to study control is that there are many beautiful theoretical results and some really neat devices.

Control has for a long time been confined to engineering but it is increasingly clear that the ideas and concepts have a much wider use. The concepts of feedback and control are thus essential in understanding biological and economical systems. We illustrate this with a quote from the book Way Life Works : The Science Lover's Illustrated Guide to How Life Grows, Develops, Reproduces, and Gets Along – by Mahlon Hoagland, Bert Dodson.

> Feedback is a central feature of life: All organisms share the ability to sense how they are doing and to make changes in "mid-flight" if necessary. 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. This apparent purposefulness, largely unconscious, operates at every level - from the interaction of proteins in cells to the interaction of organisms in complex ecologies.

It is thus reasonable to claim that control not only makes our lives more comfortable it is also essential for our existence.

The rest of this chapter gives a brief history of the development of the field of control. The richness of the field is then illustrated by a number of examples from a wide range of applications ranging from industrial applications to biology.

1.2 A Brief History

Although there are early examples of the use of feedback in ancient history, the development of automatic control is strongly connected to the industrial revolution and the development of modern technology. When new sources of power were discovered the need to control them immediately arose. When new production techniques were developed there were needs to keep them operating smoothly with high quality.

The Centrifugal Governor

The centrifugal governor is one of the most celebrated early feedback systems. It was first used to control the speed of windmills. Later around 1788 it was used by James Watt to control the velocity of steam engines. A textile mill is a typical application, one steam engine drives several spinning wheels and looms. The power from the steam engine is transmitted to the spinning wheels and looms via belt-drives. It is highly desirable to keep the speed of the steam engine constant because changes in speed may cause thread breaks and require adjustments of the looms. It was observed that engine speed changed with changes in the load, for example when a loom was connected to the drive belt. The centrifugal governor was introduced in order to keep the speed constant. Figure 1.1 shows a typical system. It consists of two balls hinged on a rotating shaft which is connected to the output shaft of the steam engine. When the speed increases, the balls swing out. This motion is connected to the valve which admits steam into the engine via mechanical links. The connection is made in such a way that steam flow increases when the velocity decreases. The system is a feedback system because changes in the velocity are fed back to the steam valve. The feedback is negative because the the steam supply is increased when the velocity decreases.

The improvement obtained when using a centrifugal governor is illustrated in Figure 1.2. The figure shows that the velocity drops when an additional loom is connected to the drive belt. The figure also shows that the velocity drop is significantly smaller when a centrifugal governor is used.

It is possible to change the characteristics of the governor by changing the mechanism that transmits the motion of the balls to the steam valve. To describe this we introduce the notion of gain of the governor. This is the ratio of the change in steam valve opening (Δu) to the change in the angle (Δv) of the velocity. See Figure 1.2 which shows how the velocity responds to changes in the load. The figure shows that the largest velocity error decreases with decreasing gain of the controller but also that there is a tendency for oscillations that increases with increasing gain. The centrifugal governor was a very successful device that drastically simplified the operation of steam driven textile mills.

The action of the basic centrifugal governor can crudely be describe with the equation

$$u = k(V_r - V) + b$$

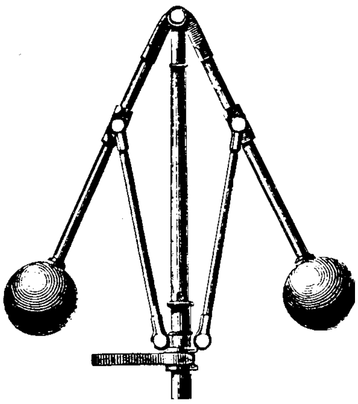

Figure 1.1 The centrifugal governor, which has been used to control the speed of engines since the beginning of the industrial revolution. When the axis spins faster the balls move away from the axis. The motion is transfered to the engine to reduce its power. The governor has also become an icon of the field of control.

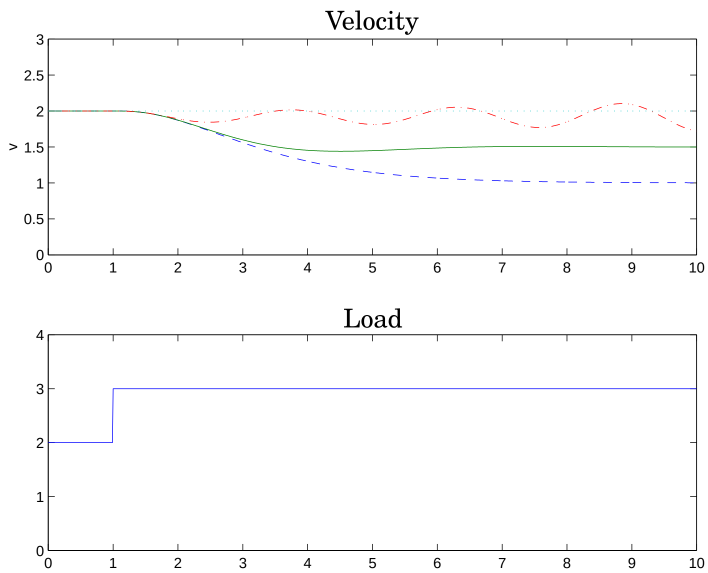

Figure 1.2 Response of the velocity to changes in the load of an engine controlled by a governor with different values of the gain, $k = 0$ (dashed), $k = 1$ (full), $k = 12$ (dash-dotted).

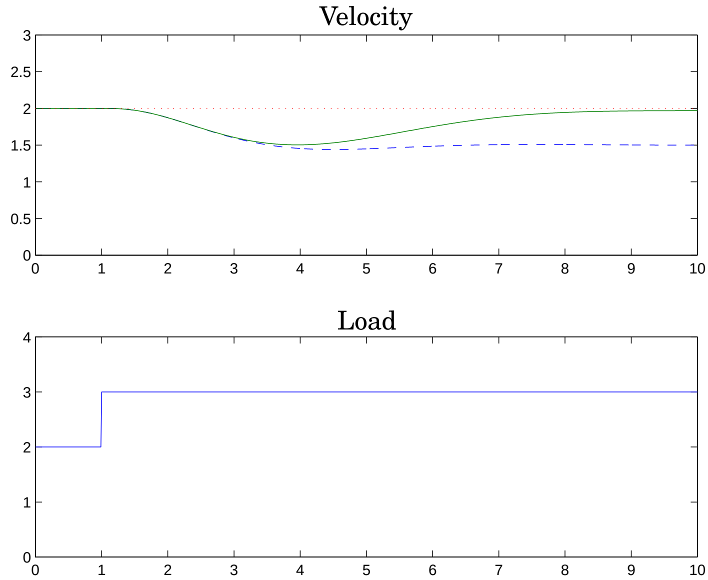

Figure 1.3 Response of velocity to changes in the load of an engine controlled by a governor having proportional control (dashed line) and PI control (full line).

where u is the opening of the steam valve, V_r is the desired speed, V the actual speed and k and b are constants. This is called a proportional controller because the control action u is proportional to the error. The parameter b is a bias term that was adjusted manually to make sure that the V was equal to V_r. Siemens made a clever invention which eliminated the bias adjustment. His governor can mathematically be described by the equation

$$u = k(V_r - V) + k_i \int_0^t (V_r(\tau) - V(\tau))d\tau \tag{1.1}$$

the bias term was thus replaced by a term proportional to the integral of past errors. A controller described by (1.1) has the amazing property that the velocity V is always equal to the desired velocity V_r in steady state. This is illustrated by Figure 1.3 which shows the behavior of an engine with control actions proportional to the integral of the error. In standard terminology the Siemens governor is called a PI controller, indicating that the control action is proportional to the error and the integral of the error. Integral action has some amazing properties that will be discussed further in Section 2.2.

The Emergence of Control

Apart from the centrifugal governor the early applications of control include, autopilots for ships and aircrafts, the electronic feedback amplifier and process control. The fundamental similarities between the different systems were not noticed. Control emerged around 1945 as a result of intensive military research in the period 1940-1945. During the Second World War it became apparent that science and technology could have a major impact on the war effort. Among the goals were development of radar and fire control systems. Control was a central feature element of these systems. In many countries groups of scientists and engineers were gathered in research institutes. It was realized that feedback was essential for such diverse systems as autopilots for ships and aircrafts, electronic amplifiers, systems for orientation of radar antennas and guns, and industrial production of uranium. Control was born out of this multidisciplinary effort.

The first manifestation of control was called servo-mechanism theory. This theory used block diagrams as a tool for abstraction. This clearly showed the similarity between the widely different systems. The mathematical tools were Laplace transforms and the theory of complex variables. Analog computers were used for simulation and controllers were implemented as analog computers. It is significant that one of the first books was edited by three persons, a mathematician, a physicist and and engineer from a control company.

One factor that strongly contributed to the emergence of control was that many of the results from the military efforts were disseminated very quickly. After the war it became apparent that control was very useful in practically all branches of engineering and the ideas spread like wild fire around the world. One result was that education in control was introduced as an essential element of engineering education practically all over the world. A characteristic feature of control is that it is not confined to a particular branch of engineering such as mechanical, electrical, chemical, aeronautical, and computer. Control appears in all disciplines, it is in fact the first systems science.

In the late 1950's there were organizational activities which created meeting places for researchers and practitioners of control. Conferences were organized and journals on control started to appear. The International Federation of Automatic Control (IFAC) was the key organization but the traditional engineering organization also introduced special interest groups on control. The first World Congress of IFAC in Moscow in 1960, where control engineers and scientists from all over the world met for the first time, was a landmark. Another effect was the industrialization of control which created companies specializing in control equipment.

The Second Wave

Any field could have been proud of the development that took place from 1940 to 1960, a second wave of development starting in the late 1950s. The driving forces were new challenging applications and a strong stimulation from mathematics. The major drivers for the development were the space race and the use of digital computers for industrial process control. Notice that Sputnik was launched in 1957 and the first computer installed to control an oil refinery started on-line control in 1959. There was a very vigorous development of both theory and practice of control. Digital computers replaced analog computers both for simulation and implementation of controllers. A number of subspecialities of control were also developed. It turned out that a wide range of mathematics fitted the control problems very well, and a tradition of a high respect for mathematical rigor emerged.

Today control is a well established field with a solid body of theory and very wide applications. Practically all controllers are today implemented in digital computers. There are also many challenges. Control is increasingly becoming mission critical which means that great attention has to be given to safety and reliability. The complexity of the control systems are also increasing substantially.

There are tremendous challenges in the future when we can visualize a convergence of control, computing and communication which will result in large interconnected systems. It is also reasonable to guess that a deeper understanding of control in the field of biology will be very exciting. In this sections we will present a number of applications together with some historical notes. The idea is to illustrates the broad applicability of control and the way it has impacted on many different fields.

1.3 Process Control

Many products that are essential for us in daily life are manufactured by the process industries. Typical examples are oil, petrochemical, paper, pharmaceuticals. These industries use continuous manufacturing processes. Process control is an essential part of these industries, the processes can not be run without control systems. We will start by discussing the centrifugal governor.

Advances in Process Control

Key issues in process control are to:

- Keep essential quality variables at specified values.

Figure 1.4 Microprocessor based single loop controller. By courtesy of ABB Industrial Systems.

- Minimize use of energy and raw material.
- Make rapid changes of production or grades.

The control systems are key factors to obtain these goals. There have been major advances since the centrifugal governor appeared.

In the centrifugal governors the actions of sensing, computing and actuation were executed by purely mechanical devices. A lot of ingenuity went into the designs which were often patented. Feedback actually had a crucial role in the design of the controllers. By using feedback it was possible to construct controllers that had a stable well defined behavior from components with a large variability. The technology of controlling engine speed by governors was applied to all types of engines. When electricity emerged in the end of the 19th century there was a similar need to control the speed of generator for hydroelectric generators. Since there was little communication between different fields ideas like integral action were reinvented several times. Major advances were made in the 1930s and 40s. Controllers then appeared as special devices separated from sensors and actuators. New industries , such as Fisher Control, Foxboro, Honeywell, Leeds and Northrup, Taylor Instruments, which supplied sensors, actuators, controller and complete system emerged. The controllers were implemented in different technologies, mechanic, pneumatic and electronic. Controllers for many loops were located in central control rooms. The technology of using digital computers for process control emerged in the 1960s. Computer control is the standard technology for process control.

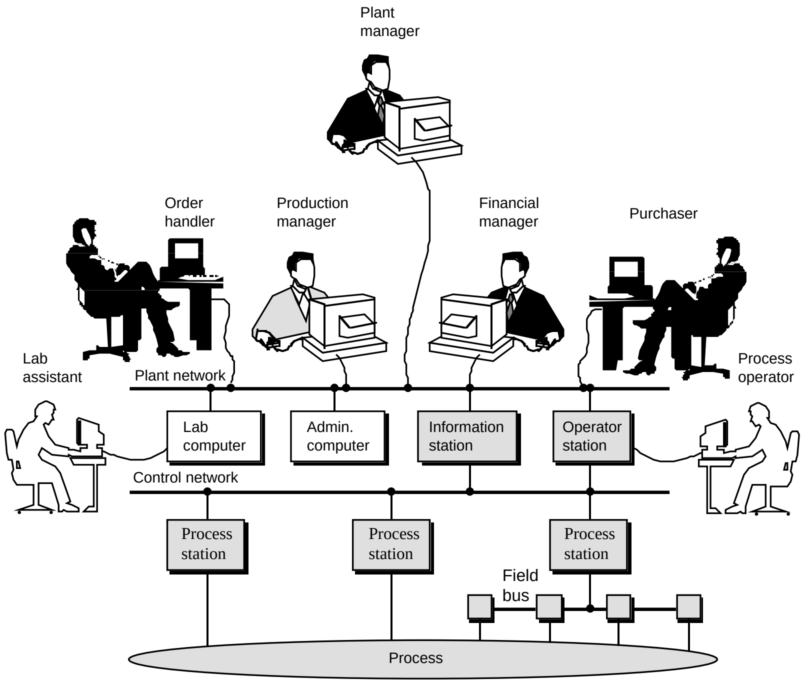

Figure 1.5 Modern industrial systems for process control, like the Advant OCS tie computers together and help create a common uniform computer environment supporting all industrial activities, from input to output, from top to bottom. (By courtesy of ABB Industrial System, Västerås, Sweden.).

New companies have often emerged when major technology changes occurred, there have also been many mergers. Today there are a few large companies that supply control world wide, ABB, Honeywell, Siemens and Toshiba.

The processes vary significantly in scale, a small process unit may have 20-100 control loops but a complete paper mill may have up to several thousand control loops. A wide variety of systems are used for control. There are microprocessor based controllers for single or multiple loops, see Figure 1.4, systems based on personal computers (PC), programmable logic controllers (PLCs) and distributed control systems consisting of many computers connected in a network, see Figure 1.5 Large plants may have 5000-10000 control loops, organized in an hierarchical structure. Most (about 90%) of the lower level control loops are still PID control, which is implemented in the digital computer.

The Role of Sensors

Sensors are key elements of control systems because they provide information about the process. The variables of primary interest in process control are closely related to product quality. Unfortunately it is very difficult to measure such variables on-line. Control is therefore often done indirectly by controlling secondary variables such as level, temperature and pressure. The on-line measurements are often augmented by laboratory analysis of samples of the product.

There are interesting synergies between development of sensors and control. Feedback is often used in sensors. When new sensors become available there are opportunities to develop new control systems. One example is the emergence of sensors for on-line measurement of basis weight and moisture in the 1960s which led to emergence of two new companies Accuray and Measurex which specialized in control of paper machines. There is a similar window of opportunity today because new sensors for measuring composition and surface structure based on infrared and near infrared spectroscopy are available.

1.4 Manufacturing

Process control is continuous manufacturing. Feedback has also had a major impact on manufacturing of discrete parts. Numerically controlled machine tools developed at the Control Systems Laboratory at MIT in the 1950s was a first step where control was used to improve precision of mechanical machining. Welding is highly automated using control and vision systems. Machines for manufacturing systems based on machining with lasers and electrical arcs depend heavily on use of control.

Large manufacturing operations are made on transfer lines, where the parts are moved along a line to stations which perform the operations. The individual stations do operations such as drilling, machining and polishing. A line for making car engines has 10 to 20 machines, which are separated by buffers. Each machine has around 10 stations. The there are simple continuous control systems in each station for tasks such as positioning. A complete transfer line has a few hundred feedback loops. The major control problem is, however, the safe control of the whole operation, for example, to determine when a part should be moved, when a drilling operation should start, when a tool should be removed. This is a discrete control problem where the information comes from on-off signals such as limit switches. The control actions are also discrete, to start a motor, to start drilling, to stop drilling, to change a tool. The systems are very complex because of the large number of signals involved. A single

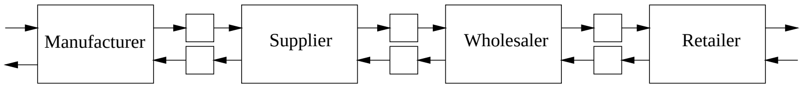

Figure 1.6 Schematic diagram of a simple supply chain consisting of manufacturing, suppliers, wholesaler and retailer, with delays between the different operations.

station in a transfer line may have 5 to 10 000 discrete inputs and output and 10 to 20 continuous control loops. Logic control was originally done by relay systems. When microprocessors were developed in the 1970s the relays were replaced by programmable logic controllers (PLCs). Design of the discrete control system is typically done in an ad hoc manner based on past experience. A general solution is still missing.

Supply Chains

There are also other uses of control in manufacturing, namely control of the complete business operation and complete supply chains. Manufacturing and distribution involve large flows of materials. Raw materials have to be supplied to the manufacturer, the products have to be supplied to the consumers. This is illustrated in Figure 1.6 which shows a simple supply chain. The diagram shows only a few levels, in practice a system may contain thousands of outlets and storages and it may deal with a very large number of products. Several important functions in the system like quality control, sales, and production management are not shown in the figure.

Manufacturing facilities and sales operations have traditionally been quite inflexible in the sense that it is difficult to change products and production rates. In such a system it is necessary to have many buffer storages to match production rates to fluctuating sales and supplies of raw materials and parts from sub-contractors. Production rate is largely determined in an open loop manner based on prediction of sales. An elaborate bookkeeping system is also required to keep inventories of parts.

A different system is obtained if the production time can be reduced and if the production can be made so flexible that products can be changed rapidly. It is then possible to produce a part when an order is obtained. Such a system may be regarded as a closed loop system where production is determined by feedback from sales. An advantage of such a system is that inventories can be vastly reduced. Administration of the system is also greatly simplified because it is no longer necessary to keep track of many products and parts in storage. A system of this type contributed significantly to the efficiency of Walmart. They relied on human sensors in the form of managers who analyzed the sales when the shops were

Figure 1.7 Remote robot surgery using the ZEUS system from Computer Motion Inc. The doctors on the left are in New York and the patient and the robots are in Strasbourg, France. Courtesy of Computer Motion Inc. Goleta

closed every evening to make new orders.

1.5 Robotics

The origin of the industrial robot is a patent application for a device called Programmed Article Transfer submitted in 1956 by the engineer George Devol. The robotics industry was created when Devol met Joseph Engelberger and they founded the company Unimation. The first industrial robot, Unimate, was developed by in 1961. A major breakthrough occurred in 1964 when General Motors ordered 66 machines Unimate robots from Unimation. In 1998 there were about 720 000 robots installed. The majority of them, 412 000 are in Japan. Robots are used for a wide range of tasks: welding, painting, grinding, assembly and transfer of parts in a production line or between production lines. Robots that are used extensively in manufacturing of cars, and electronics. There are emerging applications in the food industry and in packaging. Robots for vacuuming and lawn moving as well as more advanced service robots are also appearing.

Robots are also started to be used used in medical applications. This is illustrated in Figure 1.7 which shows robots that are used for an endoscopic operation. One advantage of the system is that it permits doctors to work much more ergonomically. Another advantage is that operations can be done remotely. The system in the figure is from a robot operation, where the patient was in Strasbourg, France and the doctors in New York.

Another aspect of robots

The word robota is a Slavic word that means work, often meaning of slave work. It came into prominence in a novel from 1918 and a play from 1921 by the check author Karel Capek called *Rossum's Universal Robots*. The play is about robot workers who revolt and kill their human ruler. Another literary aspect of robot is the famous robot trilogy by Isaac Asimov from 1940 who introduced the three laws of robotics

First Law: A robot may not injure a human being, or, through inaction, allow a human being to come to harm.

Second Law: A robot must obey orders given it by human beings, except where such orders would conflict with the First Law.

Third Law: A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Asimovs book captured the imagination of many and there is clear evidence that Engelberger was inspired by Asimovs writing. Other examples about imaginative speculations about robots are found in Arthur Clarkes book *2001 A Space Odyssey* where the robot HAL takes over the operation of a space ship and R2- D2 in the Star Wars series.

There are currently some very interesting developments in robotics. Particularly in Japan there is much research on humanoid and animaloid robots. There are very advanced humanoid robots in research laboratories, there are also robots that mimic snakes, cats, birds and fish. A robot dog AIBO and a service robot have been commercialized by Sony.

Design Issues - Task Based Control and Automomy

Design of robots is a typical multidisciplinary task which is a mixture of mechanical engineering, electronics, computer science, artificial intelligence, and control. It is a typical example of the challenges posed by new industries. Control systems are essential parts of a robot.

The control problems for industrial robots are servo problems, typically to control position of an arm or the force exerted by a tool. There are also other forms of feedback based on force sensors and vision.

Humanoid robots require a much more advanced task based control. It is necessary to provide them with functions for obstacle avoidance, path planning, navigation and map making. Since the robots have to be highly autonomous they must also have capabilities for learning, reasoning and decision making. Development of systems with such capabilities is a challenging research task.

1.6 Power

The power industry started to develop in the end of the 19the century and accelerated rapidly in the 20th century. Availability of power has improved quality of life tremendously. In 2000 the total amount of electric energy generated in the world was about about 15 000 TWh. It is expected to grow by a factor of 3 in 60 years where the main growth is outside the OECD countries.

Control is an essential element in all systems for generation and transmission of electricity. Many central ideas in control were developed in this context. When electricity emerged in the end of the 19th century the generators were typically driven by water turbines. Since alternating current (AC) was the preferred means for transmission there was immediately a need to control the speed of the generators to maintain constant frequency. Derivative and integral control appeared early as did stability critiera. The development paralleled the work on centrifugal governors but it was done independently and it had a stronger engineering flavor. One of the earliest books on control, with the title *Die Regelung der Kraftmaschinen*, was published by Tolle as early as 1905. It was discovered that the performance of hydroelectric power stations was severely limited by dynamics of the water duct. The power decreased rapidly initially when the valve was opened and then increased slowly. This property made the systems difficult to control. It is an example of what is now call a non-minimum phase dynamics.

As the demand for electricity grew many generators were connected in a network. These networks became larger and larger as more generators and consumers were connected. Figure 1.8, which is a schematic picture of the network for the Scandinavian countries, is an example of a network of moderate size. Sweden, Norway and Finland has much hydroelectric power, Sweden and Finland has nuclear power and Denmark has wind and thermal power. In Sweden the hydroelectric power is generated in the north, but most of the consumption is in the south. Power thus has to be transmitted over long lines. The system in the different countries are connected via AC and DC lines. There are also connections to Germany and Poland.

It is difficult to store energy and it is therefore necessary that production and consumption are well balanced. This is a difficult problem because consumption can change rapidly in a way that is difficult to predict. Generators for AC can only deliver power if the generators are synchronized to the voltage variations in the network. This means that the rotors of all generators in a network must line up. Synchronism is lost when the angles deviate too much.

Matching production and consumption is a simple regulation problem

Figure 1.8 The Nordel power grid which supplies electric energy for the Scandinavian countries. The squares represent hydroelectric stations and the triangles represent thermal stations, both nuclear and conventional, and circles denote transformers. The lines denote major power lines. Only the major components are shown in the system.

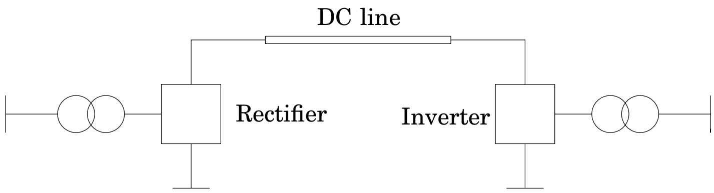

Figure 1.9 Schematic diagram of an HVDC transmission link. The system is fed from the right by AC which is converted to DC by the rectifier and transmitted over a DC line to the inverter which converts it to AC.

for one generator and one consumer, but it is a more difficult problem in a highly distributed system with long distances between consumption and generation. To have a reliable system it is highly desirable to avoid transmission of information over long distances. Control should therefore be done locally at each station based on the information available at the station. Several interesting control principles have been developed to do this. Control of each generator must be based on information that is locally available. Because of reliability requirements it is not possible to rely on information that is transmitted over wide distances.

High Voltage DC Transmission Systems

Even if most of the electricity is transmitted as AC there are some situation where it is advantageous to transmit direct current (DC). One example is when electric power has to be transmitted using under water cables. DC links can also be used to connect two asynchronous power grids. The systems for transmission of high voltage DC (HVDC systems) have many interesting properties. A schematic picture of such a system is shown in Figure 1.9. The AC is rectified to generate DC which is transmitted and converted to AC by the inverter. The rectifier and the inverter consists of semiconductor switches that permit very rapid changes of the power. It is possible to switch the direction of 600MW in fractions of a second. Control of such a system in a safe precise way is a great challenge.

Control of Networked Power Systems

An interconnected power system of the type shown in Figure 1.8 is a complicated system. The behavior of such systems is not easy to predict. Problems when interconnecting two systems were encountered by Edison. He found that it was not possible to connect two turbine driven generators when both generators had controllers with integral action. The system will drift so that one generator takes all the load. This was one of the

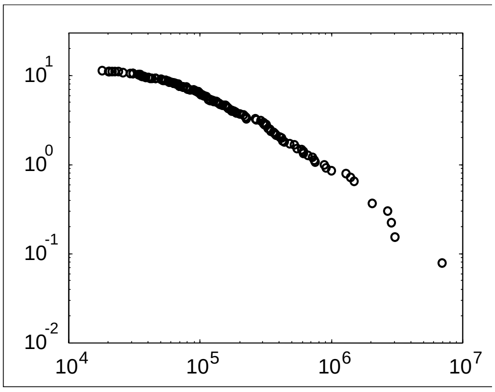

Figure 1.10 Power outages in the US 1984-97. The horizontal axis shows the number of persons N affected by the outages and the vertical axis shows the yearly frequency of outages that influence more than N persons. Notice that the scales on both axes are logarithmic.

first observations that problems may occur when several regulators are connected to an integrated system.

Edisons observation led to interesting developments of control theory. In current practice one large generator in the network controls the frequency using a controller with integral action. The other generators use proportional control. The amount of power delivered by each generator is set by the gain of the proportional controller. Each generator has separate voltage control.

There have been many other surprises in interconnected systems. In the Nordel system it has been observed that a moderated increase of power load in the north could result in large oscillations in the power transmission between Sweden and Denmark in the south. Oscillations have been observed when modern trains with switched power electronics have put in operation. An understanding of such phenomena and solutions require knowledge about dynamics and control.

Safety and Reliability

The power systems are generally very reliable. Customers will have power even when generators and lines fail. This is achieved by good engineering of the system based on redundancies. Networked generators contribute significantly to the reliability of the system because it is possible for a large number of generators to take up the load if one generator fails. The drawback is however that there may be massive failures in the system which also has occurred. This is illustrated in Figure 1.10 which shows the statistics of power failures.

1.7 Aeronautics

Control has often emerged jointly with new technology. It has often been an enabler but in some cases it has had a much more profound impact. This has been the case in aeronautics and astronautics as will be discussed in this section.

Emergence of Flight

The fact that the ideas of control has contribute to development of new technology is very nicely illustrated by the following quote from a lecture by Wilbur Wright to the Western Society of Engineers in 1901:

> “ 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 compromise between stability and maneuverability by building an airplane, Kitty Hawk, that was unstable but maneuverable. The pioneering flight was in 1905. 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. The fact that this plane was unstable was a strong impetus for the development of autopilots based on feedback.

Autopilots

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

Figure 1.11 Picture from Sperry's contest in Paris. Sperry's son is at the stick and the mechanic walks on the wings to introduce disturbances. Notice the proximity to the ground.

autopilot is the first use of feedback in aeronautical engineering. Sperry won a prize in a competition for the safest airplane in Paris in 1912. Figure 1.11 is a picture from the event. The autopilot is a good example of how feedback can be used to stabilize an unstable system.

Autonomous Systems

Fully automatic flight including take off and landing is a development that naturally follows autopilots. It is quite surprising that this was done as early as 1947, see Figure 1.12. The flight was manually supervised but the complete flight was done without manual interaction.

Autonomous flight is a challenging problem because it requires automatic handling of a wide variety of tasks, landing, flying to the normal flight altitude, navigation, approaching the airfield, and landing. This requires a combination of continuous control and logic, so called hybrid control. In the flight by Robert E. Lee the logic required was provided by an IBM computer with punch cards. The theory of hybrid systems is still in its infancy. There are many similarities with the problem of designing humanoid robots. They require facilities for navigation, map making, adaptation, learning, reasoning. There are many unsolved research prob-

Robot-Piloted Plane Makes Safe Crossing of Atlantic

No Hand on Controls From Newfoundland to Oxfordshire—Take-Off, Flight and Landing Are Fully Automatic

By ANTHONY LEVIERO

Special to THE NEW YORK TIMES.

WASHINGTON, Sept. 22—A Douglas C-54 Skymaster with a mechanical brain landed without human aid near London today after a robot directed hop from Newfoundland.

The revolutionary flight across the Atlantic, effected by the push of the button, was hailed by Air Force spokesmen said. Two ships somewhere in the Atlantic furnished bearings to the Skymaster's brain. She had 3,700 gallons of fuel aboard.

Great Britain several weeks ago had asked as a favor that the Air Force send the Skymaster there to make demonstration flights for

Figure 1.12 Excerpt from article in New York Times on September 23, 1947, describing the first fully automatic transatlantic flight.

lems in this field.

The use of autonomous systems is rapidly increasing. The Boeing 777 and the Airbus are modern examples of systems with a high degree of autonomy. There have also been extensive developments of Unmanned Air Vehicles (UAV).

Integrated Process and Control Design

Flight control is a good illustration of the value of integrated process and control design. The Wright brothers succeeded where others failed, because they made an unstable airplane that was maneuverable. Aircrafts that were both stable and maneuverable were built later. There are still substantial advantages in having unstable aircrafts that rely on a control system for stabilization. Modern fighters obtain their performance in this way. A schematic picture of two modern jet fighters are shown in Figure 1.13. The positions of the center of mass CM and the center of pressure are key elements. To be stable the center of pressure must be behind of the center of mass. The center of pressure of an aircraft shifts backwards when a plane goes supersonic. If the plane is stable at sub-sonic speeds it becomes even more stable at supersonic speeds. Very large forces and large control surfaces are then required to maneuver the airplane. A more balanced design is obtained by placing the center of pressure in front of

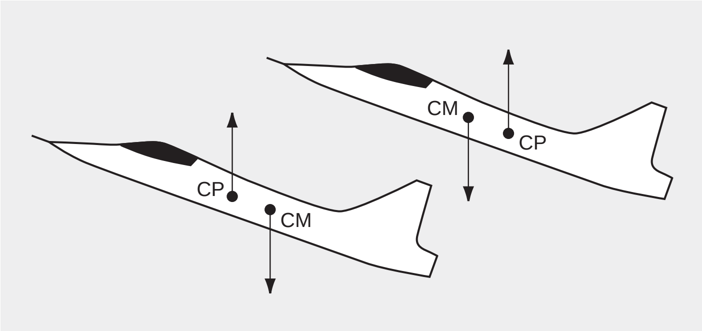

Figure 1.13 Schematic digram of two aircrafts. The aircraft above is stable because it has the center of pressure behind the center of mass. The aircraft below is unstable.

the center of mass at sub-sonic speeds. Such a plane will however by unstable at sub-sonic speeds, i.e. at take off and landing. This imposes severe constraints on the safety and robustness of the control system, but the aircraft will have superior performance. When the automatic control system becomes a critical part of the process it may also become mission critical which means that the system will fail if the controls fails. This induces strong demands on the reliability of the control system.

The development of aeronautical and aerospace engineering has often gone hand in hand with the development of feedback control. It was realized that control has to be considered up front in the design at the same time as, structures, engines, aerodynamics. A very interesting illustration of this is the recent development of high performance military aircrafts. Most aircrafts built today are designed to be stable.

Control is also mission critical for rockets and satellites.

1.8 Electronics and Communication

Electronics emerged in the in 1906 with the invention of the audion, a prototype of the vacuum tube, by Le De Forest. This was the start of the revolutionary development of the electronics industry which has had a big impact on the way we live. Control has had a major impact on this industry as well. The first application of feedback in electronics was a patent on vacuum tube amplifiers by the rocket pioneer Robert Goddard in 1912, but the most influential development is undoubtedly the negative feedback amplifier. Before dealing with this we will, however, briefly

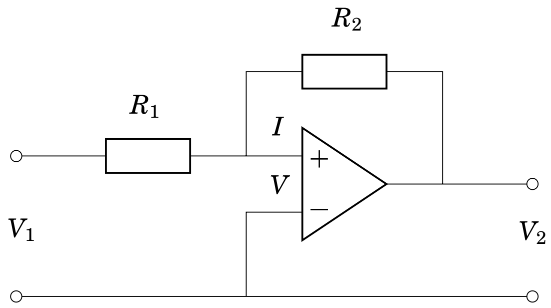

Figure 1.14 Schematic diagram of Armstrongs super-regenerative receiver.

discuss an application of positive feedback.

The Super-regenerative Amplifier - Positive Feedback

Vacuum tubes were expensive and it was highly desirable to use as few amplifiers as possible. In 1915 Armstrong suggested to use positive feedback to obtain a high amplification. A schematic of Armstrongs amplifier is shown in Figure 1.14.

Assume that the current I into the amplifier is zero, then the current through the resistors R_1 and R_2 are the same. It then follows from Ohms Law that

$$\frac{V_1 - V}{R_1} = \frac{V - V_2}{R_2} \tag{1.2}$$

Let G be the open loop gain of the amplifier, it then follows that

$$V_2 = GV \tag{1.3}$$

Eliminating the variable V between Equations (1.2) and (1.3) gives the following equation for the ratio of the output and input voltages

$$\frac{V_2}{V_1} = G\frac{R_1}{R_1 + R_2 - GR_2} \tag{1.4}$$

This equation gives the gain of the amplifier with positive feedback. The gain is very large if the resistors are chosen so that $R_1 + R_2 - GR_2$ is small. Assume for example that $R_1 = 100k\Omega$, $R_1 = 24k\Omega$, and $G = 5$. The formula above shows that the gain of the feedback system is 25. With $R_2 = 24.5k\Omega$ the gain will be 50, and with $R_1 = 25k\Omega$ the gain is infinite.

Regeneration was the word used for feedback at the time, and Armstrongs amplifier was called a super-regenerative amplifier because it obtained high gain by positive feedback. Armstrongs invention made it

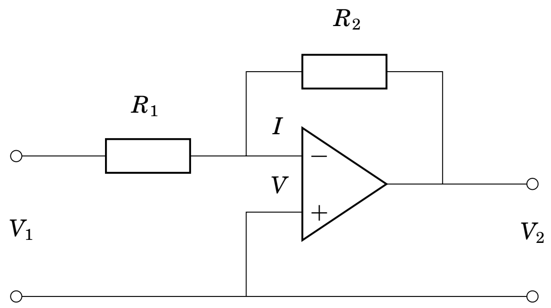

Figure 1.15 An amplifier with negative feedback.

possible to build inexpensive amplifiers with very high gain. The amplifiers were, however, extremely sensitive and they could easily start to oscillate. Prices of vacuum tubes also dropped and the interest in the amplifier dwindled. Next we will discuss another use of feedback in an amplifier which still has profound consequences.

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 Black's invention of the feedback amplifier in 1927. Black used negative feedback which reduces the gain but makes the amplifier very insensitive to variations in tube characteristics. Black's invention made it possible to build stable amplifiers with linear characteristics despite nonlinearities of the vacuum tube amplifier.

A schematic diagram of a feedback amplifier is shown in Figure 1.15. Assume that the current I into the amplifier is zero, the current through the resistors R_1 and R_2 are then the same and it follows from Ohms Law that Equation (1.2) holds. Let the gain of the amplifier be G it follows that

$$V_2 = -GV \tag{1.5}$$

Eliminating the variable V between Equations (1.2) and (1.5) gives the following equation for the ratio of the output and input voltages

$$\frac{V_2}{V_1} = -\frac{R_2}{R_1}\frac{1}{1 + \frac{1}{G}(1 + \frac{R_2}{R_1})} \tag{1.6}$$

This equation gives the gain of the amplifier with negative feedback. Since the gain G is a very large number, typically of the order of 10^5 or 10^8, it follows from this equation that the input-output property of the amplifier is essentially determined by resistors R_1 and R_2. These are passive components which are very stable. The properties of the active components appear in parameter G. Even if G changes significantly, the input-output gain remains constant. Also notice that the relation between V_{out} and V_{in} is very close to linear even if the relation between V_{out} and V, equation 1.5, is strongly nonlinear. □

Like many clever ideas the idea of the feedback amplifier seems almost trivial when it is described. It took, however, six years of hard work for Black to come up with it. The invention and development of the feedback amplifier was a key step in the development of long distance communications. The following quote from the presentation of the IEEE Lamme Medal to Black in 1957 gives a perspective on the importance of Black's invention of the feedback amplifier:

> It is no exaggeration to say that without Black's invention, the present long-distance telephone and television networks which cover our entire country and the transoceanic telephone cables would not exist. The application of Black's principle of negative feedback has not been limited to telecommunications. Many of the industrial and military amplifiers would not be possible except for its use. ... Thus, the entire explosive extension of the area of control, both electrical and mechanical, grew out of an understanding of the feedback principle. The principle also sheds light in psychology and physiology on the nature of the mechanisms that control the operation of animals, including humans, that is, on how the brain and senses operate.

It is interesting to observe that while Armstrong used positive feedback Black was using negative feedback.

Feedback quickly became an indispensable companion of electronics and communication. The applications are abundant. Today we find interesting use of feedback in power control in system for cellular telephony. A handset must naturally use enough power to transmit so that it can be heard by the nearest station, using too much power will increase interference with other handsets, necessitating use of even more power. Keeping the power at the correct level gives a large pay-off because the batteries will last longer.

Hewlett's Stabilized Oscillator

Many control problems are solved by linear systems. There are, however, problems where nonlinearities are essential. One of them is the design of

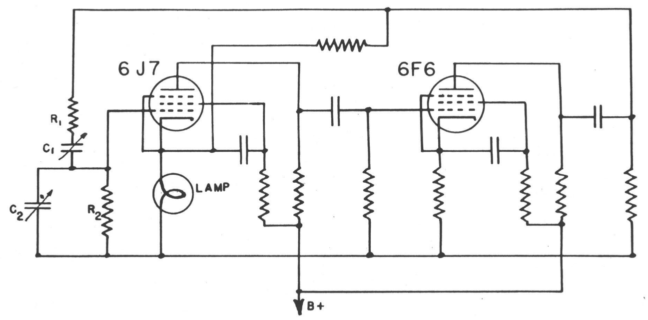

Figure 1.16 Circuit diagram of William Hewletts oscillator that gives a stable oscillation through nonlinear feedback using a lamp.

an oscillator that gives a signal with a constant amplitude. This problem was solved very elegantly by William Hewlett in his PhD thesis at Stanford University in 1939. Hewlett simply introduced a nonlinear element in the form of a lamp in the circuit, see Figure 1.16. Hewletts oscillator was the beginning of a very successful company HP, that Hewlett founded with David Packard.

CD Players and Optical Memories

The CD player is an interesting device which critically depends on several high performance control loops. The yearly sales of CD players was about 100 million units in the year 2000. That makes it one of the most common control systems.

A schematic picture of the main components of a CD player is shown in Figure 1.17. The information on the disc is read by light from a laser diode that is reflected by the disc to an optical assembly with photo transistors. Processing of the signals from the transistors gives the radial track error and a a focus signal. The laser diode, the optical assembly and the actuator are mounted on a sled which can be moved precisely over a small range. The sled is moved by another servo which permits large motions.

A block diagram of the major servos are shown in Figure 1.18. There are three critical servo loops. The focus servo concentrates the laser spot in the disc information layer. The tracking servo positions the laser spot on the track. The sled servo moves the sled so that the tracking system is in operating range. The tracking and the sled servos are also used to switch between tracks. The servos are all based on error feedback since only the error signal is available from the sensors. The major disturbance is due to misalignment of the track. In a CD player this is due to an

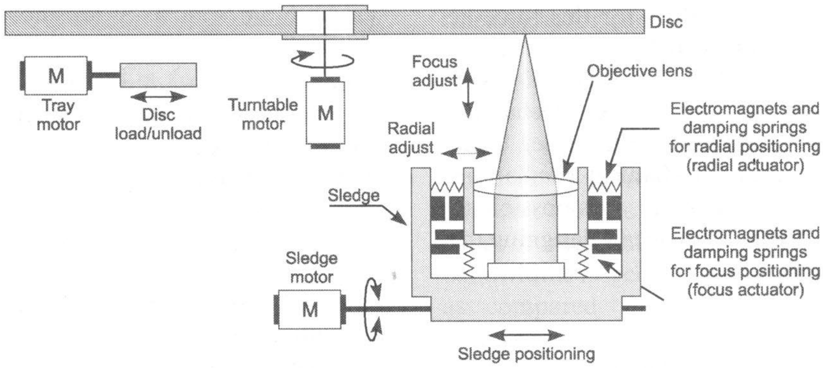

Figure 1.17 Schematic picture of a CD player.

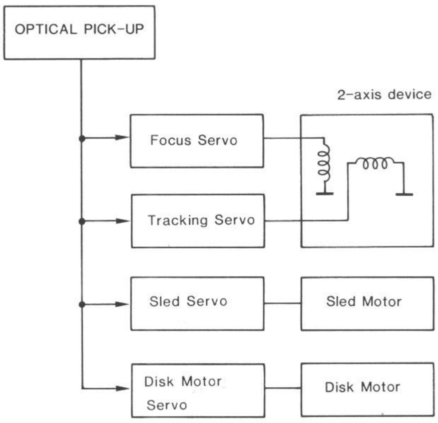

Figure 1.18 Block diagram of the major servos in a CD player.

off set of the center both due to manufacturing variations and errors in centering of the CD. The disturbance is approximately sinusoidal. The tracking accuracy of the systems is quite remarkable. For a typical CD player, which stores 650 M-bytes, the variations in a track are typically $200\,\mu m$, the track width is $1.6\,\mu m$ and the tracking accuracy is $0.1\mu m$. The tracking speed varies from 1.2 to $1.4 m/s$. The servos used in the Digital Versatile Disc (DVD) and in optical memories are very similar to the servos in a CD but the precision is higher. For a Digital Versatile Disc (DVD) the variations in a track are typically $100\,\mu m$, the track width is

1.6 μm and the tracking accuracy is $0.022 \mu m$. The tracking speed varies in the range $3.5 m/s$. The quality of the major servo loops have a direct impact on the performanc of the storage system. A better tracking servo permits a higher storage density in an optical drive. An improved servo for switching tracks is immediately reflected in the search time in an optical memory.

1.9 Automotive

Cars are increasingly being provided with more and more control systems. The possibility of doing this are due to availability of cheap microprocessors and sensors and good control technology. The automotive industry has also been a strong driving force for the development of micro-controllers and sensors.

Reducing Emissions

California introduced a standard that required a substantial reduction of emissions for internal combustion engines. To achieve this it was necessary to introduce feedback in the engine based on measurement of the oxygen in the exhaust. The following quote from a plenary lecture by William E. Powers a former vice president of Ford at the 1999 World Congress of IFAC is illuminating.

> The automobiles of the 1990s are at least 10 times cleaner and twice as fuel efficient as the vehicles of the 1970s. These advancements were due in large part to *distributed microprocessor-based control systems*. Furthermore the resultant vehicles are safer, more comfortable and more maneuverable.

Cruise and Traction Control

Most new cars are provided with a system for cruise control that keeps the speed constant. This is a common control system that practically everyone comes in contact with. More advanced systems, called adaptive cruise control, are now appearing. The system is called adaptive even if it is a regular servo system that keeps the distance to the car in front at a constant value. The distance is measured by radar. In this context it is interesting to note that in 1995 Dickmann's modified a Mercedes to make it fully autonomous. The system had vision sensors and could make automatic lane changes. The car has been tested with human supervision at high speed in Germany and France.

Systems for anti-lock braking and traction control have also been introduced. These systems were used in an unexpected fashion in the Mercedes

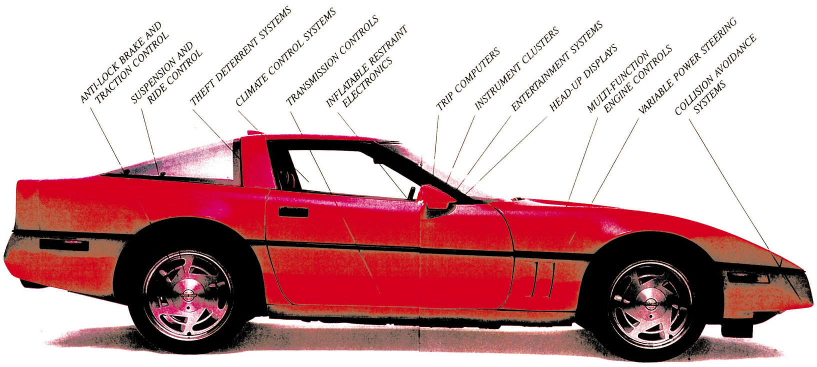

Figure 1.19 Some of the control systems in a car.

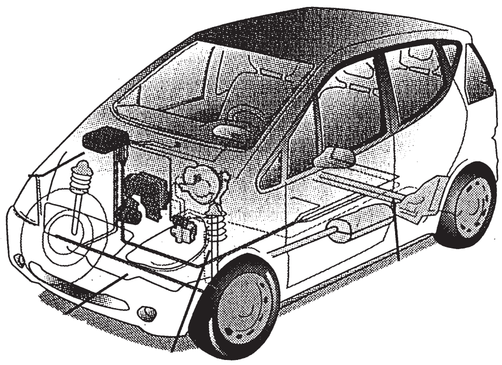

Figure 1.20 The Mercedes A-class is a small car where control helped to solve a serious problem.

A class, see Figure 1.20. This is a small car which achieves a high degree of safety through a thick floor that can be deformed in case of an accident. A consequence of this is that the center of gravity is high. When the car was introduced it was discovered that the car had a tendency to fall over in sharp turns. This difficult problem was solved by providing the car with the most sophisticated traction control system available in the company at the time. Together with minor changes of tires a severe difficulty was overcome.

Technology Drivers

The automotive industry is an important driver for technology because of the large number of produced parts and hard requirements for low cost. Several interesting developments took place when computers started to be used for computer control of engines. To save costs the microcomputer and input-output devices that connect sensors and actuators were merged on one chip, so called micro controllers. These devices made computer control cost effective in many other fields. The total number of micro-controllers for embedded systems now far exceed the number of microprocessors manufactured each year. The automotive applications also required new sensors and actuators. New accelerometers and gyros based on MEMS devices were developed, electric actuators based on new magnetic materials have also been developed. The sensors and actuators use feedback internally to obtain robustness and performance.

Autonomous Driving

There have been several attempts at developing autonomous vehicles. In 1995 Dickmanns demonstrated a fully autonomous Mercedes Benz with vision sensors. Under human supervision the car drove autonomously on the Autobahn from Munich to Copenhagen. Experiments with autonomous driving have also been done in California in the Path program.

1.10 Computing

There has been a strong symbiosis between control and computing. Computing devices are integral parts of a controller and computing and simulation are used extensively in design and validation of a control system.

Analog Computing

Early controllers, such as the centrifugal governor, were implemented using mechanical devices. Integral action was implemented using the ball and disc integrator invented by Lord Kelvin. In the process industries analog computing was instead done using pneumatic devices. The key elements were pneumatic amplifiers, restrictions and volumes. Feedback was used extensively to obtain linear behavior from the nonlinear devices.

The early development of control was severely hampered by the lack of computing, many clever graphical methods were developed to obtain insight and understanding using modest computing. The situation is summarized very clearly in the following quote from Vannevar Bush from 1923.

> "Engineering can proceed no faster than the mathematical analysis on which it is based. Formal mathematics is frequently inadequate for numerous problems pressing for solution, and in the absence of radically new mathematics, a mechanical solution offers the most promising and powerful attack wherever a solution in graphical form is adequate for the purpose. This is usually the case in engineering problems."

Bush later built the the first mechanical differential analyzer. The key elements were the ball and disc integrator and the torque amplifier. It could be used to integrate a handful of differential equations. This computer was used by Ziegler and Nichols to devise tuning rules for PID controllers.

Bush's work laid the foundation for analog computing which developed rapidly when cheap electronic amplifiers became available. This coincided with the emergence of control and analog computing became the standard tool for simulation of control systems. The analog computers were however large expensive systems that required a large staff to maintain. The use of analog computing was thus limited to persons with access to these rare resources. Analog computing was also used to implement the controllers in the servomechanism era and through the 1960s. Controllers were also implemented as small dedicated analog computers.

Even if computer control is the dominating technology for implementing controllers there are still niches where analog computing is used extensively. One area is micro-mechanical systems (MEMS) where mechanics and control are integrated on the same chip. Analog computing is also used for systems with extremely fast response time.

Computer Control

When digital computers became available they were first used for computing and simulation. The early computers were large and expensive and not suitable to be embedded in controllers. The first computer controlled systems was installed by TRW at the Port Arthur refinery in Texas in 1959. This initiated a development which started slowly and accelerated rapidly with the advances in computing. Today practically all controllers are implemented as computer controlled systems.

Real Time Computing

Use of computers for in control systems imposes demands on the architecture of systems and software, because of the requirement of fast response to external events. It turns out that the features that were useful for control were also useful for other purposes. There are in particular severe requirements on the operating system to provide rapid response to external events. It is also necessary to make sure that the system operates without interruptions. Special real time operating systems therefore emerged. When implementing control systems there is a clear need to understand both control algorithms and software. It is also necessary to have systems with a high degree of reliability. The luxury of restarting the computer (CTRL+ALT+DEL) if something strange happens is not a feasible solution for computer control.

Simulation

Simulation is an indispensable tool for the control engineer. Even if systems can be designed based on relatively simple models it is essential to verify that the system works in a wide range of operations. This is typically done by simulating the closed loop system. To be reliable the simulation requires a high fidelity model of the system, sometimes parts of the real system is actually interfaced with the simulator, so called hardware in the loop simulation. If a controller is build using a dedicated computer it can also be verified against the simulation. Simulation can also be used for many other purposes, to explore different systems configurations, for operator training and for diagnostics. Because of the advances in computers and software simulation is now easily available at every engineers desk top. Development of a suitable model for the process requires a major effort.

The Internet

The Internet was designed to be an extremely robust communication network. It achieves robustness by being distributed and by using feedback. The key function of the system is to transmit messages from a sender to a receiver. The system has a large number of nodes connected with links. At each node there are routers that receives messages and sends them out to links. The routers have buffers that can store messages. It is desirable to operate the system to exploit capacity by maximizing throughput subject to the constraint that all users are treated fairly. There are large variations in traffic and in the lengths of the messages. Routers and links can also fail so the system may also be changing. The Internet depends critically on feedback to deal with uncertainty and variations. All control is decentralized

The routers receive messages on the incoming lines and distributes them to the outgoing lines. Messages are stored in a buffer if the router can not handle the incoming traffic. In the simplest scheme a router that has a full buffer will simply drop the incoming messages. Information about congestion is propagated to the sender indirectly through the lost messages.

Flow control is done by the senders. When a message is received by a receiver it sends an acknowledgment to the sender. The sender can detect lost messages because the messages are tagged. A very simple algorithm for traffic control was proposed by Jacobson 1990. This algorithm is called Additive Increase Multiplicative Decrease (AIMD) works as follows. The transmission rate is increased by additively as long as no messages are lost. When a message is lost the transmission rate is reduced by a factor of 2.

The Internet is a nice illustration that a very large distributed system can be controlled effectively by reasonably simple control schemes. There are many variations of the basic scheme described above. Many technical details have also been omitted. One drawback with the current scheme is that the control mechanism creates large variations in traffic which is undesirable. There are many proposals for modifications of the system.

1.11 Mathematics

There has always been a strong interplay between mathematics and control. This has undoubtedly contributed to the success of control. The theory of governors were developed by James Clarke Maxwell in a paper from 1868, about 100 years after James Watt had developed the governor. Maxwell realized that stability was connected to the algebraic problem of determining if an algebraic equation has roots in the right half plane. Maxwell turned to the mathematician Routh to get help. An analogous situation occurred in the development of water turbines where Stodola turned to the mathematician Hurwitz for assistance.

Mathematics played a major role in the development of servomechanism theory in the 1940s. There were a number of outstanding mathematicians at the Radiation Laboratory at MIT, there were also outstanding mathematicians at Bell Laboratories at the time. An interesting perspective can be obtained by comparing the major advances in the theory of feedback amplifiers with the meager advances in process control and to speculate what could have happened if there had been mathematicians working on the process control problems.

There was a very fruitful interaction between mathematics and control in the 1960s in connection with the space race when optimal control was developed. The Russian mathematician Pontryagin and his coworkers developed the maximum principle, which can be viewed as an extension of the Euler-Lagrange theory in calculus of variations. The American mathematician Richard Bellman developed dynamic programming, which can be regarded as an extension of the Hamilton-Jacobi theory in calculus of variations. In both cases the developments were directly inspired by the control problems.

In the US the eminent topologist and past President of the American Mathematical Society argued for a strong effort in applied mathematics. With support form the Office of Naval Research he established a research center devoted to nonlinear ordinary differential equations and dynamics at Princeton. The center was later moved to RIAS and later to Brown University where it became the Lefschetz Center for Dynamical Systems. Later centers were also created at the University of Minnesota.

The Russian activity was centered in Moscow at the famous Institute of Automatic Control and Telemechanics and in smaller institutes in St Petersburg and Sverdlovsk. In Peking a strong center was established under the supervision of the Academy of Sciences. A very strong center INRIA was also created in Paris under Professor Jean Jacques Lions.

A wide range of mathematics such as dynamical systems, differential geometry and algebra has been important for the development of control theory after 1960.

Numerical Mathematics

A standard computational problem is to solve ordinary differential equations by time-stepping methods. As solutions often vary significantly over the range of integration, efficient computation requires step length control. This is done by estimating the local error and adjusting the step length to make this error sufficiently small. Most solvers for differential equations have carried out the correction in a simplistic fashion, and the adjustment has often been mixed with other algorithmic elements.

Recently a drastic improvement has been made by viewing step length adjustment as a feedback problem. Substantial improvements of performance can be obtained by replacing the heuristic schemes for step length adjustment, that were used traditionally, with a scheme based on a PID controller. These advantages are achieved without incurring additional computational costs. This has resulted in more reliable software, as well as software with much better structure. Knowledge of basic control schemes has thus proven very beneficial. It is likely that the same idea can be applied to other numerical problems.

1.12 Physics

Feedback has always had a central role in scientific instruments. An early example is the development of the mass spectrometer. 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 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 the particles W and Z was first demonstrated. Another use of feedback called repetitive control was developed by Nakano for particle accelerators. The key idea was to obtain very precise control by exploiting the fact the particles move in circular orbits.

The atomic force microscope is a more recent example. The key idea is to move a narrow tip on a cantilever beam across the surface and to register the forces on the tip. Such systems rely on feedback systems for precise motion and precise measurement of the forces.

Adaptive Optics

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. This is the idea of adaptive optics. A schematic picture of a system for adaptive optics is shown in Figure 1.21. The reference signal is a bright star or an artificial laser beam projected into the atmosphere. The actuator which is shown simply as a box in the figure is obtained by reflecting the light on a mirror that can be deformed selectively. The mirror can have from 13 to 1000 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.

In normal feedback terminology adaptive optics is a regular feedback system, feedback is used to compensate for variations in diffraction in the atmosphere. The word adaptive is used in because it can also be said that the system adapts for variations in the atmosphere. In the control community the word adaptive is often used with a different meaning.

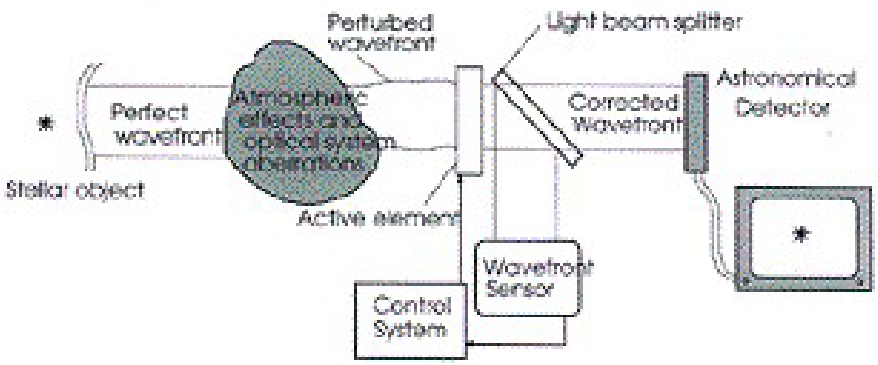

Figure 1.21 Schematic diagram of a system for adaptive optics.

Quantum Systems

Control of quantum systems is currently receiving a lot of interest. Molecular dynamics is a very spectacular application. The idea is to use modulated laser light to break up bonds in molecules to obtain ions which can react with other ions to form new molecules. This is done by tailoring the laser pulses so that they will break specific bonds between atoms. This is precision surgery at the molecular level, quite different from the methods used in conventional chemistry.

1.13 Biology

It was mentioned already in Section 1.1 that feedback is an essential mechanism in biology. Here are a few examples.

The Pupillary Reflex

The human eye has an effective system to control the amount of light that is let into the eye. The light intensity is measured and the pupil opening is adjusted. This control systems is easily accessible for experimentation. Extensive investigations have been made by changing the light intensity and measuring the pupil opening. The results show that it is a very effective feedback system.

Human Posture

The human body hasmany feedback systems. They allow us to stand upright, to walk, jump and balance on ropes. They also adjust the sensitivity of our eyes and ears, enabling us to see and hear over a wide range of intensity levels. They maintain a constant body temperature and a delicate balance of chemical substances in our body. As an illustration we will discuss the system that allows us to stand upright. The key features of the system are known although several details are poorly understood. The primary sensors are the semicircular canals located in the mastoid bone close to the ear. The sensors consist of toroidal canals that are filled with liquid. Neurons connected to hairs in the canals give signals related to the motion of the head. The major actuators are muscles in feet, legs, knees, hips and arms. There are also sensory neurons in the feet and the muscles. There is local feedback from pressure sensors in the feet and sensors in the muscles to the actuating muscles. This loop has a reaction time of about 20 ms. The interconnection between sensors and actuators is made in the spinal cord. These interconnections are responsible for fast feedback like reflexes. The reaction time is of the order of 100 ms. There is also a high level feedback loop that receives information from the vestibular system, which gives information about the position and orientation of our body parts in space. The sensory information is processed in the cerebellum and transmitted to the muscle neurons in the spinal cord. This feedback has a reaction time of about 250 ms.

This system for control of posture illustrates that feedback can be used to stabilize an unstable system. It also shows that there are very reliable biological feedback systems which are essential to everyday life. A particularly interesting feature is that the system has learning capabilities. Think about a child learning to stand up and walk or learning to bicycle. These functions are far superior to those of any technical system.

A Simple Experiment

A simple experiment on one of the systems in the body can be executed manually with very modest equipment. Take a book with text and hold it in front of you. Move the text sideways back and forth and increase the speed of motion until the text is blurred. Next hold the text in front of you and move the head instead. Notice the difference in the speeds when the text gets blurred. You will observe that higher speeds are possible when you move your head. The reason for this is that when you move the text, the information about the motion comes via the processing of the image at your retina, but when you move your head the information comes from the semicircular canals. The feedback from the visual processing is much slower because it uses higher functions in the brain.

There are many other nice control systems in the human body. Top performing athletes such as tennis players have interesting abilities for very advanced motion control that involves much interaction with vision. Humans also have interesting learning abilities.

1.14 Summary

This chapter has given glimpses of how control and feedback have influenced the development of technology and how control is used. The examples show that control has emerged concurrently with new technology, that it has had a major influence on technology and that it sometimes has been an enabler. A large number of control systems ranging from small micro devices to large global systems for generation and distribution of electricity and large communication systems have also been given. The wide range of uses of control also point to some difficulties. The ideas of control and feedback are abstract which makes them less obvious. It is therefore common that the ideas are neglected in favor of hardware which is much easier to talk about.