

**MODULE 5: SOME BASIC APPLICATIONS**  
**WEEK 5**

**Learning Outcomes:**

After completing this course you are expected to demonstrate the following:

1. Produce a simple example that assurance what we study is not just a collection of abstractions, but it something that helps us to understand many important, real problems.

**A. Engage**

**Trivia**

Although we stress the abstract and mathematical nature of formal languages and automata, it turns out that these concepts have widespread applications in computer science and are, in fact, a common theme that connects many specialty areas

**B. Explore**

Video Title: **Automata in Real Life**

YouTube Link: <https://www.youtube.com/watch?v=sD9eIQRn6mk>

Module Video Title: **Week 5 – Automata in Real Life**

**C. Explain**

Formal languages and grammars are used widely in connection with programming languages. In most of our programming, we work with a more or less intuitive understanding of the language in which we write. Occasionally though, when using an unfamiliar feature, we may need to refer to precise descriptions such as the syntax diagrams found in most programming texts. If we write a compiler, or if we wish to reason about the correctness of a program, a precise description of the language is needed at almost every step. Among the ways in which programming languages can be defined precisely, grammars are perhaps the most widely used.

The grammars that describe a typical language like Pascal or C are very extensive. For an example, let us take a smaller language that is part of a larger one.

**D. Elaborate**

The rules for variable identifiers in C are:

1. An identifier is a sequence of letters, digits, and underscores.
2. An identifier must start with a letter or an underscore.
3. Identifiers allow upper- and lower-case letters.

Formally, these rules can be described by a grammar.

$$\begin{aligned}
 \langle id \rangle &\rightarrow \langle letter \rangle \langle rest \rangle \mid \langle undrscr \rangle \langle rest \rangle \\
 \langle rest \rangle &\rightarrow \langle letter \rangle \langle rest \rangle \mid \langle digit \rangle \langle rest \rangle \mid \langle undrscr \rangle \langle rest \rangle \mid \lambda \\
 \langle letter \rangle &\rightarrow a|b|\dots|z|A|B|\dots|Z \\
 \langle digit \rangle &\rightarrow 0|1|\dots|9 \\
 \langle undrscr \rangle &\rightarrow -
 \end{aligned}$$

In this grammar, the variables are  $\langle id \rangle$ ,  $\langle letter \rangle$ ,  $\langle digit \rangle$ ,  $\langle undrscr \rangle$ , and  $\langle rest \rangle$ . The letters, digits, and the underscore are terminals. A derivation of  $a0$  is

$$\begin{aligned}
 \langle id \rangle &\Rightarrow \langle letter \rangle \langle rest \rangle \\
 &\Rightarrow a \langle rest \rangle \\
 &\Rightarrow a \langle digit \rangle \langle rest \rangle \\
 &\Rightarrow a0 \langle rest \rangle \\
 &\Rightarrow a0.
 \end{aligned}$$

The definition of programming languages through grammars is common and very useful. But there are alternatives that are often convenient. For example, we can describe a language by an accepter, taking every string that is accepted as part of the language. To talk about this in a precise way, we will need to give a more formal definition of an automaton. We will do this shortly; for the moment, let us proceed in a more intuitive way.

An automaton can be represented by a graph in which the vertices give the internal states and the edges transitions. The labels on the edges show what happens (in terms of input and output) during the transition.

For example, **Figure 5.1** represents a transition from State 1 to State 2, which is taken when the input symbol is  $a$ . With this intuitive picture in mind, let us look at another way of describing C identifiers.



**Figure 5.1**

**Figure 5.2** is an automaton that accepts all legal C identifiers. Some interpretation is necessary. We assume that initially the automaton is in State 1; we indicate this by drawing an arrow (not originating in any vertex) to this state. As always, the string to be examined is read left to right, one character at each step. When the first symbol is a letter or an underscore, the automaton goes into State 2, after which the rest of the string is immaterial. State 2 therefore represents the “yes” state of the accepter. Conversely, if the first symbol is a digit, the automaton will go into State 3, the “no” state, and remain there. In our solution, we assume that no input other than letters, digits, or underscores is possible.

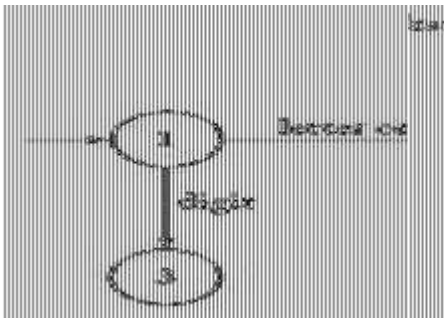


Figure 5.2

Compilers and other translators that convert a program from one language to another make extensive use of the ideas touched on in these examples. Programming languages can be defined precisely through grammars, and both grammars and automata play a fundamental role in the decision processes by which a specific piece of code is accepted as satisfying the conditions of a programming language. The above example gives a first hint of how this is done; subsequent examples will expand on this observation.

A binary adder is an integral part of any general-purpose computer. Such an adder takes two-bit strings representing numbers and produces their sum as output. For simplicity, let us assume that we are dealing only with positive integers and that we use a representation in which stands for the integer

$$x = a_0a_1 \cdots a_n$$
$$v(x) = \sum_{i=0}^n a_i 2^i.$$

This is the usual binary representation in reverse. A serial adder processes two such numbers  $x = a_0a_1 \dots a_n$ , and  $y = b_0b_1 \dots b_n$ , bit by bit, starting at the left end. Each bit addition creates a digit for the sum as well as a carry digit for the next higher position. A binary addition table (**Figure 3.0**) summarizes the process.

		$b_i$	
		0	1
$a_i$	0	0 No carry	1 No carry
	1	1 No carry	0 Carry

Figure 5.3

A block diagram of the kind we saw when we first studied computers is given in **Figure 5.4**. It tells us that an adder is a box that accepts two bits and produces their sum bit and a possible carry. It describes what an adder does, but explains little about its internal workings. An automaton (now a transducer) can make this much more explicit.

The input to the transducer is the bit pairs  $(a_i, b_i)$ , the output will be the sum bit  $d_i$ . Again, we represent the automaton by a graph now labelling the edges  $(a_i, b_i)/d_i$ . The carry from one step to the next is remembered by the automaton via two internal states labelled “carry” and “no carry.” Initially, the transducer will be in state “no carry.” It will remain in this state until a bit pair  $(1, 1)$  is encountered; this will generate a carry that takes the automaton into the “carry” state. The presence of a carry is then taken into account when the next bit pair is read. A complete picture of a serial adder is given in Figure 5.4. Follow this through with a few examples to convince yourself that it works correctly.

As this example indicates, the automaton serves as a bridge between the very high-level, functional description of a circuit and its logical implementation through transistors, gates, and flip-flops. The automaton clearly shows the decision logic, yet it is formal enough to lend itself to precise mathematical manipulation. For this reason, digital design methods rely heavily on concepts from automata theory.

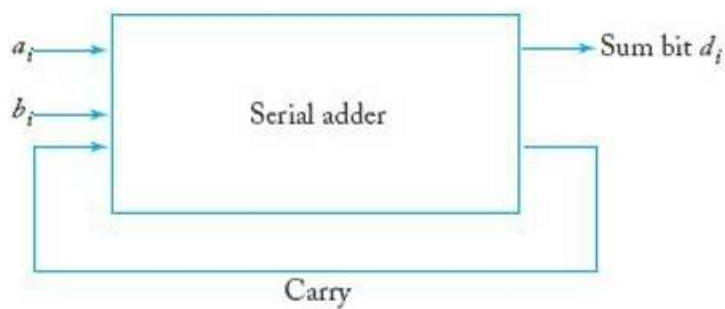


Figure 5.4

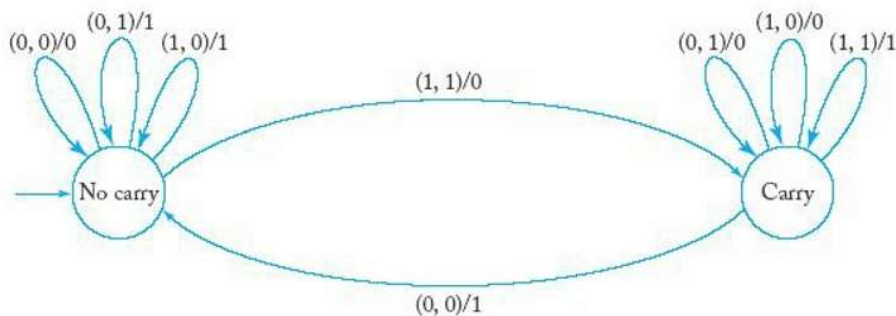


Figure 5.5

E. Evaluation

ASSESSMENT:

**Instructions:** You may write your answer on the Activity Sheet (ACTS) provided in this module.

CONTENT FOR ASSESSMENT:

**Activity Number 4** (20-points).

- 1. Give a grammar for the set of integer numbers in C.
- 2. Design an accepter for integers in C.
- 3. Give a grammar that generates all real constants in C.

References

- 1. *An Introduction to Formal Language and Automata, 5<sup>th</sup> edition, Petez Linz*

Facilitated By:

Name	:	
MS Teams Account (email)	:	
Smart Phone Number	:	