# Chapter 9
# Sensing and Actuation

The terms sensing and actuation are used to refer to getting information about the world and to affecting physical objects, respectively. In cyber-physical systems, an interesting aspect of exploring sensing and actuation is that it provides us with a natural opportunity to learn more about how computational components work today, and in particular, which ones can be realized directly using semiconductor based circuits, and which ones require other intermediate steps to realize.

## 9.1 Everyday Input and Output

Sensing and actuation occur in simple forms in many computational systems such as a desktop, laptop, or any device that we traditionally think of as a digital computer. Electronically, the simplest way to get an input into a computer is through a switch, such as the home button on a smartphone or a particular key on a keyboard. A switch is a device that either allows or blocks a current depending on external input, such as the physical position of a lever or button. A switch can, in principle, be used directly to send a binary signal into a digital circuit. In principle, a keyboard button or an OFF/ON switch can use such simple circuitry. In practice, and for a variety of reasons, additional circuits may be useful to improve the quality of the signal and to ensure the safety of various subsystems. The point is, having a device that provides input to a computational system, in its simplest form, can be quite straightforward. Maybe more importantly, every signal that we sense will need to go through such a step to enter into the computational system.

Once this signal has reached the digital circuit, it can be processed in one of two ways. Either the signal will be held by a latch so that changes in its value can only be observed at a clock tick or it can be read by a signal that will use its value directly.

The question now is how do we turn a digital output into a physical action of some sort. As it turns out, one of the simplest ways to observe the output physically is to use one of the technologies that is most widely used today, namely, the Light

W. M. Taha et al., *Cyber-Physical Systems: A Model-Based Approach*,
https://doi.org/10.1007/978-3-030-36071-9_9

Emitting Diode (LED). For many microprocessors, all what would be required would be to connect an LED followed (in series) with a small resistor to the wire carrying the digital signal that we wish to observe. Then, when there is a high (voltage) signal on the wire, the LED turns on, and it turns off when the signal is low. As an aside, it should be noted that some microprocessors use HIGH to represent 0, whereas others use it to represent 1.

## 9.2 Symmetry: LEDs and Photo-Voltaic Cells

It would have been elegant if we could somehow observe such an output by a change in the position of a button. That would give us a nice symmetry that seems natural when we are using actuation and sensing primarily for communication. Alas, whereas an OFF/ON switch may be the easiest way to get an input into a computational component, building the mechanism needed to get a switch to move is, relatively speaking, non-trivial. LEDs, on the other hand, do give us an example of some of the simplest ways to provide an input and to observe an output from a digital system: An LED itself is photo-sensitive, and can therefore be used for sensing light as well as emitting it. Isolated in a lit environment, an LED will have a voltage across two connectors. This voltage can then be amplified to detect the presence or absence of light.

The fact that LEDs can be used for both sensing and actuation makes them particularly interesting devices, as they are highly flexible as input, output, and communication devices. At some point, certain models of one of the predecessors of the smart phone, the Personal Digital Assistant (PDA) supported communication between such devices via infrared. Remote controls for many home appliances have for a long time used infrared light for one-direction communication. More recently, there has been growing discussion of light fidelity (LiFi, in analogy to WiFi) as a communication medium. Fiber optics are currently one of the highest bandwidth mechanisms for communicating between computer systems and over long distances. In addition, fiber optic buses are used in high performance computing systems to connect CPU cores. The ease with which light can be generated and processed by semiconductor devices makes it possible, in principle, that future CPUs may use light within chips. Media reports in 2019 included news that Intel is working on a chip with optical interconnects for Neural Network applications (See To Probe Further).

Another interesting aspect of LEDs is that the presence of light creates a voltage potential that can be used to harvest energy. This is in fact what photo-voltaic cells do—they can be viewed as a minor variation on the LED. A challenge with photo-voltaic cells is that they must be grouped and connected electrically with care, so as to support higher aggregate voltages as well as to enable buildup of higher currents. In addition, being significantly exposed to an open environment, they must be packaged in a way that allows them to operate effectively over a long period of time without degradation in efficiency. In light of what we have just learned, the attentive reader will understand that design of LED lighting is challenging for similar reasons.

### 9.2.1 Diodes

To understand why connecting computational and physical components is interesting, it helps to know some basics of how modern computational components are realized. Most of us have heard of Silicon, and that computer chips are made from it. Most of us have probably also heard of semiconductors, and are also wondering why something with such a curious name could be so important.

*Conductors* are materials where electricity can flow easily. Metals are a classic example of a conductor. Whether or not a material can conduct electricity depends on its atomic structure, and in particular, whether it facilitates or prevents the movement of electrons. At the atomic level, conductors are characterized by having an overlap between orbits called conduction bands and orbits called valence bands. Conduction bands are where electrons can move freely. *Insulators* are materials where electricity cannot flow. Plastics are a common example of an insulator. In terms of bands, these are insulators that have a “big” gap between the energy level of conduction and valence bands.

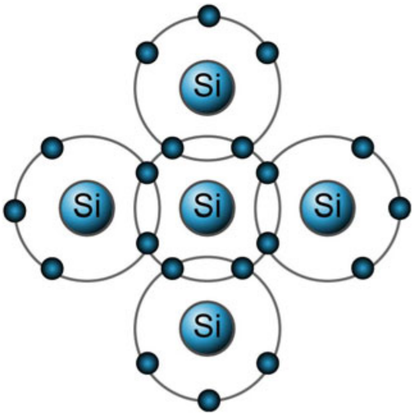

**Fig. 9.1** A crystal of pure Silicon

*Semiconductors* are interesting not because they have a fixed conductivity somewhere between being a conductor or an insulator, but rather, because they can be used to build devices that act as either conductors or insulators depending on an external control signal. In the simplest case, this signal can be electrical. There are many ways in which such effects can be materialized. Diodes are arguably one of the simplest examples, because a diode can act as a conductor or an insulator depending on the direction of the electrical potential we apply across this device itself. Thus, the control signal is the voltage across the device, and the effect we observe is how much current flows across the device as a result of this voltage. Unlike a resistor, the current that will flow through the device will vary dramatically depending on the direction of the voltage. To understand how this works let us take a closer look at how they are built in a semiconductor.

For simplicity, we will consider Silicon as a starting point. Silicon atoms have four electrons in their outer valence shell, and they form crystals by connecting with four other surrounding atoms (Figure 9.1). As a result, this creates a situation where in their outer valence shell each has eight electrons, which is a stable size for that shell. This keeps the electrons in place and makes Silicon an insulator at room temperature. Things become much more interesting when an impurity is added to Silicon, disrupting this stable form slightly and, in the process, giving it very attractive properties. This modification, which is made at the fabrication time, is called *doping*, and can be used to introduce either one free or one missing electron in the crystal lattice (Figure 9.2). Both types of doping, called n-type and p-type semiconductors, respectively, change the conductivity characteristics of the original crystal. But more importantly, when they are put next to each other, they create what is called a *junction*. This type of junction is the basis for creating a wide range of semiconductor devices, such as diodes, transistors, and photo-voltaic cells.

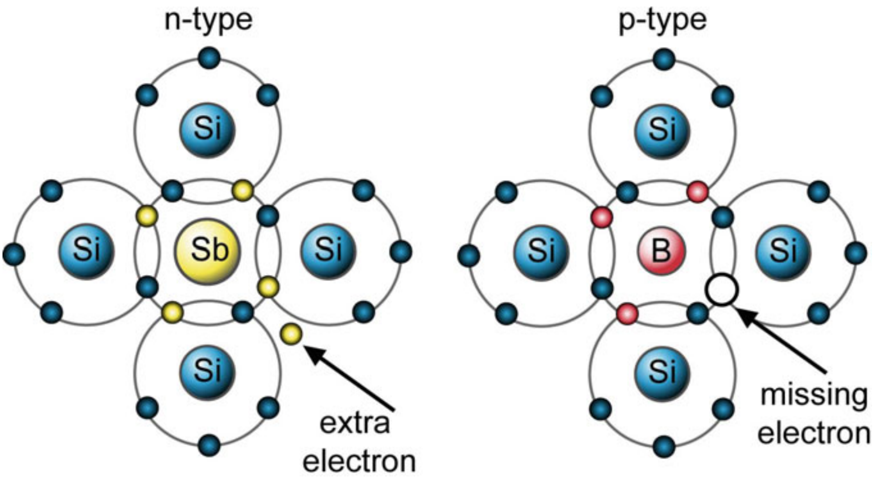

**Fig. 9.2** How n- and p-type doping introduces free and missing electrons

One of the most interesting effects that arises at such a junction is the formation of what is called a thin *depletion region*, which results from the natural migration of the free electrons from one side to fill the hole created by the missing electrons on the other side. The result is that this junction lacks free electrons and is therefore not a conductor. This migration of electrons creates a voltage potential that any electron wanting to travel against needs to overcome. What is more, the size of this depletion region is sensitive to the voltage across the junction, and applied in one direction, this region will grow (and the voltage), but in the other, it will shrink (and the voltage buildup will be negligible). This effect is what gives junctions their ability to allow the flow of current in one direction and not the other, thereby giving us the diagram as a device (Figure 9.3).

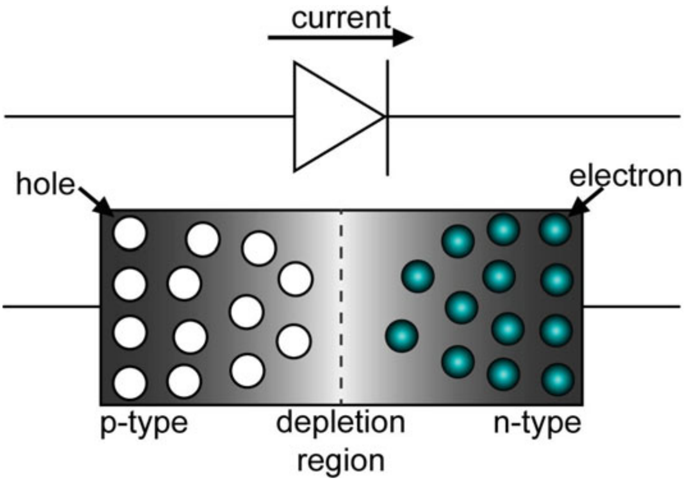

**Fig. 9.3** An n/p junction creates a diode, the most elementary semiconductor device

Diodes by themselves can have a wide range of applications as electric circuit components, including in demodulation of radio signals and building digital logic circuits. For our purposes in this chapter, they help us get a basic appreciation of how semiconductor technologies work, and will help us understand why light is possibly the easiest non-electric physical media that we can connect to a digital circuit.

### 9.2.2 *The Photo-Voltaic Effect*

One of the most interesting aspects of what happens across depletion regions is the involvement of light. When an electron moves from the n-type side to the p-type, it is possible that an electron moves down from the conduction band to a valence band. This is sometimes called a recombination, as it is when an electron meets a “missing” electron in the valence band. The difference in energy between the conduction band and the valence band results in the emission of a photon (Figure 9.4). Depending on its energy, such a photon can form visible light.[1] For traditional circuit applications, such gaps are avoided for efficiency. For LEDs, the device is designed to maximize

---

[1] It should be noted this simpler account is more applicable for semiconductors such as Germanium. For Silicon, other physical effects play a more prominent role than photons.

the chance of the occurrence of such events, and to produce light at a particular frequency.

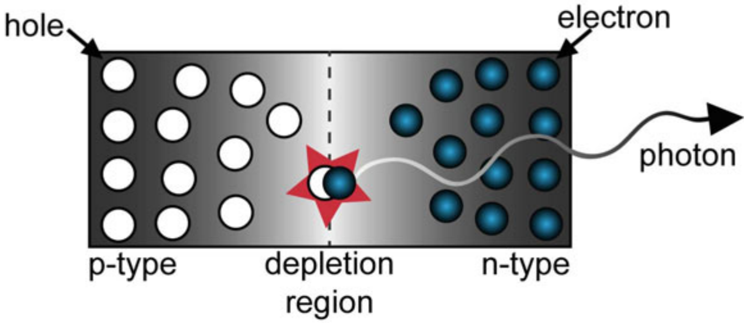

**Fig. 9.4** Light Emitting Diode (LED)

An even more interesting effect is that there is also a dual dynamic: In the abundance of photons, such a junction can have electrons flowing in the opposite direction, increasing the voltage potential across the junction. The voltage potential can allow us to use this junction as a photo-voltaic cell that can be used to detect the presence of light. In the abundance of light and with appropriately configured circuitry, such cells can also be used to harvest electrical energy from this light. The basic dynamic at the level of atomic physics is called the photo-voltaic effect, and is closely related to the photo-electric effect, for which Einstein was awarded the Nobel Prize.

### *9.2.3 Transistors and Amplifiers*

Junctions therefore allow us to build devices such as diodes, LEDs, and photo-voltaic cells. They also allow us to build another important semiconductor device, namely, the transistor. In its simplest form, a transistor can be made by juxtaposing three semiconductors segments with different doping, such as p-type followed by n-type followed by p-type. This configuration creates two junctions and a device with three terminals. Many useful effects can be realized using this device. For example, a small change in the voltage (or current) provided by the middle terminal can have a significant effect on the current that can flow across the two other terminals. This effect can be employed to realize circuits that can amplify the amplitude of a signal by several orders of magnitude. To perform this functional reliably, more than one transistor is used to build an *operational amplifier*, which functions as explained in the chapter on Control. In the context of generating and sensing light, operational amplifiers can be used to boost a digital off/on signal to drive a light emitting diode that delivers brighter light (and can therefore travel further) or amplify a low light signal coming in from a photo-voltaic cell to register clearly as a signal in a digital circuit (Figure 9.5). Amplifiers similarly play an important role in the accuracy with which we can sense external signals, and with which we can drive external devices. They are also used in both analog-to-digital and digital-to- analog converters.

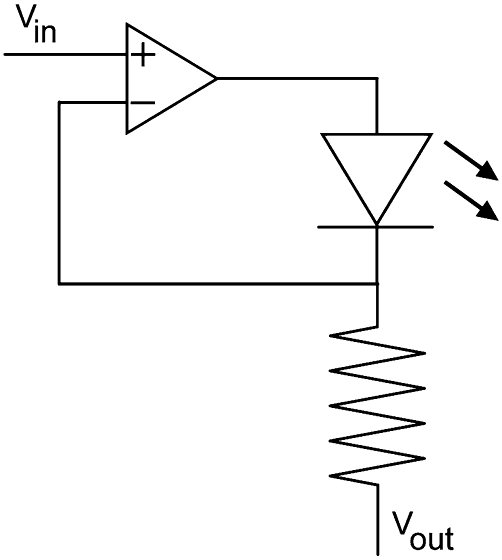

**Fig. 9.5** Operational Amplifiers have numerous applications. In this circuit, one is used to drive an LED

## 9.3 Analog-to-Digital Conversion (ADC)

To transfer an analog signal into a digital computational component we need an analog-to-digital converter (ADC). To transfer a signal from a digital computational component we need a digital-to-analog converter (DAC). Both circuits are best understood as analog circuits and it is simplest to think of the digital value as being represented by the minimum voltage and the highest voltage (such as 0 and 15 V) and the analog signal as being able to have any value in between. For simplicity, we will also assume that we have four bits to represent the signal. This means that we can only represent $2^4$ or 16 values.

A basic strategy for converting the analog signal to a digital one is to start with a simple ladder circuit made of a series of equal resistors that goes from the high voltage to the lowest voltage. In the case of our 16 level circuit, we would use 16 resistors. As long as the resistance on each is equal, the voltage drop across each of them will be equal. This will give us a source for 16 different voltages going from the lowest value of 0 V to the highest of 15 V. Starting from the 1 V point and going up we can start building 15 circuits by feeding this signal into the negative input to

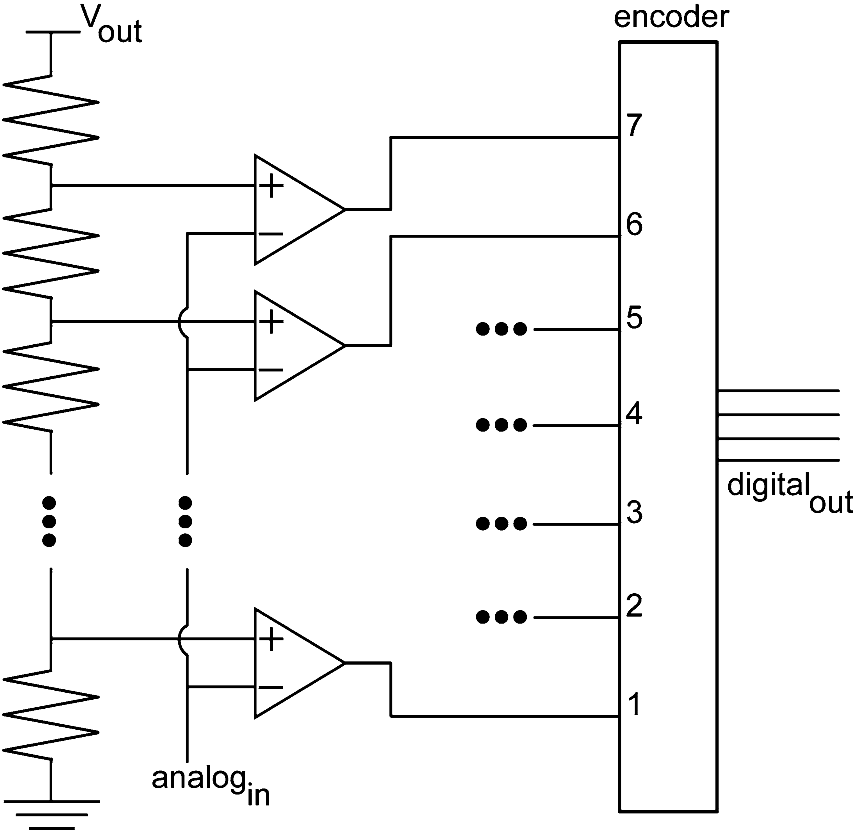

**Fig. 9.6** A ladder circuit to convert from analog to digital

the operational amplifier, and the signal we want to measure to the positive one. This way, the output of each such amplifier will give us a high signal as soon as the input signal is higher that than this voltage, and will produce a very low signal otherwise. We can then treat these output signals as digital signals and collect them in one of several ways, including simply adding them or putting them through a simpler circuit called a priority encoder, which identifies the "highest" of the 15 lines and converts its number into a four-bit representation. Figure 9.6 depicts an example of such a circuit. The following model illustrates the behavior of a ladder circuit:

```
initially
Vs = 1:16,input = 0,input’ = 1,output = 0
always
input’ = 2,
output = sum 1 for v in Vs if input > v
```

The effect of this circuit is essentially the same as computing the `floor` of the input value, which is a more direct model of quantization. Rounding can be viewed as a model of basic analog-to-digital conversion. Depending on the application we can choose to build circuits that realize other rounding operations such as those that computing the ceiling or the closest integer value. Also, if we have more bits or if we have a smaller range of input values, we can let each integer represent a fraction. Again, the effect of such a circuit can be modeled more directly with a rounding function, but we would have to multiply the input first by the denominator of the fraction and then divide the resulting integer by that fraction to recover the value that we are representing.

## 9.4 Digital-to-Analog Conversion (DAC)

A basic strategy for the dual process, namely, digital-to-analog conversion (DAC), also makes use of an operational amplifier. In this case, a classic circuit called the *summing amplifier* is used, whereby an operational amplifier's positive end is connected to ground and the negative end is connected to an input node.The input node is connected via a (denominator) resistance to the output of the amplifier to provide a feedback signal. In addition, the input node is connected to any number of resistors that are connected to the bits encoding analog signal we want to generate. This configuration provides a mechanism for making the output take the value of the sum of (one over) all the other resistances connected to the negative input of the amplifier for the bits set to a high voltage. Figure 9.7 depicts an example of such a circuit.

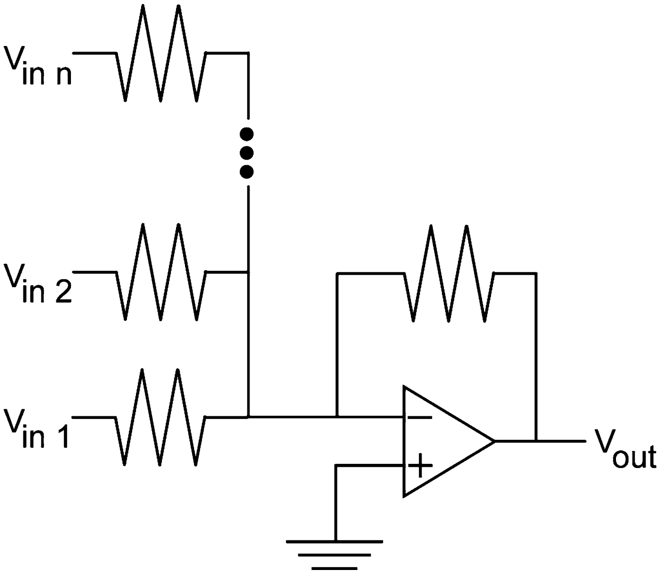

**Fig. 9.7** A summing amplifier to convert from digital to analog

## 9.5 Sensing Temperature

Now that we understand the basics of how an analog signal can be mapped to a digital one, we can now turn to how the analog signals themselves can be generated. Of course, we have already considered light. One of the most commonly measured parameters is temperature. Applications include air-conditioning systems, almost every battery inside a smart device such as a smartphone or a computer, and various mechanical and chemical processes. Interestingly, it is now quite common to measure temperatures inside CPUs to avoid overheating and to respond by stabilizing temperatures by varying workload distribution.

Temperature can generate an electrical effect in a number of ways. Thermocouples are junctions of two different types of metal that produce a temperature-sensitive voltage due to what is called the thermoelectric effect (Figure 9.8). An alternative is a thermistor, which is a device that has a resistance that depends on temperature. Virtually all materials change conductivity depending on temperatures, and materials that have more variation are more suitable for this application. This contrasts to materials chosen for building traditional components, which are chosen to minimize variation with temperature.

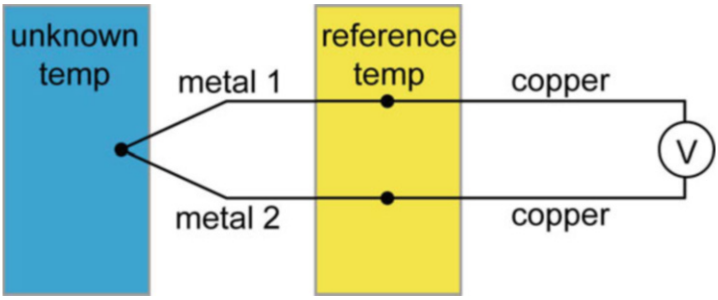

**Fig. 9.8** Example of a circuit design for a thermocouple

## 9.6 Sensing Position

Another type of measurement that is commonly needed in a cyber-physical system is relative position. In the simplest case, a switch can be used to measure closed/open positions, as done in refrigerators and laptops. A more continuous measurement can be made using a device called a *rheostat*, which is a variable resistance device that changes resistance as one of the electrical terminals of the device moves along the resistive material, thereby changing the length that the current travels through the materials, and as a result, changing the total resistance (circuit element illustrated in Figure 9.9). This is a simple and reliable way to measure relative position, and can be used in both linear and angular configurations. However, it does require physical connectivity to the point which we wish to track. Internally, it also has moving parts that slide against each other, which over time can lead to significant wear and tear. For this reason, light (sometimes infrared) is used instead to detect affinity, and indirectly position. More commonly used in practice are rotary encoders that can measure either relative or absolute position using a variety of physical phenomena.

A wide range of techniques can be used for measuring position remotely, that is, without physical connectivity. Depending on the environment, one or more cameras can be used for providing positional information. In indoor environments, ultrasonic sensor can be used. In outdoor environments, systems such as the Global Positioning System (GPS) or LIDAR may be used depending on the demands of the situation.

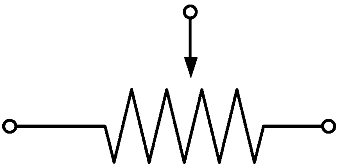

**Fig. 9.9** Circuit notation for a rheostat, a basic device for sensing position

## 9.7 Actuating Mechanical Systems

When it comes to actuating mechanical systems, one of the most direct ways to achieve this is by powering an electric motor. Specialized operational amplifier designs may be used to generate the necessary electric power to drive a Direct Current (DC) motor. Most motors require a particular voltage level to be operated correctly. For this reason, typically, the main parameter that we control in actuating such a motor is how much power is delivered by rapidly turning the power ON and OFF according to a chosen ratio. For example, if we want to deliver no power the signal is OFF $100\%$ of the time. If we want to send full power the signal is ON $100\%$ of the time. If we want $50\%$ power we mix OFF and ON signals in equal proportion. In essence, this provides us with a mechanism for controlling speed. Using feedback control and various mechanical gearing combinations, this approach can also be used to control position.

## 9.8 Chapter Highlights

1. Sensing and Actuation
    (a) Provide the link between computational and physical components in cyber-physical systems
    (b) Switches as one of the simplest input mechanisms
    (c) Missing symmetry
    (d) Why symmetry matters
2. Light as the medium closest to today's implementation technology for the cyber-part
    (a) Can serve as both input and output
    (b) Diodes, LEDs, and Photo-Voltaic cell
3. More on semiconductors
    (a) How semiconductors work
    (b) Transistors as the "next up" from diodes in terms of complexity
    (c) Transistors as the building block for operational amplifiers
    (d) Pervasive role of operational amplifiers in electronics
4. Building the interfaces
    (a) Temperature affects everything
    (b) Measuring relative positions
    (c) Actuating motors

## 9.9 Study Problems

1. Modify the example of Section 9.3 to use 5 bits and increments of 0.5 to represent the continuous input.
2. Simplify the result of the previous exercise by using the `floor` function.
3. Present a circuit for digital-to-analog conversion based on the strategy explained in Section 9.4 for the example discussed in the previous section. Derive the equation for the output as a function of the value of the input bits to show that this circuit will indeed function as a digital-to-analog converter.

## 9.10 To Probe Further

- Tech Lapse's article entitled Intel is working on optical chips for more efficient AI.