CHAPTER 1

DISCRETE EVENT SYSTEMS AND AUTOMATA

1 Discrete Event Systems

1.1 Short Introduction

Discrete event systems are systems whose dynamic behaviour is driven by asynchronous occurrences of events. Some examples are:

- a manufacturing plant with machines, workers, buffers, etc.;
- a bank with different types of customers and services (desks, ATMs, etc.);
- an airport with passengers in different states (check-in, security control, gate, boarding, etc.);
- a computer system with processes accessing to resources;
- a road system with cars, roads, crosses, traffic lights, etc.;
- a hospital with different types of patients and wards;
- a fast-food restaurant with a staff and different types of customers.

Broadly speaking, discrete event systems can be found in a variety of fields, such as control, computer science, automated manufacturing, and communication, information and transportation networks. In what follows notions regarding modeling, analysis and simulation tools for discrete event systems will be provided. Concepts of automata (deterministic and non-deterministic, untimed and timed) will be introduced as models of discrete event systems. Markov chains also will be addressed as an important class of timed automata.

1.2 Some ideas from Dynamical Systems Theory

The reader must be familiar with basic Dynamical Systems Theory concepts (System formal definition, notions of state, input and output, state transition function and so on). System is one of those primitive concepts (like set or mapping) whose understanding might best be left to intuition rather than an exact definition. Nonetheless, we can provide three representative definitions found in the literature:

- An aggregation or assemblage of things so combined by nature or man as to form an integral or complex whole (Encyclopedia Americana).
- A regularly interacting or interdependent group of items forming a unified whole (Webster's Dictionary).
- A combination of components that act together to perform a function not possible with any of the individual parts (IEEE Standard Dictionary of Electrical and Electronic Terms)

There are two salient features in these definitions. First, a system consists of interacting "components", and second a system is associated with a "function" it is presumably intended to perform. It is also worth pointing out that a system should not always be associated with physical objects and natural laws. For example, system theory has provided very convenient frameworks for describing economic mechanisms or modeling human behavior and population dynamics. As scientists and engineers, we are primarily concerned with the quantitative analysis of systems, and the development of techniques for design, control, and the explicit measurement of system performance based on well-defined criteria. Therefore, the purely qualitative definitions given above are inadequate. Instead, we seek a model of an actual system. Intuitively, we may think of a model as a device that simply duplicates the behavior of the system itself. To be more precise than that, we need to develop some mathematical means for describing this behavior. In particular, from the input-output point of view, a system can be represented as in Figure 1.1, where:

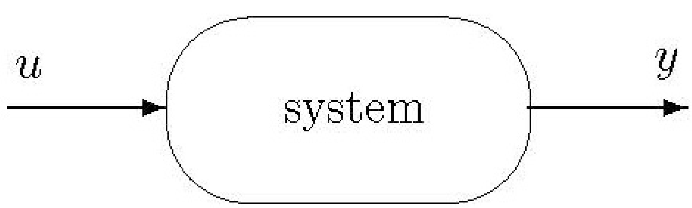

Figure 1.1: Queueing system state transition

- u denotes the system inputs, or independent variables;
- y denotes the system outputs, or dependent variables.

Inputs are variables that can be varied independently of the system, and are fed to the system to modify its behavior. Outputs are variables describing how the systems reacts to the inputs. Different systems may react differently to the same inputs. For instance, the same force may determine a different acceleration depending on the mass of the body to which it is applied. In this sense, the outputs depend on the inputs applied and on the system itself.

Definition 1.1. A system is termed static when the outputs at any time instant depend only on the inputs applied at the same time instant. Models of static systems are just functions mapping the inputs at time t to the outputs at time t:

$$y(t) = f(u(t)) \tag{1.1}$$

An example is the Ohm's Law, mapping the current i passing through a resistor to the voltage v across the resistor:

$$v(t) = Ri(t) \tag{1.2}$$

where R is the resistance.

Definition 1.2. A system is termed dynamic when the outputs at any time instant depend on the whole past history of the inputs.

Differential equations are models of dynamic systems. Consider for instance the current-voltage relation of a capacitor

$$i(t) = C\frac{d\,v(t)}{d\,t} \tag{1.3}$$

If we integrate both size starting from t_0 on we have

$$v(t) = v(t_0) + \frac{1}{C}\int_{t_0}^{t} i(\tau)\,d\tau \tag{1.4}$$

where it is clear that the voltage (system output) at time t depends on the past values of the current (system input) up to time t.

Notion of state

Consider the following problem:

Problem 1 *Given a system, is it sufficient to know the story of the inputs $u(t)$ for all $t \geq t_0$ in order to determine uniquely the outputs $y(t)$ for all $t \geq t_0$?*

For static systems, the answer is yes. What about dynamic systems?

Example 1.1. A constant braking force is applied to a car. We are interested in the space necessary to stop the car. This implies that, in our problem, the car is the system, the braking force is the input, and the displacement is the output. We can write the second Newton's law of motion for the car:

$$M\,a(t) = -f \tag{1.5}$$

where M is the mass of the car, $a(t)$ is the car acceleration, and f is the modulus of the braking force. Recalling that the acceleration is the first derivative of the velocity, and integrating both sides of 1.5 between t_0 (initial time) and $t > t_0$, we get

$$v(t) = v(t_0) - \frac{f}{M}(t - t_0) \tag{1.6}$$

where v denotes the velocity of the car. Recalling that the velocity is the first derivative of the position, and integrating both sides of 1.6 between t_0 (initial time) and $t > t_0$, we further gethere v denotes the velocity of the car. Recalling that the velocity is the first derivative of the position, and integrating both sides of (1.6) between t0 (initial time) and t > t0, we further get

$$\Delta x(t) = v(t_0)\,(t - t_0) - \frac{f}{2\,M}\,(t - t_0)^2 \tag{1.7}$$

where $\Delta x(t)$ denotes the displacement. It can be observed that the displacement does not depend only on the input (the braking force f), but also on the value of the velocity of the car at time t_0. This implies that the knowledge of the input only, is not sufficient to determine uniquely the output. Also the value $v(t_0)$ is needed. Notice that the velocity is neither an input, nor an output of the system. It is rather an additional variable whose value at time t_0 is instrumental in order to determine uniquely the output for a given input. △

We understand that, for dynamic systems, the answer to Problem 1 is no. This leads to the following definition.

Definition 1.3. The state of a dynamic system is a set of variables whose values at time t_0 are necessary to determine uniquely the outputs $y(t)$ for all $t \geq t_0$, given the inputs $u(t)$ for all $t \geq t_0$.

Typically, the state of a system will be denoted by x. Providing the correct definition of state for a system is crucial in any modeling task.

Example 1.2. The queueing system of Figure 1.2 is formed by a single server preceded by

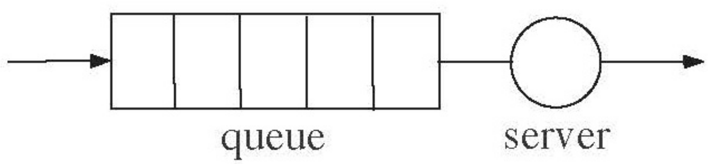

Figure 1.2: Queue system

a queue (or storage space). The total capacity of the queueing system is K, therefore the number of slots in the queue is $K - 1$. Assume that one is interested in whether the server is idle or busy, and to this aim defines the following variable:

$$x = \begin{cases} 0 & \text{if the server is idle} \\ 1 & \text{otherwise} \end{cases} \tag{1.8}$$

The question is whether 1.8 can be taken as a definition of state for the system. The answer is no, and the reasoning is simple. Assume that $x = 1$. This implies that the server is busy. What happens when the current service terminates? Using only the information contained in $x = 1$, it is not possible to determine uniquely whether the next value will be $x^+ = 0$ or $x^+ = 1$. On the basis of Definition 3, the variable in 1.8 is therefore not a state for the queueing system. Notice that the next value will be $x^+ = 0$ if no customer is waiting in the queue, and $x^+ = 1$ otherwise. It is apparent that full information about the number of customers in the system is needed for a proper definition of state. In Example 1.4 this aspect will be clarified. △

Definition 1.4. A system is termed with *continuous state* when the state takes values in a continuous set. It is termed with *discrete state* when the state takes values in a discrete/-countable set.

There exists a class of systems, called *hybrid* systems, for which some components of the state are continuous, and others are discrete.

Time Driven vs. Event Driven dynamic systems

An important classification is between *time-driven* and *event-driven* systems.

Definition 1.5. A system is *time-driven* when all its variables (inputs, state and outputs) may change at any time instant with regular/synchronous ticking.

The above definition implies that time is an independent variable for time-driven systems. There is an external clock, and all system variables may change at any clock tick. Figure 1.3 shows an example of this type, where the observed variable is the ambient temperature in a

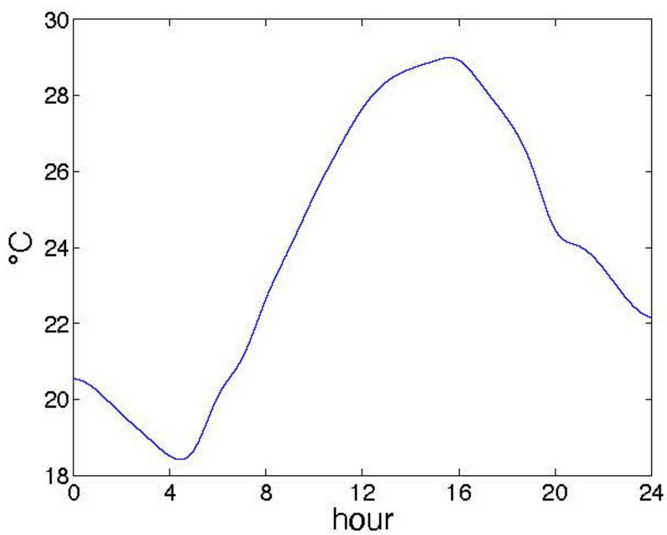

Figure 1.3: Time dependent signal

summer day. Differential equations are typical models used to describe time-driven systems.

Definition 1.6. A system is *event-driven* when all its variables (inputs, state and outputs) changes only upon the (typically irregular/asynchronous) occurrence of events.

It will be clear later that, for event-driven systems, "time" (meaning the time instants when the events occur) is a dependent variable of the system. An example of sample path (or state trajectory) of an event-driven system is shown in Figure 1.4. Its most evident characteristic is that it is piecewise constant, with the time instants of the state jumps determined by the asynchronous occurrence of events.

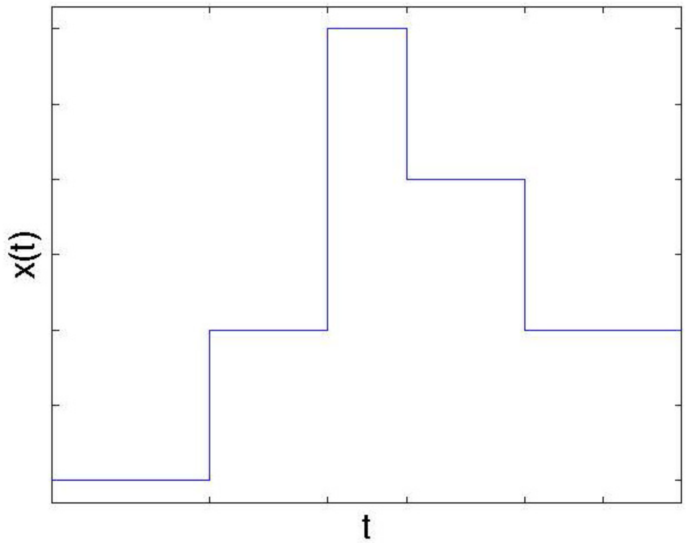

Figure 1.4: Event based signal

Example 1.3. A typical example of event-driven system is the queueing system of Example 1.2. If the state is defined as the number of customers in the system, it is clear that the state changes only upon the occurrence of events, in particular:

- the state is increased by one when the arrival of a new customer occurs (except when the system is full);
- the state is decreased by one when the service of a customer terminates, and the customer leaves the system.

△

An event should be thought of as occurring instantaneously. Events can be of different types:

- specific actions (e.g., somebody presses a button, a customer arrives in a queue, a job is terminated, etc.);
- spontaneous occurrences (e.g., the failure of a component, an interruption of service, etc.);
- fulfillment of logic conditions (e.g., a warning threshold is exceeded, etc.).

In the following, a generic event will be typically denoted by e.

Discrete Event System

We are now ready for the definition of *discrete event system.*

Definition 1.7. A *discrete event system* (DES) is a dynamic, event-driven system with discrete state.

It follows that a DES is characterized by:

- a discrete set E of events;
- a discrete state set X;
- an event-driven state transition function.

There exist real systems that can be naturally viewed as discrete event systems. A typical example is provided next.

Example 1.4. Consider again the queueing system of Example 1.2, and define the state x

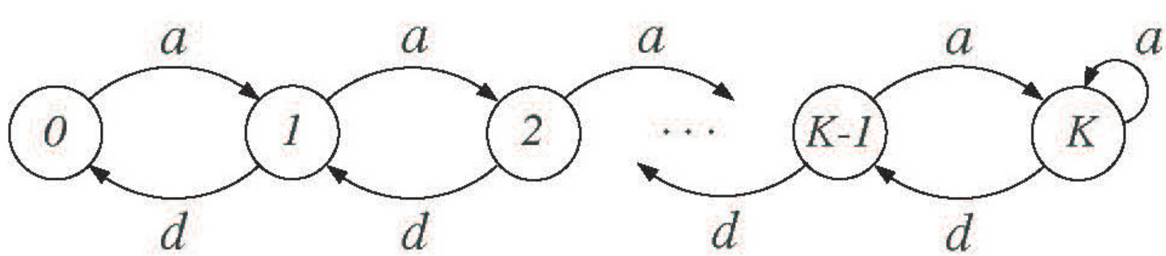

Figure 1.5: Queueing system state transition

as the number of customers in the system. It turns out that x may take values in the discrete set

$$X = \{0,\ 1,\ 2, \ldots,\ K-1, K\}$$

Moreover, the state x may change only upon the arrival of a new customer, or upon the termination of a service (assuming that customers depart from the system after service). We can thus define the event set

$$E = \{a,\ d\}$$

where a denotes the arrival of a new customer, and d denotes the termination of a service in the server. The event-driven dynamics of the queueing system can be represented through the state transition diagram in Figure 1.5. In the state transition diagram, the nodes represent the states of X, whereas labeled arcs represent the state transitions: an arc from x to x^+ with label e means that, when the current state is x and the next event is e, the next state will be x^+. Notice that in state $x = K$ the queueing system is full, and therefore arrivals of new customers (that are still possible) are rejected due to lack of space (the state of the system does not change). Notice also that the state transition diagram contains the information that event d is not possible in state $x = 0$ (there is no arc labeled with d exiting from node 0). Indeed, when the queueing system is empty, no termination of service is possible. $\triangle$

In other cases, real systems that are naturally time-driven, can be modeled as event-driven for particular applications.

Example 1.5. A cart moves along a track. Sensors are located at three points of the track (they are denoted by A, B and C in Figure 1.6). Each sensor sends an impulse when the cart crosses the corresponding point, in both directions. For the sake of simplicity, it is assumed that the cart never changes direction when it is across a sensor. The cart can be naturally

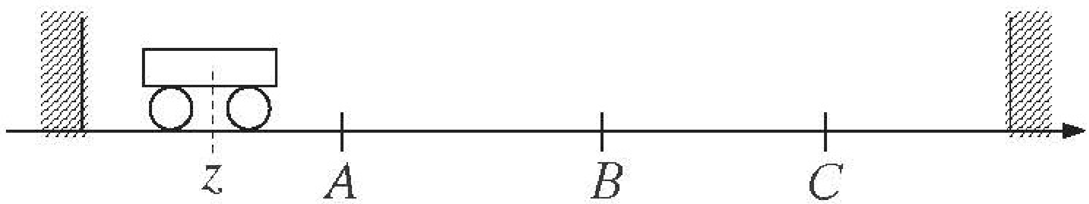

Figure 1.6: Cart with Sensors

seen as a time-driven system, whose motion is modeled by the second Newton's law:

$$M\,\ddot{z}(t) = f(t) \tag{1.9}$$

where M is the mass of the cart, z is the position of the cart along the track, and f is the tractive force applied to the cart. In principle, by integrating twice 1.9, and given the initial position $z(t_0)$, the initial velocity $\dot{z}(t_0)$, and the applied force $f(t)$ for all $t \geq t_0$, it is possible to know from the model the exact position $z(t)$ of the cart at any $t \geq t_0$. In practice, uncertainty on the aforementioned quantities may lead to discrepancies between the real system and the model. However, in some applications, it is neither required nor feasible to know the exact position of the cart. It is sufficient to know the interval in which the cart is localized. To this aim, the input provided by the sensors can be used. Define the state of the system as follows:

$$x = \begin{cases} 0 & \text{if } z(t) < A \\ 1 & \text{if } A \leq z(t) < B \\ 2 & \text{if } B \leq z(t) < C \\ 3 & \text{if } z(t) \geq C \end{cases}$$

Thus, we have a discrete state set

$$X = \{0,\ 1,\ 2,\ 3\}$$

Moreover, define the set of events

$$E = \{a,\ b,\ c\}$$

where a means that an impulse is received from the sensor located at A, and so on. With these definitions, we can construct the state transition diagram in Figure 1.7 Notice that the dynamics of the system so described, turns out to be an event-driven dynamics. △

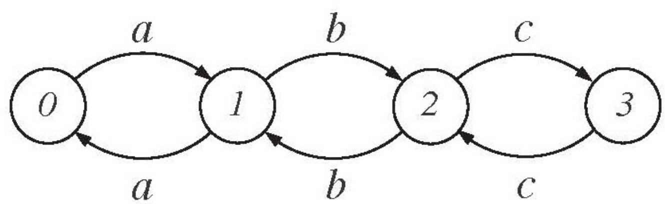

Figure 1.7: Cart System state Transitionn