Control Systems Tehnical Note #68

TIMING CHARACTERISTICS OF DISCRETE EVENT SYSTEMS

June 1989
Revised: November 1989

Kevin M. Passino[1] and Panos J. Antsaklis[2]
Dept. of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, IN 46556

[1]The author acknowledges the partial support of the University of Notre Dame Arthur J. Schmitt Fellowship and the Jet Propulsion Laboratory.
[2]The author acknowledges the partial support of the Jet Propulsion Laboratory.

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

TIMING CHARACTERISTICS OF DISCRETE EVENT SYSTEMS

Kevin M. Passino and Panos J. Antsaklis
Dept. of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, IN 46556

ABSTRACT

A discrete event system (DES) is a dynamical system whose evolution in time develops as the result of the occurrence of physical events at possibly irregular time intervals. Although many DES's operation is asynchronous others have dynamics which depend on a clock or some other complex time schedule. Here we provide a formal representation of the advancement of time for logical DESs via "interpretations of time". It is shown how these interpretations of time provide a common framework to characterize deadlock, simultaneous events, and the advancement of time in various DESs in the literature. We also show that the interpretations of time along with a "timing structure" provide a framework to study principles of the advancement of time for hierarchical DESs (HDESs). In particular, it is shown that for a class of HDESs the event occurrence rate is higher at the lower levels. A relationship between event rate and event aggregation is shown. For another HDES the event rate is shown to be higher at the higher levels but this depends on the definition of the hierarchy.

1.0 INTRODUCTION

One feature of discrete events systems (DESs) that is used to distinguish them from other dynamical systems is that they often operate in an asynchronous fashion. This is, however, not always the case. Sometimes a DES may operate in a synchronous fashion relative to a clock or according to some other complex timing schedule. The variety of timing possibilities can lead to confusion in (i) understanding the dynamics of DESs, (ii) determining how to separate the system timing characteristics from those physical properties that must be modelled to study the design objectives, and (iii) the treatment of deadlock and simultaneous events. Hence, there is the need to formalize the representation and discussion of the advancement of time in DESs. This is the main topic of this technical note which is a summary and extension of results obtained in [Passino 1989;Knight and Passino 1989], including the first steps towards formalizing timing characteristics of hierarchical DESs (HDESs). Comparisons to relevant literature are made throughout the note.

We focus on timing characteristics of (controlled) DESs that can be accurately modelled with

$$P=(X,U,Y,\delta,\lambda,X_0) \tag{1}$$

where if $\mathbb{P}(X)$ denotes the power set of X,

(i) X is the set of plant states x,

(ii) U is the set of plant inputs u,

(iii) Y is the set of plant outputs y,

(iv) δ:U×X$\rightarrow$$\mathbb{P}$(X) is the plant state transition function,

(v) λ:X$\rightarrow$Y is the plant output function, and

(vi) $X_0 \subset X$ is the set of possible initial plant states.

The plant state transition function (a partial, point to set function) specifies for each current input u and state x the set of possible next states $x' \in \delta(u,x)$. The output function specifies for the current state x, the current output symbol $y=\lambda(x)$. Formally, P is equivalent to a directed graph with node set X and edges $x \rightarrow x'$ labelled with u for each triple (u,x,x') such that $x' \in \delta(u,x)$. The states x are labelled with their corresponding output $y=\lambda(x)$. A *run* of P is defined as a sequence of triples $(u_0,x_0,y_0),(u_1,x_1,y_1),(u_2,x_2,y_2)$, ... such that $x_0 \in X_0$, u_0 is the initial input, $x_{k+1} \in \delta(u_k,x_k)$, and $y_k=\lambda(x_k)$. Notice that since it is possible that $\delta(u_k,x_k)=\emptyset$ for all $u_k \in U$ at some $x_k \in X$ a run may have a finite length. The model P can be classified as being a "logical DES model" [Ramadge and Wonham 1989] since, often, it is used to study logical DES properties. Next, we summarize the contents of this technical note.

In Section 2, "index sets" and "index sequences" are shown to provide one possible mathematical characterization of how real time evolves for the logical DES P. After discussing how events can be used for DES synchronization, "interpretations of time" are introduced and shown to be able to formally represent asynchronous, partially asynchronous, and synchronous time. Utilizing our mathematical framework we discuss deadlock and simultaneous events, and show how it can be used to formalize the timing characteristics of a wide variety of DESs studied in the literature.

Motivated by the work of Gershwin [Gershwin 1987] and the observations in [Saridis 1983; Antsaklis, Passino, and Wang 1989] on timing characteristics of hierarchical systems, in Section 3 we provide a study of timing characteristics of HDESs. We introduce a "timing structure" composed of interpretations of time, and input and output triggers to specify how various components of the HDES can influence (be influenced by) the timing characteristics of other components of the HDES. As an example, we use this framework to discuss the timing characteristics of a standard controlled DES. In Proposition 1 it is shown that for a particular HDES if the lower levels have an asynchronous (synchronous) interpretation of time then the higher levels will have an asynchronous (general synchronous) interpretation of time. We propose a definition for DES "event rate" for components of the HDES. Via Theorem 1 and Corrollary 1 it is shown that for a particular HDES the event rate will be higher at higher levels in the HDES.

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

In Theorem 2 we show a relationship between "event aggregation" and event rate, where as events are more aggregated at the higher levels in the HDES we have lower event rates. Finally, via Theorem 3 and Corrollary 2 we give an example of an HDES where the event rate at the higher level is greater that the event rates at the lower levels. But then it is explained how this depends on how the hierarchy is defined.

The results in this note are preliminary. Some additions that need to be made to this work include the development of conditions under which we are guaranteed to have lower event rates at the higher levels for a wider class of HDESs and to examine different ways to define and study properties of various hierarchies in a system. Also, there is a need to fully illustrate the results with examples; certain manufacturing systems will probably best serve our purposes.

2.0 CHARACTERIZING THE ADVANCEMENT OF TIME IN DESs

When a physical plant is modelled as above, the meaning of the advancement of time must be defined. If Z is an arbitrary set then Z^* denotes the set of all finite strings of elements from Z. In order to discuss timing issues for P, an *index set* J and *index sequences* $\alpha \in J^* \cup J^{\mathbb{N}}$ are utilized similar to the approach in [Sain 1981]. The index set J is thought of as a set of times. Let $\mathbb{R}^+$ denote the set of positive real numbers and $\mathbb{R}_+ = \mathbb{R}^+ \cup \{0\}$, the set of non-negative reals and let $\mathbb{N}$ denote the natural numbers. Note that $\mathbb{N}$, $\mathbb{R}_+$, or $\mathbb{R}$ could be candidates for the set J. The index sequences $\alpha \in J^* \cup J^{\mathbb{N}}$ are sequences of time instants that can be of finite or infinite length. If $\alpha \in J^* \cup J^{\mathbb{N}}$ let $|\alpha|$ denote the number of elements in the string α. An index sequence (function) $\alpha \in J^* \cup J^{\mathbb{N}}$ is said to be *admissible* if

(i) it is order preserving, i.e.

(a) if $\alpha \in J^{\mathbb{N}}$ for all $k_1, k_2 \in \mathbb{N}$, $k_1 \leq k_2$ implies that $\alpha(k_1) \leq \alpha(k_2)$ and

(b) if $\alpha \in J^*$ for all $k_1, k_2 \in \mathbb{N}$ with $k_1, k_2 \in [0, |\alpha|-1]$, $k_1 \leq k_2$ implies that $\alpha(k_1) \leq \alpha(k_2)$, and

(ii) it is injective.

Only admissible index functions will be used in the sequel but the possibility of using non-admissible ones is considered in Section 2.3 simultaneous events are discussed.

Following [Sain 1981], the state of the plant $x \in X$ is associated with the index $\alpha(k)$ for some $\alpha \in J^* \cup J^{\mathbb{N}}$ and is denoted with $x(\alpha(k))$, meaning "the state at time $\alpha(k)$". Similarly, inputs $u \in U$ and outputs $y \in Y$ are associated with that same index and denoted with $u(\alpha(k))$ and $y(\alpha(k))$ respectively. The transition to a state in the set $\delta(u,x)$ can be thought of as leading to the next state, with "next" quantified with the index sequence α as $\alpha(k+1)$. With this, the transition function is given as $x(\alpha(k+1)) \in \delta(u(\alpha(k)), x(\alpha(k)))$ which is often

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

abbreviated as $x_{k+1} \in \delta(u_k,x_k)$. The output is often denoted with $y_k=\lambda(x_k)$ for $k\in \mathbb{N}$. Notice that if for some $x_k\in X$ and all $u_k\in U$, $\delta(u_k,x_k)=\emptyset$, then $\alpha(k+1)$ is undefined. Each run of P $(u_0,x_0,y_0),(u_1,x_1,y_1),\ldots$ has an associated index sequence $\alpha(0),\alpha(1),\ldots$ specifying the time instants at which the triples are defined.

Consider Figure 2.1 which illustrates one possible index sequence $\alpha\in J^*\cup J^{\mathbb{N}}$ which specifies the time instants at which the inputs $u\in U$, states $x\in X$, and outputs $y\in Y$ of P are defined.

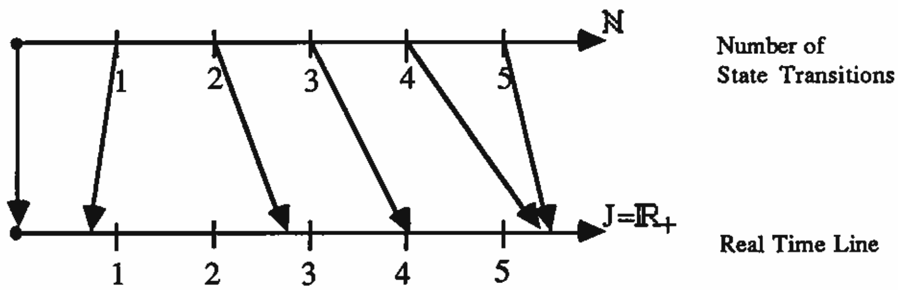

Figure 2.1 Relationships Between State Transitions and Real Time

Intuitively, we think of the top line in Figure 2.1 as representing the number of state transitions that have occurred since the system was started up. The bottom "real time line" represents the index set J, where in Figure 2.1 $J=\mathbb{R}_+$. For a given J, restrictions are put on the allowable sequences of real time instants at which the variables can be defined by requiring the $\alpha\in J^*\cup J^{\mathbb{N}}$ to satisfy the admissibility requirements. All $\alpha\in J^*\cup J^{\mathbb{N}}$ must be order preserving. In terms of Figure 2.1 this means that the arrows specified by α pointing from $k\in \mathbb{N}$ to the real time line will not cross over one another and that time will progress in a nonnegative direction. Also, the fact that α is required to be injective ensures that no two arrows will point to the same time instant on the real time line so that time will advance in a positive direction, i.e., for $x(\alpha(k+1))\in \delta(u(\alpha(k)),x(\alpha(k)))$, $\alpha(k+1)>\alpha(k)$.

The manner in which we think of the timing characteristics of the real plant is represented in the model P. Clearly physical plants with different timing characteristics have, in general, different models P. Based on the characterization of time with index sets and index sequences the meaning of certain logical DES properties can change. For instance, if a certain order of changes in the plant variables is always known to occur (proven in analysis) then when we reflect this information back to the real plant it takes on special meanings depending on the timing characteristics. For example, the sequence of changes may have occurred successively at time instants of one second apart; they may have occurred at random time instants, one after another, etc. Clearly the meaning of

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

logical DES properties depends in an important way on the timing characteristics of the plant.

2.1 DES Synchronization

A DES P often *activates* or *triggers* other DESs to act. For instance, in the case where P represents a plant, P may trigger a controller to generate an input to P. In this case, the trigger often represents certain changes that occur in the plant. There are many different possibilities for how to define the trigger. For instance, in some plants the controller may act when there are changes in the plant output, while in others the controller may be triggered by changes in the plant state. In still other systems the controller may be triggered to act when certain patterns of changes occur. Here, we consider the case where *events*, to be defined next, are used as the trigger. It must be stressed that other triggers are possible. Some of them can be defined by re-defining an event.

Events are thought of as representing certain plant characteristics that change, for instance, instantaneously, and asynchronously. Events could be given by pairs (x,x'), where $x' \in \delta(u,x)$. This is the sort of event used in [Ramadge and Wonham 1987] except that the authors define a finite set of "event labels" σ (which are are also used as the outputs of the plant) for the state transitions, and hence refer to events by triples consisting of event labels, states, and next states. Many other definitions for "events" are possible. What is considered an event is often a matter of taste which is dependent on the particular problem at hand. (See, for instance, the discussion in [Ostroff 1986].) Moreover, for some problems it may be convenient to think of events as taking a certain fixed amount of time to occur, or as occurring at regularly spaced intervals, rather than just instantaneously and asynchronously.

Here, we let $E \subset X \times X$ denote the set of *events* e, where

$$E=\{(x,x') \in X \times X: x' \in \delta(u,x)\} \tag{2}$$

An event (x,x') is said to *occur* if the state transition from x to $x' \in \delta(u,x)$ takes place. Hence, often we use the terms "event" and "state transition" interchangeably. The time instant at which the event occurs is, in general, unspecified. Let $\alpha \in J^* \cup J^N$ and $\alpha(k)$ be the time instant at which the current state $x \in X$ is defined. Let $x' \in X$ denote a possible next state $x' \in \delta(u,x)$, and $\alpha(k+1)$ the time instant at which it is defined. The event $e=(x,x')$ can be thought of as occurring *over* the interval $(\alpha(k),\alpha(k+1)]$, a segment of the time line. It is not necessary to require that the event occur instantaneously at any particular time instant in the interval $(\alpha(k),\alpha(k+1)]$. For convenience, however, we shall for the remainder of this note, assume that the event occurs (is defined) at the time instant $\alpha(k+1)$ where the next state is defined.

Based on the above formulation, in [Passino 1989] the author defines a discrete event regulator system (DERS) and describes how time advances. Other discussions on the timing of the DERS are provided in [Knight and Passino 1989]. The timing of the DERS is similar to that which was set up in the Ramadge-Wonham framework in [Ramadge and Wonham 1987]. In their "controlled generators" the finite set of events label state transitions and are the outputs of the plant. As outputs, the events are used to "drive" the state transitions in their controller, called a "supervisor". In [Passino 1989] and [Knight and Passino 1989] the events are used as a trigger but the outputs of the plant are not necessarily the events.

2.2 Interpretations of Time

The pair $I=(A,J)$ where J is an index set and $A \subset J^* \cup J^{\mathbb{N}}$ will be referred to as an *interpretation of time* since it specifies the meaning of the advances in time, i.e. it specifies the time instants where the variables of P are defined. In general, a system P is said to have a particular interpretation of time $I=(A,J)$ as long as the time instants associated with the elements of the runs of P are elements of J and the index sequences associated with the runs of P are elements of A. The *admissible interpretation of time* will be denoted with $I_{ad}=(A_{ad},J_{ad})$ where J_{ad} is an index set and

$$A_{ad}=\{\alpha \in J_{ad}^* \cup J_{ad}^{\mathbb{N}} : \alpha \text{ is admissible}\}. \tag{3}$$

Notice that for any admissible interpretation of time after an event occurs another may not eventually occur; hence deadlock can be modelled directly. Most often we can choose $J_{ad}=\mathbb{R}_+$ and this is what we will assume here. Next, different interpretations of time are used to characterize different manners in which time can advance in DESs.

It is common to consider the timing characteristics of DESs relative to a *clock*. By a "clock" we mean a device which has a fixed interval $T \in \mathbb{R}^+$ between *ticks* and which does not stop ticking.

Definition 1: The *asynchronous interpretation of time* is $I_a=(A_a,J_a)$ where $J_a=\mathbb{R}_+$ and

$$A_a=\{\alpha \in A_{ad} : \alpha(0)=0\}.$$

According to convention $J_a=J_{ad}=\mathbb{R}_+$ with the time instant of zero corresponding to the case where no state transitions have occurred. I_a represents the situation where the plant P is asynchronous (out of sync, not synchronous) with a clock. For I_a the time instants at which the plant variables are defined are at non-uniform (irregular) distances from one another along the time line $\mathbb{R}_+$. Notice that $A_a \subset A_{ad}$ if $J_{ad}=J_a$.

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

Definition 2: The *partially asynchronous interpretation of time* is $I_{pa}=(A_{\gamma\beta},J_{pa})$ with $J_{pa}=\mathbb{R}_+$ and $A_{\gamma\beta}=\{\alpha\in A_a : \alpha(k)+\gamma\leq\alpha(k+1)\leq\alpha(k)+\beta\}$ for $\gamma,\beta\in\mathbb{R}^+$ where $\beta\geq\gamma$.

I_{pa} represents the case where we know that the time instants where the next state is defined is constrained to occur at least γ, and no more than β time units later. Notice that $A_{\gamma\beta}\subset A_a$ if $J_{pa}=J_a$.

Definition 3: The *general synchronous interpretation of time* is $I_s=(A_T,J_s)$ with $J_s=\mathbb{R}_+$ and $A_T=\{\alpha \in A_a : \alpha(k+1)=\alpha(k)+nT \text{ where } n\in \mathbb{N}-\{0\}\}$ with $T\in\mathbb{R}^+$.

For the general synchronous interpretation of time, the time instants at which the plant variables x, u, y are defined are at distances nT, for $n\in\mathbb{N}-\{0\}$, from one another along the time line $\mathbb{R}_+$. Notice that, in general, a state transition may not occur between any two particular ticks of the clock (since n>0), and that after each state transition occurs another may not eventually occur. When n=1 we shall refer to I_s simply as the *synchronous interpretation of time*. For the synchronous interpretation of time if is not necessarily the case that $|A_T|=1$ since any finite length index sequence is possible. Notice that $A_T\subset A_{\gamma\beta}$ provided that $\gamma\leq T\leq\beta$.

Also notice that for asynchronous time if $\alpha_a\in A_a$ and the current time is $\alpha_a(k)$, then the next time is $\alpha_a(k+1)=\alpha_a(k)+r$ where $r\in\mathbb{R}^+$. On the other hand, for general synchronous time if $\alpha_s\in A_T$ the current time is $\alpha_s(k)$ then the next time is $\alpha_s(k+1)=\alpha_s(k)+r'$ where $r'\in R_1$ and $R_1=\{nT: n\in\mathbb{N}-\{0\}\}$ for a given $T\in\mathbb{R}^+$. Since R_1 is equinumerous with $\mathbb{N}-\{0\}$, a proper subset of $\mathbb{R}^+$, it is the case that $\text{card}(\mathbb{R}^+)>\text{card}(R_1)$. Hence, no matter what the time interval T, the number of possible "next" time instants is always greater if an asynchronous interpretation of time is used rather than a synchronous one. This helps to clarify the intuitions we have about the relationships between the synchronous and asynchronous interpretations of time. It is clear that synchronous time cannot be used if the underlying system can only be accurately represented with an asynchronous interpretation. However, it is possible that the synchronous interpretation of time with T very small may result in an accurate model for some asynchronous systems. This will depend on the particular plant to be modelled and the design objectives to be studied.

It is not, in general, required that the timing characteristics of the plant be defined relative to a clock although they are often treated as such. In general, in a manner similar to that with the clock, the plant may be in sync (out of sync) with changes in other systems.

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

The knowledge of this fact may influence the meaning of certain logical DES properties in a manner similar to how it did for the clock. Intuitively, the representation of the timing characteristics with the interpretations of time (i) clarifies how to go about modelling physical systems because it separates complex timing characteristics from the physical characteristics of the plant which must be represented with P, and (ii) helps to show how the timing characteristics of the physical plant can change and still have a valid model P. The model P will always reflect the particular manner in which the timing characteristics of the real plant are understood so analysis performed using the model P is performed independent of the interpretation of time. The results of the analyses, however, often have a special meaning based on the interpretation of time.

2.3 Timing Characteristics of Various DESs

In this Section we discuss several timing issues encountered in DES studies. First, the issue of deadlock and simultaneous events is discussed; then examples of places where these timing issues have been addressed in the literature are given. It is also shown how the interpretations of time can be used to discuss the timing characteristics of a wide variety of DESs found in the literature.

Deadlock and Simultaneous Events

A DES P is said to *deadlock* if P enters any state $x \in X$ such that for all $u \in U$, $\delta(u,x)=\emptyset$. It is important to note that the interpretations of time defined above allow for the direct modelling of deadlock since the index sequences can have finite length. Deadlock issues are addressed in most of the work on logical DES in one way or another. For an introductory treatment in a Petri net framework see [Peterson 1981]. Another notable study on deadlock in DES is given in [Li and Wonham 1988].

Next, the issue of *simultaneous events* is discussed. Let $I_{ad}=(A_{ad},J_{ad})$ be the admissible interpretation of time. Due to the injective part of the admissibility requirement for all $\alpha \in A_{ad}$, the variables x, u, and y are defined at time instants which are distinct from one another. For instance, if $x(\alpha(k))$ is the current state and $x(\alpha(k+1)) \in \delta(u(\alpha(k),x(\alpha(k)))$ is the next state then $\alpha(k+1)>\alpha(k)$. Since the time instants are distinct, when state transitions occur it is guaranteed that time will advance (although it may be a very small amount). The other important implication is that using the definition of events E it is automatically assumed that events occur at distinct times, i.e., simultaneous events are not allowed because the index sequences are required to be admissible.

Suppose that the injective part of the admissibility requirement is omitted so that for $\alpha \in J^* \cup J^{\mathbb{N}}$ it is possible that $\alpha(k+1)=\alpha(k)$ for any $k \in \mathbb{N}$ such that $\alpha(k)$ and $\alpha(k+1)$ are

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

defined. This will allow events to occur simultaneously at a particular time instant. In fact, for $\alpha \in J^{\mathbb{N}}$ it will allow even an infinite number of events to occur at one time instant resulting in the possibility that time will not advance. Normally, to treat simultaneous events, only a finite number are allowed to occur at a single time instant; hence, other events representing the case that "several events occur at once" can often be defined. So the problem of dealing with simultaneous events is often transformed to the case where only a single event occurs at each time instant so that time is guaranteed to advance and the admissible interpretation of time as defined in (3) above can be used. Examples of DES studies where (a finite number of) simultaneous events are allowed can be found in [Krogh 1987], [Li and Wonham 1987], and [Ichikawa and Hiraishi 1987].

Interpretations of Time in DES Studies

The asynchronous interpretation of time, as formalized above (Definition 1), agrees with the standard interpretation of time used in most logical DES studies. (Sometimes, however, it is known that P will not deadlock so it is known that all index sequences $\alpha \in J^{\mathbb{N}}$.) For instance, the work that has evolved from the Ramadge-Wonham formulation [Ramadge and Wonham 1987], research on Petri nets (See, for instance the references in [Peterson 1981]), and many studies in temporal logic all depend on this asynchronous interpretation of time. Indeed, the asynchronicity of a DES is often quoted as one of the characteristics that distinguishes them from other dynamical systems such as those modelled with differential equations. The partially asynchronous interpretation of time is used in [Bertsekas and Tsitsiklis 1989] in the study of computer networks.

The synchronous interpretation of time has also been used in certain DES studies. Indeed, this is often thought to be the normal way of thinking about time for automata-theoretic models [Hopcroft and Ullman 1979] since such models are often thought to operate in a conventional computer driven by a clock. It is important to note that the knowledge that a particular plant is synchronous can be exploited in DES studies because in this case another characteristic of the DES is known. For instance, in [Ostroff 1986] an "extended state machine" is shown to be able to model a wide variety of DES including a clock. From this, the author defined a useful "real time temporal logic" for studying "hard" time constraints in DES.

Another notable study where time is thought of as being synchronous is given in [Gershwin 1987]. There, Gershwin defines a hierarchy where at each level i, events occur at different frequencies, say $f_{ei}=1/T_{ei}$ where the T_{ei} are fixed positive intervals. Inherent in this, is the assumption that time is synchronous at each level i. At the higher levels of the hierarchy he defined the frequency of the events as smaller while at the lower levels, the frequency of event occurrence is higher. He thinks of the hierarchy as being split according to its "spectra" of event frequencies. The timing characteristics of each specific level of Gershwin's hierarchy of events can be formalized using the interpretations of time defined above. At each level i, the synchronous interpretation of time (n=1) can be used with time interval T_{ei}.

Gershwin's work is also valuable in that it highlights an important property of the timing of event occurrences in HDESs: *that events seem to occur more often at higher levels in the hierarchy*. In the next Section we verify this timing principle for a wide class of HDESs.

3.0 TIMING CHARACTERISTICS OF HIERARCHICAL DESs

Motivated by the work of Gershwin described in Section 2.0, in this Section we shall study the evolution of time in HDESs. The formation of a control theory for HDESs is just beginning [Zhong and Wonham 1988,1989] even though such systems occur quite frequently. Some principles of the evolution of time in hierarchical systems have been postulated but not fully investigated in [Albus, Barbera, and Nagel 1981; Saridis 1983; Valavanis 1986; Mesarovic, Macko, and Takahara 1970; Antsaklis, Passino, and Wang 1989; Passino and Antsaklis 1988]. What these researchers have recognized is that "things usually occur at the higher rates at the lower levels in a hierarchical system". We shall verify this intuition for one class of HDESs here.

We shall focus on HDESs that have as components two types of DESs, G_j, for $j=1,2,...,m$ and P_i for $i=1,2,...,n$ all defined via (1) except with different timing characteristics. We introduce what we call a *timing structure* which will define how the various components of the hierarchical system influence (are influenced by) the timing characteristics of other components of the HDES. The definition of the timing structure is based on the interpretations of time defined in Section 2.0 and what will be called *input* and *output triggers*. Each P_i for $i=1,2,...,n$ in the HDES has timing characteristics that are specified via their own interpretation of time $I_{pi}=(A_{pi},J_{pi})$. Each G_j for $j=1,2,...,m$ has timing characteristics that depend on P_i for $i=1,2,...,n$ and G_k for $k\neq j$ as we now discuss.

Let E_{pi} denote the set of events for P_i, and E_{gj}, the set of events for G_j both defined in a similar manner to the events E for P in (2). The *output triggers* for P_i $i=1,2,...,n$ and G_j $j=1,2,...,m$ are defined by the events E_{pi} and E_{gj} respectively. The *input triggers* for the G_j are defined by

$$\tau_j:E_{p1}\times E_{p2}\times \cdots \times E_{pn}\times E_{g1}\times E_{g2}\times \cdots \times E_{gk}\times \cdots \times E_{gm}\rightarrow\{0,1\} \qquad (4)$$

where $k\neq j$ and $\tau_j(\cdot)=1$ (=0) indicates that an event $e_{gj}(\alpha(k+1))\in E_{gj}$ where $e_{gj}(\alpha(k+1))=(x_g(\alpha(k)),x_g(\alpha(k+1)))$ is forced (not) to occur in G_j. Let $\alpha_{pi}(k+1)$ and

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

$\alpha_{gk}(k+1)$ denote the time instants at which events $e_{pi} \in E_{pi}$ and $e_{gk} \in E_{gk}$ $(k \neq j)$ occur respectively. The time instant at which $e_{gj}(\alpha(k+1))$ occurs (is defined) is given by

$$\alpha(k+1)=\max_{i,k\neq j}\{\alpha_{pi}(k+1),\alpha_{gk}(k+1)\}. \tag{5}$$

This time instant corresponds to the time instant at which the last event occurred which caused $\tau_j(\cdot)=1$.

Whereas the interpretation of time is always specified for the P_i, the interpretations of time for the G_j are specified in terms of the other G_k, $k\neq j$ and P_i. The interpretation of time for any G_j is found by executing all possible runs (in all possible orders) of the P_i for $i=1,2,\ldots,n$ and G_k for $k=1,2,\ldots,m$ where $k\neq j$. Then via equations (4) and (5) the time instants and hence index sequences and interpretations of time for the G_j are specified. We shall study HDESs where there is at least one P_i and the interpretations of time for the G_j can be defined in terms of P_i. The following example will serve to clarify how the interpretation of time for a G_j can be specified via a P_i.

Consider the controlled DES shown in Figure 3.1.

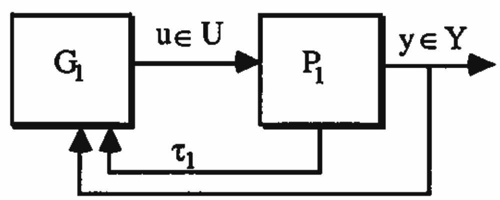

Figure 3.1 Discrete Event Control System

We have $\tau_1:E_{p1}\rightarrow\{0,1\}$ and for the standard control configuration it is most often assumed that $\tau_1(e_{p1})=1$ for all $e_{p1}\in E_{p1}$ so that each time an event occurs in P_1, G_1 is forced to act by having an event in G_1 occur (it is assumed that one always exists). Clearly, then if $I_{p1}=(A_{p1},J_{p1})$ is the interpretation of time for P_1 and $I_{g1}=(A_{g1},J_{g1})$ for G_1 where $J_{g1}=J_{p1}$, then $A_{g1}=A_{p1}$. The interpretation of time for the plant and controller are the same. In this way we think of specifying the interpretation of time for G_1 by I_{p1} and P_1 via τ_1. Intuitively, for general τ_1 we see that we can expect fewer events to occur in G_1 than in P_1. This timing characteristic is treated in some detail next.

Consider the HDES shown in Figure 3.2. Let the admissible interpretation of time for P_1 be $I_{p1}=(A_{p1},J_{p1})$ with $J_{p1}=\mathbb{R}_+$ and for G_j, $j=1,2,\ldots,m$ be $I_{gj}=(A_{gj},J_{gj})$.

<u>Proposition 1</u>: If P_1 has the asynchronous (synchronous) interpretation of time then G_j for $j=1,2,\ldots,m$ has an asynchronous (general synchronous) interpretation of time.

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

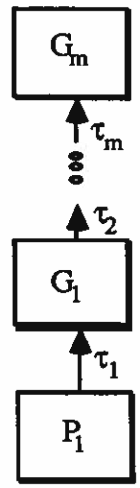

Figure 3.2 Hierarchical DES with m+1 Levels

Proof: Choose $J_{gj}=J_{p1}$ for all j. Assume that a run is made in P_1 and that its corresponding index sequence is $\alpha_{p1}\in A_{p1}$. The corresponding runs in G_j have index sequences $\alpha_{gj}\in A_{gj}$ for all j. Clearly, the I_{gj} are admissible if we choose $\alpha_g(0)=0$ and the result for the asynchronous case follows immediately. For the synchronous case, the map τ_j may mask certain events ($e_{p1}\in E_{p1}$ or $e_{gj}\in E_{gj}$) from forcing events to occur at the next level up ($e_{g1}\in E_{g1}$ or $e_{gj+1}\in E_{gj+1}$) so events may not occur between every two ticks of the clock (but still only at the ticks) resulting in a general synchronous interpretation of time for each G_j. ∎

Notice that even if all of the runs in P are of infinite length there may be a deadlock at the higher levels of the hierarchy (particularly at G_1). If G_j for j=1,2,...,m deadlocks then G_k for k>j will also deadlock.

Theorem 1: For any possible run made by P_1 with an index sequence $\alpha_{p1}\in A_{p1}$, the corresponding runs in G_j for j=1,2,...,m have index sequences $\alpha_{gj}\in A_{gj}$ where α_{gj} is a subsequence of both α_{gk} where k<j and α_{p1} .

Proof: The input trigger τ_1 for G_1 may mask event occurrences in P_1. If event $e_{p1}\in E_{p1}$ occurs in P_1 at time $\alpha_{p1}(k)$ and $\tau_1(e_{p1})=0$, but later another event $e'_{p1}\in E_{p1}$ occurs at time $\alpha_{p1}(k+1)$ and $\tau_1(e'_{p1})=1$ then the index sequence α_{g1} will contain $\alpha_{p1}(k+1)$ but not $\alpha_{p1}(k)$. Since this is true for all $k\in \mathbb{N}$, α_{g1} is a subsequence of α_{p1}. A similar argument holds for the higher levels. ∎

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

Define $\#(P_i,T_u)$ and $\#(G_j,T_u)$ to be the number of events that occur per unit time $T_u=(r_1,r_2]$ where $(r_1,r_2]\subset\mathbb{R}^+$ for P_i or G_j respectively. We shall refer to $\#(P_i,T_u)$ or $\#(G_j,T_u)$ as the *event occurrence rate* or *event rate* in P_i or G_j respectively. Notice that if P_i has a synchronous interpretation of time with time interval $T\in\mathbb{R}^+$ and we choose T_u such that $|r_2-r_1|=T$ then $\#(P_i,T_u)=1$, i.e., there is 1 event occurrence in the time interval T_u. In the case where there is deadlock we also take $\#(P_i,T_u)=1$ so our definition of event rate only pertains to P_i when it is not deadlocked and the possibilities for T_u must be chosen accordingly. If P_i has an asynchronous interpretation of time then no matter how T_u is chosen it is possible that $\#(P_i,T_u)=0$, since we cannot guarantee that an event will occur in the time interval T_u. In fact, we do not know how many events will occur in T_u. It would appear that our definition of event rate is too restrictive. This is, however, not the case since the focus here is on *comparing* the event rates of different DESs and this comparison is made relative to T_u an interval of the real time line.

Corollary 1: $\#(P_1,T_u)\geq\#(G_1,T_u)\geq\#(G_2,T_u)\geq\cdots\geq\#(G_m,T_u)\geq 0$ for all T_u.

Proof: Assume that a run is made in P_1 and that its corresponding index sequence is $\alpha_{p1}\in A_{p1}$; the corresponding run in G_1 has index sequence $\alpha_{g1}\in A_{g1}$. Each time instant $\alpha_{p1}(k)$ and $\alpha_{g1}(k)$ for $k>0$ corresponds to an event occurrence. For any T_u, $\#(P_1,T_u)\geq\#(G_1,T_u)$ since the map τ_1 may mask events. A similar argument holds for the higher levels. ∎

This means that the event rate is lower at the higher levels of the HDES shown in Figure 3.2. In the case where I_{p1} is synchronous the above results support the studies in [Gershwin 1987] where the author assumes that the event rates can be split into "spectra" according to the level in the hierarchy. A similar split can be made for the HDES of Figure 3.2.

Next we study the case where certain restrictions are put on τ_j and characterize different properties of the HDES shown in Figure 3.2. For simplicity consider the case where m=1. Assume that the state set X of P can be partitioned as

$$X=X_1\mathring{\cup}X_2\mathring{\cup}X_3\mathring{\cup}\cdots \tag{6}$$

(where $\mathring{\cup}$ denotes a disjoint union) so that if P enters a state $x\in X_i$ it will take more than $\pi\in\mathbb{N}-\{0\}$ state transitions before the state of P, say x', is such that $x'\in X_j$ where $j\neq i$. Let

$$\mathfrak{X}(\pi)=\{X_1,X_2,X_3,\ldots\}. \tag{7}$$

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

We shall study the case where events occur if P_1 switches states from being one in X_i to another in X_j where $X_j \neq X_i$ and τ_j will be defined accordingly. Such a definition for τ_j results in a type of "event aggregation" since some sequences of events will be ignored by the higher levels.

Theorem 2: Let $\tau_1(e)=1$ where $e=(x,x')$, $x,x' \in X$, and there exists $X_i, X_j \in \mathfrak{X}(\pi)$ such that $x \in X_i$ and $x' \in X_j$ where $X_i \neq X_j$ ($\tau_1(e)=0$ otherwise). In this case, $\#(P_1,T_u) \geq \#(G_1,T_u)+\pi$ if $T_u=(r_1,r_2]$ and $|r_2-r_1|$ is sufficiently large.

Proof: Assume that a run is made in P_1 and that its corresponding index sequence is $\alpha_{p1} \in A_{p1}$; the corresponding run in G_1 has index sequence $\alpha_{g1} \in A_{g1}$. Let $\alpha_{g1}(k)$ and $\alpha_{g1}(k+1)$ be two elements that make up α_{g1} such that $|\alpha_{g1}(k+1)-\alpha_{g1}(k)|$ is a maximum for all k. Let $T_u=(r_1,r_2]$ such that $|r_2-r_1|>2|\alpha_{g1}(k+1)-\alpha_{g1}(k)|$. In this case the time interval T_u will always contain 2 event occurrences in G_1. By the definition of τ_1, between any two event occurrences in G_1 there must be at least π event occurrences in P_1; therefore $\#(P_1,T_u) \geq \#(G_1,T_u)+\pi$. ∎

For the case where there are multiple levels, i.e., m>1, if the τ_j for j>1 are defined as τ_1 is in Theorem 2 (via (6) and (7)), then $\#(G_j,T_u) \geq \#(G_{j+1},T_u)+\pi_j$ (provided T_u is chosen as in Theorem 2) where π_j is the number of events that must occur in G_j before one can occur in G_{j+1}. The τ_j can be viewed as maps that cause event aggregation; consequently, Theorems 1 and 2 provide a relationship between event aggregation and event rate for one class of HDESs: *as events are aggregated to the higher levels in the hierarchy, fewer events occur*. In general hierarchical systems researchers have observed a similar inverse relationship between "time scale density" ("time granularity") and "model abstractness" [Saridis 1983; Antsaklis, Passino, and Wang 1989]. This type of relationship does not, however, hold in general. Next, we provide a HDES where the event rate is higher at a higher level in the hierarchy, but this depends on the definition of the hierarchy. For the moment assume that if a system's timing characteristics are influenced by another's, it is at a higher level in the hierarchy (this is what is assumed for the above results).

Consider the HDES shown in Figure 3.3. In this case $\tau_1 : E_{p1} \times \cdots \times E_{pn} \rightarrow \{0,1\}$. Assume that $\tau_1(e_{p1},e_{p2},\ldots,e_{pn})=1$ for all $e_{pi} \in E_{pi}$. If any P_i has the asynchronous interpretation of time then G_1 has a asynchronous interpretation of time. To get deadlock in G_1 there will have to be deadlock in all the lower level DESs P_i.

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

Theorem 3: For any possible run made by any P_i for i=1,2,...,n with an index sequence $\alpha_{pi} \in A_{pi}$, the corresponding run in G_1 has index sequence $\alpha_{g1} \in A_{g1}$ where each α_{pi} is a subsequence of α_{g1}.

Proof: The proof is similar to the proof for Theorem 1. ∎

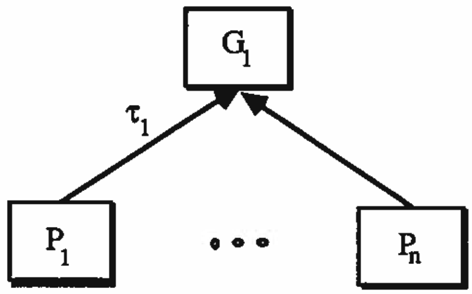

Figure 3.3 Hierarchical DES

Corollary 2: $\sum_i \#(P_i,T_u) \geq \#(G_1,T_u) \geq \#(P_i,T_u)$ for all T_u and i=1,2,...,n.

Proof: It follows that $\#(G_1,T_u) \geq \#(P_i,T_u)$ for all T_u and i=1,2,...,n in a manner similar to the proof of Corollary 1. Since it is possible that an two events in different P_i occur at the same time, the greatest number of events that can occur in G_j in one time interval T_u is given by the sum of all the events that occur in the P_i in T_u. ∎

Hence the event rate is higher for G_1 than for any of the other DESs P_i. This shows one case where a DES at a higher level has an event rate that is higher than the event rate for a DES at a lower level. This of course assumes a rather narrow definition for what it means for a system to be at a higher level in the hierarchy. In a "time scale hierarchy" one could define systems that operate with higher event rates to be at lower levels, then the above results (Theorem 3, Corrollary 2) agree with the intuition that at higher levels, event rates are lower.

Acknowledgement: The authors gratefully acknowledge the partial support of the Jet Propulsion Laboratory and a University of Notre Dame, Arthur J. Schmitt Fellowship for the first author.

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

4.0 REFERENCES

[Albus, Barbera, and Nagel 1981] Albus J.S., Barbera A.J., Nagel R.N., "Theory and Practice of Hierarchical Control", Proc. of the 23rd IEEE COMPCON, pp. 19-39, 1981.

[Antsaklis, Passino, and Wang 1989] Antsaklis P.J., Passino K.M., Wang S.J., "Towards Intelligent Autonomous Control Systems: Architecture and Fundamental Issues", J. of Intelligent and Robotic Systems, Vol. 1, pp. 315-342, 1989.

[Bertsekas and Tsitsiklis 1989] Bertsekas D.P., Tsitsiklis J.N., Parallel and Distributed Computation: Numerical Methods, Prentice-Hall, Engelwood Cliffs, NJ, 1989

[Gershwin 1987] Gershwin S.B., "A Hierarchical Framework for Discrete Event Scheduling in Manufacturing Systems", in Varaiya P., Kurzhanski A.B., ed., Discrete Event Systems: Models and Applications, Lecture Notes in Control and Information Sciences #103, pp. 197-216, Springer Verlag, NY, 1987.

[Hopcroft and Ullman 1979] Hopcroft J.E., Ullman J.D., Introduction to Automata Theory, Languages and Computation, Addison-Wesley, Mass., 1979.

[Ichikawa and Hiraishi 1987] Ichikawa A., Hiraishi K., "Analysis and Control of Discrete Event Systems Represented with Petri Nets", in Varaiya P., Kurzhanski A.B., ed., Discrete Event Systems: Models and Applications, Lecture Notes in Control and Information Sciences #103, pp. 115-134, Springer Verlag, NY, 1987.

[Knight and Passino 1989] Knight J.F., Passino K.M., "Decidability for a Temporal Logic Used in Discrete Event System Analysis", To appear in the Int. Journal of Control.

[Krogh 1987] Krogh B.H., "Controlled Petri Nets and Maximally Permissive Feedback Logic", Proc. of the 25th Allerton Conf., Univ. of Illinois at Urbana-Champaign, Sept. 1987.

[Li and Wonham 1987] Li Y., Wonham W.M., "On Supervisory Control of Real Time Discrete Event Systems", Systems Control Group Report #8707, Dept. of Electrical Eng., University of Toronto, July 1987.

[Li and Wonham 1988] Li Y., Wonham W.M., "Deadlock Issues in Supervisory Control of Discrete Event Systems", Proc. of the 22nd Annual Conf. on Information Sciences and Systems, Princeton Univ., pp. 57-63, March 16-18, 1988.

[Mesarovic, Macko, and Takahara 1970] Mesarovic M., Macko D., Takahara Y., Theory of Hierarchical Multilevel Systems, Academic Press, NY, 1970.

[Ostroff 1986] Ostroff J.S., Real-Time Computer Control of Discrete Systems Modelled by Extended State Machines: A Temporal Logic Approach, Ph.D. Dissertation, Report No. 8618, Dept. of Elect. Eng., Univ. of Toronto, Toronto Canada, Sept. 1986.

[Passino 1989] Passino K.M., Analysis and Synthesis of Discrete Event Regulator Systems, Ph.D. Dissertation, Dept. of Electrical and Computer Eng., University of Notre Dame, Notre Dame, IN, April 1989.

[Passino and Antsaklis 1988] Passino K.M., Antsaklis P.J., "Fault Detection and Identification in an Intelligent Restructurable Controller", J. of Intelligent and Robotic Systems, Vol. 1, pp. 145-161, 1988.

[Peterson 1981] Peterson J.L., Petri Net Theory and the Modeling of Systems, Prentice Hall, NJ, 1981.

K. M. Passino and P. J. Antsaklis, "Timing Characteristics of Discrete Event Systems," Control Systems Technical Note #68, Dept. of Electrical and Computer Engineering, University of Notre Dame, June 1989; Revised November 1989.

[Ramadge and Wonham 1987] Ramadge P.J., Wonham W.M., "Supervisory Control of a Class of Discrete Event Processes", SIAM J. Control and Optimization, Vol. 25, No. 1, Jan. 1987.

[Ramadge and Wonham 1989] Ramadge P.J., Wonham W.M., "The Control of Discrete Event Systems", Proceedings of the IEEE, Vol. 77, No. 1, pp. 81-98, Jan. 1989.

[Sain 1981] Sain M.K., Introduction to Algebraic System Theory, Academic Press, NY, 1981.

[Saridis 1983] Saridis G.N., "Intelligent Robot Control", IEEE Trans. on Automatic Control, Vol. AC-28, pp. 547-556, 1983.

[Valavanis 1986] Valavanis K.P., A Mathematical Formulation for the Analytical Design of Intelligent Machines, Ph.D. Dissertation, Dept. of Electrical and Computer Eng., Rensselaer Polytechnic Inst., Troy, NY, Nov. 1986.

[Zhong and Wonham 1988] Zhong H., Wonham W.M., "On the Hierarchical Control of Discrete-Event Systems", Proc. of the 1988 Conf. on Information Sciences and Systems, Princeton, NJ, March 1988.

[Zhong and Wonham 1989] Zhong H., Wonham W.M., "Hierarchical Control of Discrete Event Systems: Computation and Examples", Proc. of the Allerton Conf. on Communication, Control, and Computing, Univ. of Illinois, Urbana, Sept. 1989.