

Diagnosis of Cyclic Discrete-Event Systems Using Active Acquisition of Information

David Thorsley and Demosthenis Teneketzis

Abstract—This paper extends the active acquisition of information approach developed in [1] from the case of acyclic, timed automata to the more general case of cyclic, asynchronous automata. Conditions for the existence of optimal solutions at finite cost are presented for both logical and stochastic systems. The information state method developed in the previous paper is reduced to a “diagnoser state” method wherein actions are computed for each potential set of states, as opposed to each potential set of strings. After developing a method of finding an optimal policy, a limited lookahead algorithm is presented to produce a suboptimal solution with less intensive computation.

I. INTRODUCTION

In complex systems such as communications networks, manufacturing processes, and queueing systems, an important problem is the detection and isolation of failures. One approach to failure detection in these systems involves modelling them as discrete event systems (DES) and then verifying if these DES have the property of diagnosability (for an overview of this approach, see [2]).

In recent years, there has been interest in studying variations of the basic diagnosability problem for logical DES formulated in [3]. One area of extension has been to the realm of stochastic DES [4], [5], while another related problem is the sensor selection problem [6]–[8], where the objective is to find the minimal sets of sensors under which diagnosability is preserved when these sensors are activated for the duration of the discrete-event process.

In problems such as diagnosability and sensor selection, finding solutions to the problem depends on which events are *observable*, i.e., events for which we have a sensor to detect their occurrence. In the typical formulation of these problems, however, the term observable is a misnomer: not only does it indicate that the occurrence of an event *can* be detected by a sensor, it also indicates that the occurrence of an event *will* be detected.

If we have the ability to choose whether or not to observe an occurrence of an event, we have a problem of measurement scheduling. Such a problem is very important for situations where the act of measurement incurs a cost in terms of money or energy. If we have a wireless sensor is a network, the act of transmitting data involves using some of the small amount of energy available to the sensor and some of the bandwidth in the network. In these situations, we cannot simply purchase a sensor at the start of the process

and let it run for the duration; instead we must use the sensor selectively, and we capture this requirement by allowing a small cost to incur each time a sensor is used.

Sensor selection problems have been studied for many classes of systems outside DES, including centralized and decentralized linear stochastic systems (e.g., [9]–[12]), communication networks, [13], [14] and operations research [15]. In the context of DES, we call the class of problems related to finding a minimal-cost observation policy as the “active acquisition of information” problem (or simply active acquisition). Our objective is to minimize the cost of observing a finite-state machine when a cost is paid each time a sensor is activated, while preserving a diagnosability property similar to that of [3].

In a previous paper [1], we considered active acquisition for the special case of automata with acyclic, timed, structures. By restricting our attention to this smaller class of systems, the information structure and the methods for finding an optimal solution were simplified.

In both [1] and this paper, we consider a version of the problem where the decision as to what sensors are activated is made by a centralized diagnoser. In this paper we consider the active acquisition problem developed in [1] for general, cyclic, automata. After developing the required concepts in Sections II and III, we formulate and solve the problem for logical models in Sections IV and V, and develop a limited lookahead algorithm in Section VI. Sections VII through X formulate and solve the problem for stochastic models. Proofs of the theorems stated in this paper can be found in [16].

II. THEORETICAL BACKGROUND

The study of discrete-event systems has followed a path common to the research of a wide class of systems: the first problems in DES were developed for the case of perfectly observed centralized systems [17], and, as the subject matured, the set of problems was expanded to include the cases of centralized partial observation and partial decentralized information (e.g. [18]). Currently in DES, all problems can be placed into one of these three broad categories; the problem considered in this paper falls into the category of centralized partial information.

To analyze the problem of active acquisition we consider the similarities between DES and stochastic discrete-time systems and borrow a concept from the study of general systems theory, that of information state. The information state is a generalization of the concept of system state that

This research is supported in part by NSF Grants ECS-0080406 and CCR-0325571 and a grant from the Xerox University Affairs Committee.

The authors are with the Department of EECS, University of Michigan.

preserves two necessary properties of a meaningful concept of state: causality and recursion.

Definition 1: (From [19]) π_k is an *information state* for a (stochastic/deterministic) system if

- 1) (Causality) π_k is a function of y^k, u^{k-1}
- 2) (Recursion) π_{k+1} can be calculated from π_k, y_{k+1} , and u_k

where the notation u_k and y_k denote the inputs and outputs at stage k , and u^k and y^k denote the sequences of inputs and outputs from the initial stage to stage k .

A third requirement on the concept of information state not commonly appearing in the literature (for an exception see [20]) is that it should be sufficient for some purpose (e.g. input-output mapping, optimization, dynamic programming, etc.) For many centralized stochastic systems, we can define an information state as the conditional probability density function of the state x_k with respect to the observation and control sequences. Such an information state is sufficient for performance evaluation, e.g., optimization.

Below we illustrate the concept of information state in DES.

A. Information States in Discrete Event Systems

Consider a DES modelled by an automaton $G = (X, \Sigma, \delta, x_0)$, where

- Σ is a finite set of events
- X is a finite state space
- $\delta : X \times \Sigma \rightarrow X$ is the partial transition function
- $x_0 \in X$ is the initial state

The DES G is observed through the detection of sequences of events. Of the events in the set Σ , some are observable and others are unobservable. If we wish to know the state of G , we may not be able to do so because our observation of the system is incomplete. Thus, it may be advantageous to define an information state in the context of DES, but in order to do so we must consider the specific problem to which we wish to find a solution.

Consider a partially observed DES where certain unobservable events are classified as failure events, i.e., we define a set $\Sigma_f \subseteq \Sigma_{uo}$. For ease of notation, we assume all failures are of the same type. Our objective is to determine if such a DES is diagnosable, that is, if it is possible to detect occurrences of a failure after at most a finite delay. A solution to this diagnosability problem was first proposed in [3]. The exact conditions for diagnosability are not germane to the discussion of information states; however, the structure of the diagnoser used to determine them is.

We define a pair of failure labels: N for normal and F for failed. Using these failure labels we define a set of possible diagnoser states as:

$$Q_o = 2^{X_o \times \{N, F\}} \quad (1)$$

where X_o denotes those states in X that are reachable via an observable event. A state of the diagnoser is thus a set of labelled states of the original system.

A diagnoser for G is the finite state machine $G_d = (Q_d, \Sigma_o, \delta_d, q_0)$, where

- $Q_d \subseteq Q_o$ is the set of reachable diagnoser states
- Σ_o is the set of observable events in G
- δ_d is the partial transition function of the diagnoser
- $q_0 = \{x_0, N\}$ is the initial state of the diagnoser

The diagnoser transition function δ_d is determined from the transition function of the system and from the *label propagation function* $LP : X_o \times \Delta \times \Sigma^* \rightarrow \Delta$

$$LP(x, l, s) = \begin{cases} \{N\} & \text{if } l = \{N\} \wedge [\Sigma_f \not\subseteq s] \\ \{F\} & \text{otherwise} \end{cases} \quad (2)$$

The label propagation function keeps track of those failure events that have occurred along a particular string. From δ and LP , the diagnoser transition function can be written as:

$$\delta_d(q, \sigma) = \bigcup_{(x, l) \in q} \bigcup_{s \in L_\sigma(G, x)} \{(\delta(x, s), LP(x, l, s))\} \quad (3)$$

where $L_\sigma(G, x)$ denotes the set of string of the from $u\sigma$ that are feasible from x , where $u \in \Sigma_{uo}^*$.

The diagnoser states defined using the above construction satisfy the definition of an information state, as 1) the diagnoser state at stage k can be calculated using the initial diagnoser state π_0 , the sequence of observations s , and δ_d ; and 2) the diagnoser state at stage $k + 1$ can be calculated from π_k , an observable event σ , and δ_d .

$$\pi_k = \delta_d(\pi_0, s) \quad (4)$$

$$\pi_{k+1} = \delta_d(\pi_k, \sigma) \quad (5)$$

The set of all possible information states (or *information space*) is Q_o . Since the observation policy in this example is fixed, only those information states in the subset Q_d are reachable.

The concept of information state defined by the diagnoser state is sufficient to solve the diagnosis problem for DES proposed in [3]. Different variations on the diagnosis problem, such as safe diagnosability [21], require different conceptions of information state to capture the unique features of the problems.

The space of information states Q_d generated in this example is insufficient for the active acquisition problem we study in this paper. In the active acquisition problem, we cannot compute the reachable state space Q_d of the diagnoser until we have determined the observation policy; as our objective is to determine an optimal observation policy, it follows that we cannot use a diagnoser approach directly to solve this problem. In this paper we present methods of constructing information spaces that are appropriate for different formulations of the problem.

III. METHODS FOR ACYCLIC SYSTEMS

We now review some results for acyclic systems (see [1], [16]) that form the basis for the results that we develop in this paper for cyclic systems and are especially useful in defining limited lookahead algorithms.

We first assign a cost $\nu : \Sigma_o \rightarrow [0, \infty)$ to each observable event. If $\nu(\sigma) = 0$, then σ is said to be *freely observable*; otherwise σ is *costly observable*. The set of all costly observable events are denoted by Σ_{co} . An observation action $g(\pi)$

at information state π according to the scheduling policy g specifies the set of events that are the first observable events along any continuation from π . The cost of an observation action $g(\pi)$ at an information state π is given by $c(g(\pi)) = \sum_{\sigma \in g(\pi)} \nu(\sigma)$.

The difficulty in the active acquisition problem is determining how to systematically approach how information regarding the system behaviors evolves as events are observed. To address this difficulty, we use a maximal σ -field approach. This approach was initially proposed in [22] in the context of general informationally decentralized systems and was further used in [23]–[26].

Suppose that the length of the longest trace in the acyclic automaton is T . We wish to define a sequence of σ -fields $\mathcal{F}_0, \dots, \mathcal{F}_T$ such that all possible information states reachable under any observation policy are elements of these σ -fields. To create these σ -fields, we partition the language $\mathcal{L}(G)$ into sets of strings X_n , where every element of X_n is a set of strings that are identical under projection and such that at most n more observations are possible if we choose to observe all observable events. We define the projection P and the inverse projection P_L^{-1} in the usual manner [27].

Specifically, for $n = 0 \dots T$, we define

$$X_n = \{s \in P_L^{-1}[P(\mathcal{L}(G))] : \max_{t \in P(\mathcal{L}(G))/P(s)} \|t\| = n\} \quad (6)$$

and we define the sequence of σ -fields as follows:

$$\mathcal{F}_t = \sigma(\cup_{n=0}^t X_n) \quad (7)$$

Each element of the partition that generates these σ -fields is the “finest,” “maximal” information available to the diagnoser. Such information is available if all observations are made at all times.

As events are observed, the *information state transition function* defines how the information state changes, dependent on the observation action we have chosen. For a given observation policy g , the information state transition function is defined as $\hat{\delta}_g : \mathcal{F}_t \times \Sigma_{g(\pi),obs} \cup \epsilon \rightarrow \mathcal{F}_t$:

$$\hat{\delta}_g(\pi, \sigma) := \{st\sigma : s \in \pi \wedge t \in (\Sigma_{g(\pi),unobs})^*\} \quad (8)$$

$$\hat{\delta}_g(\pi, \epsilon) := \{st : s \in \pi \wedge t \in \Sigma_{g(\pi),unobs}^* \wedge \Gamma(\delta(x_0, st)) = \emptyset\} \quad (9)$$

where $\Sigma_{g(\pi),obs}^*$ and $\Sigma_{g(\pi),unobs}^*$ denote the set of events that are observable and unobservable, respectively, under the observation action $g(\pi)$. $\Gamma(x)$ denotes the set of events that are feasible from the state $x \in X$.

We can solve for an optimal observation policy by backwards induction. We begin by defining a cost for all $\pi \in \mathcal{F}_0$ as follows.

$$V(\pi) = \begin{cases} 0 & \text{if } \pi \text{ is } N\text{-certain} \\ 0 & \text{if } \pi \text{ is } F\text{-certain} \\ \infty & \text{otherwise} \end{cases} \quad (10)$$

where an information state is *F-certain* if $f \in s$ for all $s \in \pi$ and *N-certain* if $f \notin s$ for all $s \in \pi$ [3].

For every information state π not in \mathcal{F}_0 , we can calculate the optimal cost and an optimal observation action by solving

the following dynamic program:

$$V(\pi) = \min_{u \in 2^{\Sigma_{co}}} \{c_u + \max_{\sigma \in \Sigma \cup \epsilon} V(\hat{\delta}_u(\pi, \sigma))\} \quad (11)$$

We can solve this dynamic program sequentially, first for elements π in $\mathcal{F}_1 - \mathcal{F}_0$, then for elements in $\mathcal{F}_2 - \mathcal{F}_1$, and so on.

IV. MODELLING FORMALISM

We now formulate the active acquisition of information problem for diagnosis of cyclic systems. For ease of notation we will restrict attention to the case where there is only one failure type; the results below can be extended to the case of multiple failure types.

Definition 2: An information state π is *N-safe* if $f \notin s$ and $f \notin L/s$ for all $s \in \pi$.

Definition 3: An information state π is *safe* if π is *F-certain* or *N-safe*.

If the system is in a safe information state, we need not make any more observations since we are certain about the failure mode in the current information state and in all future states. If the information state is unsafe, we must choose an action that ensures that another event will eventually be observed or else we may never diagnose the failure.

Definition 4: An information state π is *non-diagnosable* if $\exists M \in \mathbb{N}$ such that for all $n \geq M$, $\exists t \in L/\pi$ such that $\|t\| = n$ and the information state obtained by implementing any policy g along t is uncertain.

Definition 5: A language $\mathcal{L}(G)$ is *diagnosed* by an observation policy g if, for all $s \in \mathcal{L}(G)$, the information state reached by implementing g along s is never non-diagnosable.

Definition 6: Let H denote the set of all policies that diagnose $\mathcal{L}(G)$. The language $\mathcal{L}(G)$ is *diagnosable* if H is non-empty, i.e., if there exists a policy that diagnoses $\mathcal{L}(G)$.

Define the performance criterion:

$$J(g) = \left\{ \sup_{s \in \mathcal{L}(G)} c^g(s) \right\} \quad (12)$$

where $c^g(s)$ denotes the cost of implementing policy g along the string s .

The active acquisition of information problem, henceforth also called the cyclic diagnosis problem, is defined as follows.

Problem CD. Find a policy $g^* \in H$ such that

$$J(g^*) = \inf\{J(g) | g \in H\} \text{ and } J(g^*) < \infty \quad (13)$$

A. Solution Existence

If we assume all observable events have a non-zero cost, solution existence for Problem CD can be determined using the following definition and theorem.

Definition 7: A language $\mathcal{L}(G)$ is strictly logically diagnosable with respect to Σ_o and Σ_f if:

$$(\exists N \in \mathbb{N})(n > N \Rightarrow \hat{D}^N(s) = 1 \vee D^F(s) = 1) \quad (14)$$

where the functions \hat{D}^N and D^F are defined as:

$$\hat{D}^N(s) = \begin{cases} 1 & \text{if } P_L^{-1}[P(s)] \text{ is } N\text{-safe} \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

$$D^F(s) = \begin{cases} 1 & \text{if } P_L^{-1}[P(s)] \text{ is } F\text{-certain} \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

Theorem 1: A finite-cost observation policy exists if and only if $\mathcal{L}(G)$ is strictly logically diagnosable with respect to Σ_o and Σ_f .

The condition of strict logical diagnosability is too severe for most problems as it disallows the possibility of the system running in a normal, “unsafe,” state for an indefinitely long time. Were a system to run in such a state indefinitely, a diagnosis at infinite observation cost would be incurred; however, the number of events required to occur for this cost to be incurred would also be infinite. Therefore, it would be more realistic to find a criterion for solution existence closer to the pre-existing concepts of diagnosability [3]

Definition 8: A language $\mathcal{L}(G)$ is *logically diagnosable* with respect to Σ_o, Σ_f if

$$(\exists n \in \mathbb{N})[\forall s \in \Psi(\Sigma_f)](\forall t \in L/s)[\|t\| \geq n \Rightarrow D^F(st) = 1] \quad (17)$$

where $\Psi(\Sigma_f)$ denotes the set of strings that end in a failure event. To formulate the problem so that finite-cost solution existence corresponds to the notion of logical existence, consider a performance criterion where future costs are discounted at a rate $\beta < 1$:

$$J_\beta(g) = \left\{ \max_{s \in \mathcal{L}(G)} \sum_{t=0}^{\|s\|} \beta^t c_t^g(s) \right\} \quad (18)$$

The discounted active acquisition problem, or discounted cyclic diagnosis problem, is defined as follows.

Problem CD-D. Find a policy $g^* \in H$ such that

$$J_\beta(g^*) = \inf\{J_\beta(g) \mid g \in H\} \quad (19)$$

Theorem 2: A language $\mathcal{L}(G)$ is diagnosable at finite discounted cost if and only if it is logically diagnosable with respect to Σ_o and Σ_f .

V. SOLUTION METHODS

A cyclic automaton generates an infinite number of *string-based* information states in the σ -field $2^{P_L^{-1}[P(\mathcal{L}(G))]}$. In order to derive an optimal policy as we had done in the case of acyclic automata, we reduce the string-based information states to diagnoser states, as the set of diagnoser states is guaranteed to be finite [3].

Recall from Section II that a diagnoser state is an element of the set $Q_o = 2^{X_o \times \{N, F\}}$. For each information state π in $2^{\mathcal{L}(G)}$, the diagnoser state associated with π can be computed by the function $q : 2^{\mathcal{L}(G)} \rightarrow Q_o$

$$q(\pi) = \bigcup_{s \in \pi} (\delta(x_o, s), LP(x_o, s)) \quad (20)$$

This mapping allows the infinite set of string-based information states in $2^{\mathcal{L}(G)}$ to be reduced to a finite set of diagnoser states, or *state-based information states*. We can calculate optimal policies using diagnoser states instead of string-based information states as a result of the following theorem.

Theorem 3: If multiple information states map to the same diagnoser state, the same sequence of observation actions is optimal for any string after that diagnoser state.

The reduction of information states to diagnoser states ensures that an optimal policy need only be calculated for a finite number of information states. However, the reduction of set of strings to diagnoser states sacrifices the sequentiality inherent in the strings; there is no inherent “filtration” of diagnoser states that we can use as we have in the case of acyclic systems.

Nevertheless, there are certain diagnoser states for which we can assign a cost *a priori* just as we assign costs to information states in the final maximal σ -field. If a state is safe, we are sure that no more observations are needed after reaching such a state and can assign zero cost to such a state. Furthermore, we can test all remaining diagnoser states to see if they are non-diagnosable [28] and assign infinite cost in that case.

For all $q_d \in Q_o$, define

$$V(q_d) = \begin{cases} 0 & \text{if } q_d \text{ is safe} \\ \infty & \text{if } q_d \text{ is non-diagnosable} \end{cases} \quad (21)$$

We state what it means for a state-based information state to be diagnosable in the following definition.

Definition 9: A state-based information state q is *diagnosable* if the language generated by the automaton $G' = (X \cup x', \Sigma \cup \{f, n\}, \delta', x')$ is diagnosable, where:

$$\delta'(x', f) = x \quad \text{if } (x, F) \in q \quad (22)$$

$$\delta'(x', n) = x \quad \text{if } (x, N) \in q \quad (23)$$

$$\delta'(x, \sigma) = \delta(x, \sigma) \quad \text{if } x \neq x' \quad (24)$$

In short, a state-based information state is diagnosable if the automaton is diagnosable when the initial information state is not x_o , but instead the diagnoser state specified by the information state. To apply standard diagnosability results, we append a new initial state to the automaton G and add unobservable transitions to this state that bear the failure labels associated with each component of the diagnoser state.

In order to determine whether a particular diagnoser state q_d has a zero or infinite cost, we need only test these conditions for diagnoser states consisting of at most two components. If q_d has two components and is not diagnosable, any diagnoser state that is a superset of q_d will also be non-diagnosable. If q_d consists of one component and is safe, than any diagnoser state that is the union of one-component safe states having the same label is also safe; such a diagnoser state has zero cost.

For Problem CD, the costs of the remaining diagnoser states q_d can be determined using the following equation:

$$V(q_d) = \min_{u \in 2^{\Sigma_{co}}} \{c_u + \max_{\sigma \in \Sigma_o} V(\hat{\delta}_u(q_d, \sigma))\} \quad (25)$$

For Problem CD-D, the equation that needs to be solved is given by

$$V(q_d) = \min_{u \in 2^{\Sigma_{co}}} \{c_u + \max_{\sigma \in \Sigma_o} \beta^t V(\hat{\delta}_{g(q_d)}(q_d, \sigma))\} \quad (26)$$

where the exponent t is defined as:

$$t = \min_{u \in \Sigma_{g(q_d)}, u \text{ nobis}} \{\|u\sigma\| : u\sigma \in L/\pi\} \quad (27)$$

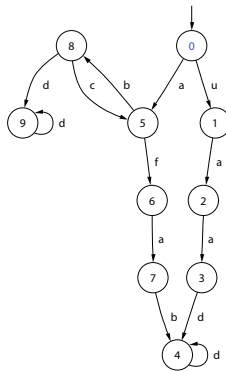


Fig. 1. An automaton used to illustrate the active acquisition method.

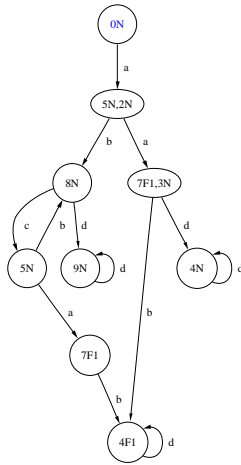


Fig. 2. The diagnoser of the system in Figure 1 obtained when all observable events are observed at all stages.

Future costs in Equation (26) are thus discounted according to the minimum number of events that may have occurred between the current observation and the next observation.

The dynamic programming equations that solve Problems CD and CD-D are sets of algebraic equations that in general must be solved simultaneously for all diagnoser states. Such equations appear in the literature as characteristics of the free-time problem in stochastics (cf. Chapter 4 of [29]) and infinite horizon expected discounted-cost problems [19].

A. Example

Figure 1 shows an example of a cyclic automaton, where the costs of each observable event are given by $\nu(a) = 1$, $\nu(b) = 2$, $\nu(c) = 3$, and $\nu(d) = 4$. The diagnoser obtained when all observable events are always observed is shown in Figure 2. The results of the two-component diagnosability tests are shown in Table I.

Since the diagnoser state $\{(3, N), (4, F)\}$ has infinite cost, any diagnoser state containing both $(3, N)$ and $(4, F)$ will also have infinite cost. Also, if two diagnoser states have zero cost and the same label, their union will also have zero cost, e.g. since $\{(3, N), (4, N)\}$ and $\{(4, N), (8, N)\}$ both have zero cost and bear only the label N , the diagnoser state $\{(3, N), (4, N), (8, N)\}$ also has zero cost.

	4F	7F	0N	2N	3N	4N	5N	8N
7F	0	—	—	—	—	—	—	—
0N	?	?	—	—	—	—	—	—
2N	?	?	?	—	—	—	—	—
3N	∞	?	?	0	—	—	—	—
4N	∞	?	?	0	0	—	—	—
5N	?	∞	∞	?	?	?	—	—
8N	∞	?	?	0	0	0	?	—
9N	∞	?	?	0	0	0	?	0

TABLE I

RESULTS OF DIAGNOSABILITY TESTS FOR REACHABLE TWO-COMPONENT DIAGNOSER STATES.

For the system in Figure 1, a finite-cost solution exists for Problem CD-D but not for problem CD. To see this, consider the cost of the information state $q = \{(8, N)\}$ under problem CD-D.

In the information state $q = \{(8, N)\}$, only the two actions to observe the event sets $\{c, d\}$ and $\{a, b, d\}$ and actions that are supersets of those actions are admissible in that they prevent the system from entering a non-diagnosable state. The equation to find an optimal action for q is therefore:

$$V(8N) = \min\{c+d+\beta V(5N), a+b+d+\beta^2 V(8N)\} \quad (28)$$

We now need to consider the cost of the information state $\{5, N\}$. Using the same arguments as above, the only two actions we need to consider are $\{a, b\}$ and $\{a, c, d\}$. The equation to find an optimal action at $\{5, N\}$ is:

$$V(5N) = \min\{a+b+\beta V(8N), a+c+d+\beta^2 V(5N)\} \quad (29)$$

If we solve these equations simultaneously, we find that the optimal action at $\{8, N\}$ is $\{a, b, d\}$ and the optimal action at $\{5, N\}$ is $\{a, b\}$. The cost of the information state $\{8, N\}$ is then:

$$V(8N) = \frac{7}{1-\beta^2} \quad (30)$$

If $\beta < 1$, the cost of $\{8, N\}$ is finite. However, if we consider problem CD, β is equal to exactly one and the cost of diagnosing the failure from this information state becomes infinite. The loop between states 5 and 8 means that it is possible for an arbitrarily large number of observations to be necessary, thus the worst-case undiscounted observation cost must be infinite.

VI. LIMITED LOOKAHEAD IN CYCLIC SYSTEMS

To overcome the difficulties inherent in cyclic systems we consider a limited lookahead method similar to the one proposed in [1] for acyclic timed automata. By restricting attention to a finite lookahead horizon, we no longer need to make the switch from string-based to state-based information states, as the limited lookahead ensures that only a finite number of strings are considered at each stage.

However, in applying the limited lookahead method to cyclic systems, we must take note of a fine distinction that does not appear in acyclic automata; namely, the distinction between *preserving the property of diagnosability* and the *actual act of diagnosing the failure*. To see this difference, consider the example in Figure 3 and suppose that $\Sigma_o =$

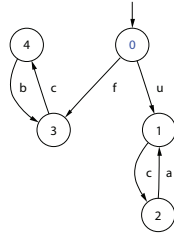


Fig. 3. An automaton where applying the acyclic limited lookahead approach directly results in the failure never being diagnosed.

$\{a, b, c\}$, $\Sigma_f = \{f\}$, $\Sigma_{co} = \{a, b\}$, and that the limited lookahead horizon is $T' = 2$. Suppose we apply the limited lookahead algorithm for acyclic automata defined in [1] without modification. At each stage, the locally optimal action is always to observe only c , as it will always be possible to pay to observe a and b beyond the lookahead horizon. Such a policy ensures that the failure event is always diagnosable, but the actual diagnosis can be put off indefinitely. The “procrastination” characteristic described for the acyclic timed model is no longer held in check by the existence of a final, finite, deadline for diagnosis.

In order to ensure diagnoses are made in a timely fashion, we consider a surrogate problem wherein we introduce a penalty for the delay in diagnosis occurring in uncertain information states. The delay in diagnosis for an uncertain information state π is defined as:

$$dly(\pi) = \max_{s \in \pi} (\|t\| : s = uft) \quad (31)$$

We require the delay penalty function $R : \mathbb{N} \rightarrow \mathbb{R}^+$ to have the following properties: 1) R is non-decreasing in \mathbb{N} , and 2) $\exists n \in \mathbb{N}$ such that $R(n) \geq c(\Sigma_{co})$. The first condition ensures that the penalty for delaying a diagnosis increases as the delay increases, while the second ensures that if the diagnosis has been delayed a sufficient length of time, it becomes optimal to make whatever observations are necessary to complete the diagnosis.

The costs of terminal information states associated with the lookahead horizon T are specified as follows. If an information state is diagnosable when all possible observations are made beyond the horizon T , then the cost of that information state is a function of the diagnosis delay defined by (31).

To construct the σ -fields used in the limited lookahead algorithm, we first create the automaton $G_{T'}$ which generates all strings in $\mathcal{L}(G)$ of length T' or less. For $n = 0 \dots T$, we define a sequence of partitions using the method for acyclic untimed automata described in Section III:

$$X'_n = \{s \in P_L^{-1}[P(G_{T'})] : \max_{t \in P(G_{T'})/P(s)} \|t\| = n\} \quad (32)$$

and we define the sequence of σ -fields as follows:

$$\mathcal{F}'_t = \sigma(\cup_{n=0}^t X'_n) \quad (33)$$

We assign a cost to all information states in \mathcal{F}'_0 as follows:

$$V'(\pi) = \begin{cases} 0 & \text{if } \pi \text{ is F-certain} \\ R(dly(\pi)) & \text{if } \pi \text{ is diagnosable} \\ \infty & \text{if } \pi \text{ is non-diagnosable} \end{cases} \quad (34)$$

We then determine an optimal observation action for every information state π in \mathcal{F}'_T using the following equation:

$$V'(\pi) = R(dly(\pi)) + \min_{g(\pi) \in 2^{\Sigma_{co}}} \{c_g(\pi) + \max_{\sigma \in \Sigma \cup \epsilon} V'(\hat{\delta}_{g(\pi)}(\pi, \sigma))\} \quad (35)$$

Upon determining an optimal action $g(\epsilon)$, we implement that action; when an event is observed, we generate a new information state π' , construct a new sequence of σ -fields as in (33), and proceed as before using equations (34)-(35) with π' as the initial information state.

In the example in Figure 3, by considering a diagnosis delay function R with the properties stated above, eventually the cost of delaying the diagnosis will exceed the cost of making the necessary observations, and thus we can ensure the failure will eventually be detected.

VII. PROBLEM FORMULATION FOR STOCHASTIC, CYCLIC AUTOMATA

In a manner analogous to the section on timed, acyclic automata, we now consider the active acquisition of information problem for stochastic, cyclic automata. The stochastic automata we consider are constructed by assigning probabilities to each transition of a logical automaton.

We start by restating Definitions 2-6 for the stochastic framework. We will consider the case where a diagnosis is made if the probability of being certain about the failure information is greater than $\alpha \leq 1$.

Definition 10: An information state π is *almost-F-certain* if $\Pr(s : f \in s \mid s \in \pi) > \alpha$.

Definition 11: An information state π is *almost-N-safe* if $\Pr(s : f \notin s \wedge f \notin L/s \mid s \in \pi) > \alpha$.

Definition 12: An information state π is *almost safe* if π is almost-F-certain or almost-N-safe.

Definition 13: An information state π is *uncertain* if $\alpha > \Pr(s : f \in s \mid s \in \pi) > 1 - \alpha$.

Definition 14: An information state π is *non-diagnosable* if $\exists N \in \mathbb{N}$ such that for all $n \geq N$, $\exists t \in L/\pi$ such that $\|t\| = n$ and the information state obtained by implementing any policy g along t is uncertain in the sense of Definition 13.

Definition 15: A language $\mathcal{L}(G)$ is *diagnosed* by an observation policy g if, for all $s \in \mathcal{L}(G)$, the information state π reached by implementing g along s is never non-diagnosable.

Definition 16: Let H denote the set of all policies that diagnose $\mathcal{L}(G)$. The language $\mathcal{L}(G)$ is *diagnosable* if H is non-empty, i.e., if there exists a policy that surely diagnoses $\mathcal{L}(G)$.

For stochastic automata we consider the expected cost instead of the worst-case cost. Define the performance criterion:

$$J(g) = \{E(c^g(s))\} \quad (36)$$

where $c^g(s)$ denotes the cost of implementing policy g along the string s . The performance criterion is thus the expected total cost of policy g .

The stochastic cyclic almost sure diagnosis active acquisition problem is defined as follows.

Problem SCASD. Find a policy $g^* \in H$ such that

$$J(g^*) = \inf(J(g)|g \in H) < \infty \quad (37)$$

VIII. SOLUTION EXISTENCE IN THE CYCLIC, STOCHASTIC CASE

Just as in the case of logical systems, we first consider conditions necessary and sufficient to ensure that a language can be diagnosed at finite cost. To find such conditions, we consider the previous work on diagnosability of stochastic discrete-event systems [5].

A. Review of Stochastic Diagnosability

The notions of stochastic diagnosability introduced in [5] replace the sure statements of the definition of diagnosability for logical automata in [3] with probabilistic almost sure statements. Of the two definitions presented in [5], the stricter is A -diagnosability.

Definition 17: (A -diagnosability) A live, prefix-closed language L is A -diagnosable with respect to a projection P and a set of transition probabilities p if

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N})(\forall s \in \Psi(\Sigma_{f_i}) \wedge n \geq N) \{ \Pr(t : D^F(st) = 0 \mid t \in L/s \wedge \|t\| = n) < \epsilon \} \quad (38)$$

where the diagnosability condition function D^F is as in Equation (16).

If a system is A -diagnosable, when a failure occurs, the probability of a continuation that does not allow the failure to be diagnosed approaches zero as the length of the continuation approaches infinity. However, we still need to be logically certain that a failure has occurred in order to call it diagnosed. In the second definition, AA -diagnosability, we weaken the requirement necessary to diagnose a failure.

Definition 18: (AA -diagnosability) A live, prefix-closed language L is AA -diagnosable with respect to a projection P and a transition probability function p if

$$(\forall \epsilon > 0 \wedge \forall \alpha < 1)(\exists N \in \mathbb{N})(\forall s \in \Psi(\Sigma_{f_i}) \wedge n \geq N) \{ \Pr(t : D_\alpha^F(st) = 0 \mid t \in L/s \wedge \|t\| = n) < \epsilon \} \quad (39)$$

where the diagnosability condition function D_α is:

$$D_\alpha^F(st) = \begin{cases} 1 & \text{if } \Pr(\omega : \Sigma_f \in \omega \mid \omega \in P_L^{-1}[P(st)]) > \alpha \\ 0 & \text{otherwise} \end{cases} \quad (40)$$

Thus a system is AA -diagnosable if almost every continuation of a certain length after a failure event leads to a state where we are almost certain that the failure has occurred with probability greater than α , for any α arbitrarily close to, but not equal to, one. Conditions necessary and sufficient to confirm A -diagnosability and sufficient to confirm AA -diagnosability are given in [5].

B. Strict- AA -diagnosability; Solution Existence for Problem SCASD

For an optimal finite-cost solution to exist when $\alpha < 1$, we wish to ensure that a diagnosis is almost surely made in a finite amount of time. We therefore define strict- AA -diagnosability, which requires that a diagnosis of either “normal” or “failed” is almost surely made.

Definition 19: A language is strictly- AA -diagnosable if

$$(\forall \epsilon > 0 \wedge \forall \alpha < 1)(\exists N \in \mathbb{N})(\forall n > N) \Pr(s : D_\alpha^F(s) = 0 \wedge \hat{D}_\alpha^N(s) = 0 \mid \|s\| = n) < \epsilon \quad (41)$$

where the function \hat{D}_α^N is defined analogously to D_α^F as:

$$\hat{D}_\alpha^N(st) = \begin{cases} 1 & \text{if } \Pr(\omega : \Sigma_f \notin \omega L/\omega \mid \omega \in P_L^{-1}[P(st)]) > \alpha \\ 0 & \text{otherwise} \end{cases} \quad (42)$$

Theorem 4: A language is strictly- AA -diagnosable if and only if it is AA -diagnosable.

While it is fairly clear that strict- AA -diagnosability should imply AA -diagnosability, the opposite implication is not intuitively obvious; the idea behind this result is as follows. If a system is AA -diagnosable and no failure occurs, the probability that the system does not reach a safe normal state becomes arbitrarily small in the long run, as the set of unsafe normal states is transient. Thus if no failure occurs, we will almost surely eventually diagnose that the system is in normal operation, and since the system is AA -diagnosable, we will almost surely eventually diagnose any failure events.

Having demonstrated that strict- AA -diagnosability and AA -diagnosability are equivalent, we can now state the conditions under which Problem SCASD has a solution when $\alpha < 1$.

Theorem 5: A language is diagnosable for all $\alpha < 1$ at finite expected cost if and only if the language is AA -diagnosable when all events in Σ_o are observed.

Using the results from [5], we can now state a sufficient condition for a stochastic automaton to be diagnosable with finite expected cost.

Corollary 1: A language is diagnosable for all $\alpha < 1$ at finite expected cost if the set of recurrent components in each logical element of its stochastic diagnoser is certain.

For the case when $\alpha = 1$, we define strict- A -diagnosability in manner similar to Definition 19.

Definition 20: A language is strictly- A -diagnosable if

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N})(\forall n > N) \Pr(s : D^F(s) = 0 \wedge \hat{D}^N(s) = 0 \mid \|s\| = n) < \epsilon \quad (43)$$

Strict- A -diagnosability is not equivalent to A -diagnosability; therefore the result corresponding to Theorem 5 when $\alpha = 1$ must be stated as:

Theorem 6: A language is diagnosable at finite expected cost for $\alpha = 1$ if and only if the language is strictly- A -diagnosable when all events in Σ_o are observed.

IX. COMMENTS ON SOLUTION METHODS FOR STOCHASTIC AUTOMATA

In general, the information state of a partially observed stochastic automaton is an element of an infinite space [4]. Thus, we cannot perform a reduction from an infinite set of string-based information states to the finite set of state-based information states, as in the logical case.

When $\alpha = 1$, we can assign costs to certain information states as follows:

$$V(\pi) = \begin{cases} 0 & \text{if } \pi \text{ is safe} \\ \infty & \text{if } \pi \text{ is not strictly-}A\text{-diagnosable} \end{cases} \quad (44)$$

An information state $\pi = s_1 + s_2 + \dots + s_n$ is defined to be strict- A -diagnosable if the language $L_\pi := \hat{P}(s_1)t_1 + \hat{P}(s_2)t_2 + \dots + \hat{P}(s_n)t_n$ is strictly- A -diagnosable, where \hat{P} is the projection of Σ onto Σ_{uo} .

An optimal observation policy can be computed by solving for every information state π the dynamic programming equation:

$$V(\pi) = \min_{u \in 2^{\Sigma_{co}}} \{c_u + \sum_{\sigma \in \Sigma_{u,obs}} V(\hat{\delta}_u(\pi, \sigma))P(\sigma | \pi, u)\} \quad (45)$$

where $P(\sigma | \pi, u)$ indicates the probability that the next observed event is σ , given the current information state of the system is state π . This probability is in general dependent on the observation policy.

For $\alpha < 1$, we initialize the dynamic program by assigning costs as follows:

$$V(\pi) = \begin{cases} 0 & \text{if } \pi \text{ is almost safe} \\ \infty & \text{if } \pi \text{ is not } AA\text{-diagnosable} \end{cases} \quad (46)$$

where an information state π is defined to be AA -diagnosable if the language L_π is AA -diagnosable.

While finding an optimal observation policy requires determining optimal actions for an infinite set of information states simultaneously, we can find a suboptimal policy by modifying the limited lookahead algorithm of Section VI to minimize the expected observation cost instead of the maximal observation cost. A full discussion of computational features of logical and stochastic problems can be found in [16].

X. DISCUSSION

This paper demonstrates how the active acquisition of information method for DES introduced in [1] can be extended to general, cyclic, automata that are either logical or stochastic. In the logical case, the size of the information space can be bounded by translating string-based information states into diagnoser states. In both cases, the use of limited lookahead algorithm can produce a suboptimal solution by applying the techniques used for acyclic systems.

Future work on this problem lies mainly in the area of developing more efficient techniques to find either optimal or suboptimal solutions that can be tailored to specific applications. By making additional assumptions on the system structure, it may be possible to derive practically useful solutions without requiring the computational intensity of the general approach developed in this paper.

REFERENCES

- [1] D. Thorsley and D. Teneketzis, "Active acquisition of information for diagnosis of discrete event systems," in *Proc. 42th Allerton Conference on Communication, Control, and Computing*, Sept. 2004.
- [2] S. Lafortune, D. Teneketzis, M. Sampath, R. Sengupta, and K. Sinnamohideen, "Failure diagnosis of dynamic systems: An approach based on discrete event systems," in *Proc. 2001 American Control Conference*, June 2001, pp. 2058–2071.
- [3] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis, "Diagnosability of discrete-event systems," *IEEE Trans. Automatic Control*, vol. 40, no. 9, pp. 1555–1575, Sept. 1995.
- [4] J. Lunze and J. Schröder, "State observation and diagnosis of discrete-event systems described by stochastic automata," *Discrete Event Dynamic Systems: Theory and Applications*, vol. 11, no. 4, pp. 319–369, 2001.
- [5] D. Thorsley and D. Teneketzis, "Diagnosability of stochastic discrete-event systems," *IEEE Trans. Automatic Control*, Apr. 2005.
- [6] R. Debouk, S. Lafortune, and D. Teneketzis, "On an optimization problem in sensor selection," *J. of Discrete Event Dynamical Systems: Theory and Applications*, vol. 12, pp. 417–445, 2002.
- [7] S. Jiang, R. Kumar, and H. Garcia, "Optimal sensor selection for discrete-event systems with partial observation," *IEEE Trans. on Systems, Man and Cybernetics, Part B*, vol. 30, no. 5, pp. 653–660, 2003.
- [8] T.-S. Yoo and S. Lafortune, "NP-completeness of sensor selection problems arising in partially observed discrete-event systems," *IEEE Trans. Automatic Control*, vol. 47, no. 9, pp. 1495–1499, Sept. 2002.
- [9] M. Athans, "On the determination of optimal costly measurement strategies for linear stochastic systems," *Automatica*, vol. 8, pp. 397–412, 1972.
- [10] H. Kushner, "On the optimum timing of observations for linear control systems with unknown initial state," *IEEE Trans. Automatic Control*, vol. 9, no. 2, pp. 144–150, Apr. 1964.
- [11] M. Khanna, "Sampling and transmission policies for controlled Markov processes with costly communication," Ph.D. dissertation, Department of Electrical Engineering, University of Toronto, 1973.
- [12] M. Andersland and D. Teneketzis, "Measurement scheduling for recursive team estimation," *J. of Optimization Theory and Applications*, vol. 89, no. 3, pp. 615–636, June 1996.
- [13] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 32, no. 2, pp. 554–568, Apr. 1996.
- [14] S. Appadwedula, V. Veeravalli, and D. Jones, "Robust and locally-optimum decentralized detection with censoring sensors," in *Proc. 5th International Conference on Information Fusion*, Annapolis, MD, USA, 2002.
- [15] X. Ding, M. Puterman, and A. Bisi, "The censored newsvendor and the optimal acquisition of information," *Operations Research*, vol. 50, pp. 517–527, May–June 2002.
- [16] D. Thorsley and D. Teneketzis, "Active acquisition of information for diagnosis and control of discrete event systems," Department of Electrical Engineering and Computer Science, University of Michigan, Tech. Rep. CGR-06-01, Jan. 2006.
- [17] P. Ramadge and W. Wonham, "The control of discrete-event systems," *Proc. IEEE*, vol. 77, no. 1, pp. 81–98, 1989.
- [18] K. Rudie and W. Wonham, "Think globally, act locally: Decentralized supervisory control," *IEEE Trans. Automatic Control*, vol. 37, no. 11, pp. 1692–1708, Nov. 1992.
- [19] P. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Englewood Cliffs, NJ: Prentice Hall, 1986.
- [20] H. Witsenhausen, "Some remarks on the concept of state," in *Directions in Large-Scale Systems: Many-Person Optimization and Decentralized Control*. Plenum Press, 1975, pp. 69–76.
- [21] A. Paoli and S. Lafortune, "Safe diagnosability of discrete event systems," in *Proc. 42st IEEE Conf. on Decision and Control*, Dec. 2003, pp. 2658–2664.
- [22] H. Witsenhausen, "On information structures, feedback and causality," *SIAM J. of Control*, vol. 9, no. 2, pp. 149–160, May 1971.
- [23] M. Andersland and D. Teneketzis, "Information structures, causality, and non-sequential stochastic control, I: design-independent properties," *SIAM J. of Control Optim.*, vol. 30, no. 6, pp. 1447–1475, Nov. 1992.

- [24] —, “Information structures, causality, and non-sequential stochastic control, II: design-dependent properties,” *SIAM J. of Control Optim.*, vol. 32, no. 6, 1994.
- [25] D. Teneketzis, “On information structures and nonsequential stochastic control,” *CWI Quarterly*, vol. 9, no. 3, pp. 241–260, 1996, special issue on Systems and Control.
- [26] D. Teneketzis and M. Andersland, “On partial order characterizations of information structures,” *Mathematics of Control, Signals and Systems*, vol. 13, pp. 277–292, 2000.
- [27] C. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*. Boston, MA: Kluwer Academic Publishers, 1999.
- [28] S. Jiang, Z. Huang, V. Chandra, and R. Kumar, “A polynomial algorithm for testing diagnosability of discrete-event systems,” *IEEE Trans. Automatic Control*, vol. 46, no. 8, pp. 1318–1320, Aug. 2001.
- [29] H. Kushner, *Introduction to Stochastic Control*. Holt, Rinehart, Winston, 1971.