Discrete Event Dynamic Systems: Theory and Applications, 12, 417–445, 2002

On an Optimization Problem in Sensor Selection*

RAMI DEBOUK† rami.debouk@gm.com
STÉPHANE LAFORTUNE stephane@eecs.umich.edu
DEMOSTHENIS TENEKETZIS teneketzis@eecs.umich.edu
Department of Electrical Engineering and Computer Science,
The University of Michigan, 1301 Beal Avenue, Ann Arbor, MI 48109-2122, U.S.A.

Abstract. We address the following sensor selection problem. We assume that a dynamic system possesses a certain property, call it Property D, when a set Γ of sensors is used. There is a cost c_A associated with each set A of sensors that is a subset of Γ. Given any set of sensors that is a subset of Γ, it is possible to determine, via a test, whether the resulting system-sensor combination possesses Property D. Each test required to check whether or not Property D holds incurs a fixed cost. For each set of sensors A that is a subset of Γ there is an *a priori* probability p_A that the test will be positive, i.e., the system-sensor combination possesses Property D. The objective is to determine a test strategy, i.e., a sequence of tests, to minimize the expected cost, associated with the tests, that is incurred until a least expensive combination of sensors that results in a system-sensor combination possessing Property D is identified. We determine conditions on the sensor costs c_A and the *a priori* probabilities p_A under which the strategy that tests combinations of sensors in increasing order of cost is optimal with respect to the aforementioned objective.

Keywords: failure diagnosis, hypothesis testing, Markovian decision problems, optimization, sensor selection

1. Introduction

We formulate and analyze an optimization problem in sensor selection that arises in failure diagnosis, detection and hypothesis testing, and is motivated by economic or energy-related considerations. Below we describe two classes of problems in the aforementioned areas that highlight the issues we are concerned with.

Consider a problem in failure diagnosis. Assume that a dynamic system is diagnosable when a set Γ of sensors is used. By diagnosable we mean that any failure of interest can be detected and isolated within a finite time after its occurrence by a prespecified failure diagnosis scheme. Let us assume that there is a cost function associated with the subsets of Γ. We wish (i) to identify a least expensive combination of sensors $A^* \subseteq \Gamma$ under which the system is diagnosable; (ii) to determine the minimum number of tests required to identify such a set A^*, assuming that we use a fixed prespecified test for diagnosability.

Consider a M-ary hypothesis testing problem. Assume that when a set Γ of sensors is available

*This research was supported in part by NSF grants ECS-0080406 and CCR-0082784, by the Department of Defense Research & Engineering (DDR&E) Multidisciplinary University Research Initiative (MURI) on ''Low Energy Electronics Design for Mobile Platforms'' and managed by the Army Research Office (ARO) under grant ARO DAAH04-96-1-0377, and by AFOSR grant F49620-96-1-0028.
†Rami Debouk is currently with GM R&D and Planning, Mail Code 480-106-390, 30500 Mound Road, P.O. Box 9055, Warren, MI 48090-9055, U.S.A.

$$P_\Gamma(\text{correct detection}|h_i, y_1, y_2, \ldots, y_T) > \alpha, \quad i = 1, 2, \ldots, M \tag{1}$$

where $P_\Gamma(\text{correct detection}|h_i, y_1, y_2, \ldots, y_T)$ denotes the probability that the i-th hypothesis is correctly identified when it is true, given T observations $y_1, y_2, \ldots, y_T$ collected from the set Γ of sensors, and α, $0 < \alpha < 1$, expresses the quality of performance required by the system over a T-horizon problem. Again, a cost function is associated with the subsets of Γ. We wish (i) to identify a least expensive combination of sensors, call it A^*, such that

$$P_{A^*}(\text{correct detection}|h_i, y_1, y_2, \ldots, y_T) > \alpha, \quad i = 1, 2, \ldots, M \tag{2}$$

and (ii) to determine the minimum number of tests required to identify such a combination of sensors A^*, assuming that a prespecified test is used to determine whether any set B of sensors satisfies

$$P_B(\text{correct detection}|h_i, y_1, y_2, \ldots, y_T) > \alpha, \quad i = 1, 2, \ldots, M \tag{3}$$

The costs associated with the sensors could have various interpretations. These costs could model the actual economic costs of using the given sensors. In this case the problem of finding a least costly set of sensors that diagnoses the system or satisfies (2) is of obvious economic interest. This objective is especially justified in situations where a combination of sensors is to be used in a large number of units (e.g., automobiles, appliances, airplanes, unmanned aerial vehicles, etc.) and the cost of that combination is to be taken as the total cost over all units. The cost of a sensor could also model its energy consumption, in which case our sensor selection problem becomes one of minimization of energy consumption in the process of diagnosing the system or satisfying (2). Furthermore, in all cases, it is desirable to minimize the expected computational (testing) cost associated with the identification of a least costly combination of sensors, because each test is time and memory consuming and requires significant computational effort.

In this paper, we formulate and analyze an optimization problem in sensor selection motivated by the above examples. We assume that a system possesses a certain Property D (e.g., the system is diagnosable) when a set Γ of sensors is used for information gathering. There are no restrictions on the class of dynamic systems considered. For each set A of sensors that is a subset of Γ, there is an *a priori* probability p_A that the system possesses Property D. There is a cost c_A associated with each set A of sensors that is a subset of Γ. Given a set A of sensors that is a subset of Γ, it is possible to determine, via a prespecified test, whether the resulting system-sensor combination possesses Property D. The objective is to determine a test strategy, i.e., a sequence of tests, to minimize the expected cost, associated with the tests, that is incurred until a least expensive system-sensor combination that possesses Property D is identified.

Related sensor selection problems have been previously studied in Bavishi and Chong (1994), Haji-Valizadeh and Loparo (1996) and Darabi and Jafari (1998) in the context of discrete event systems. In Bavishi and Chong (1994), the authors present an algorithm that determines the minimum-cost set of sensors which ensures system testability (a property defined in Bavishi and Chong (1994). In Haji-Valizadeh and Loparo (1996), Darabi and

Jafari (1998), the authors present algorithms that minimize the cardinality of the set of observable events and result in observable (normal) languages (see, for example, Cassandras and Lafortune, 1999, for definitions of these properties). The problems in Bavishi and Chong (1994), Haji-Valizadeh and Loparo (1996) and Darabi and Jafari (1998) as well as the problem investigated here have the following similar objective: the goal is to determine an optimal (according to some criterion) set of sensors or events with respect to which the system possesses a certain given property. A major difference, however, between Bavishi and Chong (1994), Haji-Valizadeh and Loparo (1996) and Darabi and Jafari (1998) and the present work is that we also determine a sequence of tests that minimizes the computational effort required to identify a least costly set of sensors that results in a system possessing the required property; this issue is not addressed in Bavishi and Chong (1994), Haji-Valizadeh and Loparo (1996) and Darabi and Jafari (1998).

In this paper we formulate the abovementioned optimization problem as a Markovian decision problem (MDP). Our objective is to obtain insight into the nature of this MDP by identifying instances where it is possible to explicitly determine optimal strategies. In this regard, we assume that any combination of l sensors has less cost than any combination of $(l+1)$ sensors for all l and that testing for Property D requires a constant cost. We also assume that any combination of l sensors results in a system possessing Property D with the same probability and the probability is an increasing function of l. Under these assumptions, we explicitly determine the sequence of tests that identifies a least (sensor) costly system-sensor combination that achieves Property D with the minimum expected number of tests.

This paper is organized as follows. In Section 2 we define one instance of the sensor selection problem. We formulate that instance of the sensor selection problem as a MDP and solve it in Section 3. We conclude with a discussion of the problem formulation and its solution in Section 4.

2. Problem Formulation

In principle, we can formulate the sensor selection problem described in Section 1 as a MDP with perfect observations (cf. Section 3.2). Such a formulation, without any further assumptions, would only allow determination of the optimal sequence of tests by computational methods, and would not provide any further insight into the structure of the optimal policy. To develop insight into the structure of the problem, we need to isolate instances where we can explicitly determine by analytical arguments the optimal sequence of tests. We present such an instance in this paper, where the analytical solution relies on specific assumptions on the costs of sensors and the *a priori* probabilities, p_A, as defined in Section 1.

We formulate and analyze the following sensor selection problem.

Problem P

We are given a dynamic system and a set Γ of sensors. We make the following assumptions:

A1 The system possesses Property D when the set Γ of sensors is used. The cardinality of Γ is denoted by K.

A2 Let $A_1, A_2, \ldots, A_{2^K-2}$ denote all elements in the power set of Γ except the empty set and Γ, and let c_{A_l} denote the total cost of all sensors in A_l, $l = 1, 2, \ldots, K'$ where $K' = 2^K - 2$. If $|A_l| < |A_m|$, $l, m \in \{1, 2, \ldots, K'\}$, then $c_{A_l} < c_{A_m}$. Furthermore, there is a total order in terms of cost among subsets of sensors that have the same cardinality. Therefore, without loss of generality, we label subsets of sensors so that

$$c_{A_1} < c_{A_2} < \cdots < c_{A_{K'}} \tag{4}$$

A3 The *a priori* probability that the system possesses Property D when a combination of i sensors is used is p_i. Furthermore,

$$0 = p_0 < p_1 \leq p_2 \cdots \leq p_{K-1} < p_K = 1 \tag{5}$$

A4 If the system with a set A of sensors is tested and found not to possess Property D, then p_B remains the same for all sets of sensors $B \supset A$. If the system with a set A of sensors is tested and found to possess Property D, then p_C remains the same for all sets of sensors $C \subset A$.

A5 A test performed off-line on any combination of sensors reveals whether the system possesses Property D when that combination of sensors is used. The cost of the test is independent of the combination of sensors on which it is performed, and is equal to c.

The objective is to determine a sequence of tests that minimizes the expected cost (associated with these tests) incurred until a least costly combination of sensors with respect to which the system possesses Property D is identified.

We briefly discuss the assumptions made in the above problem formulation. Without Assumption A1 there is no problem to solve. If the cost of any combination of l sensors is equal to the sum of the costs of its individual components and the individual sensor costs are comparable, then Assumption A2 is a reasonable assumption. However, Assumption A2 may not always be true. Moreover, under Assumption A2 ordering sensor combinations according to their cost is the same as ordering them according to their cardinality. In Assumption A3 the probabilities p_i, $i = 1, 2, \ldots, K-1$, are the result of prior knowledge based on experimental data. Assumption A3 is reasonable under the stipulation that all sensors have comparable information gathering ability. Assumption A4 can be justified as follows: the addition (subtraction) of a sensor to (from) a set of sensors increases (decreases) the diagnostic ability of the resulting set of sensors in a highly nonlinear fashion that is very difficult (if not impossible) to quantify. Thus, it is reasonable to assume that: (i) a negative test on a set A of sensors does not provide any additional information about the outcome of the test on any set B of sensors that is a superset of A; and (ii) a positive test on a set A of sensors does not provide any information on any set C of sensors that is a subset of A. Finally, Assumption A5 is true in several instances. For example, it holds in the M-ary hypothesis testing problem discussed in Section 1, and in the case when Property D is that of diagnosability and the test for diagnosability employed in Sampath et al. (1995, 1996, 1998) and Debouk et al. (2000) is used.

3. Solution

3.1. Preliminaries

We can visualize Problem P as a decision problem on a digraph (directed graph) as follows. The nodes of the digraph are elements of 2^{Γ}, the power set of Γ. Thus, each node represents a combination of sensors. Two nodes A and B are connected by a link, whose direction is from node B to node A, if and only if cardinality$(B) =$ cardinality$(A) + 1$ and $A \subset B$. Node $\emptyset$ of the network, called the empty node, corresponds to the empty subset of Γ, i.e., there are no sensors. An example of such a digraph in the case where there are four sensors is shown in Figure 1. The objective is to minimize the expected number of tests, each one performed on one node of the digraph, required to identify a node that represents a least costly combination of sensors with respect to which the system possesses Property D. Such a visualization facilitates the analysis of Problem P.

To analyze Problem P we need the following concepts.

DEFINITION 3.1 *Node B in the digraph possesses Property D if the system possesses Property D when the combination of sensors in B is used.*

DEFINITION 3.2 *Node B in the digraph possesses Property* $\overline{D}$ *if it does not possess Property D.*

The above definitions imply that the nodes of the digraph possess the following characteristics.

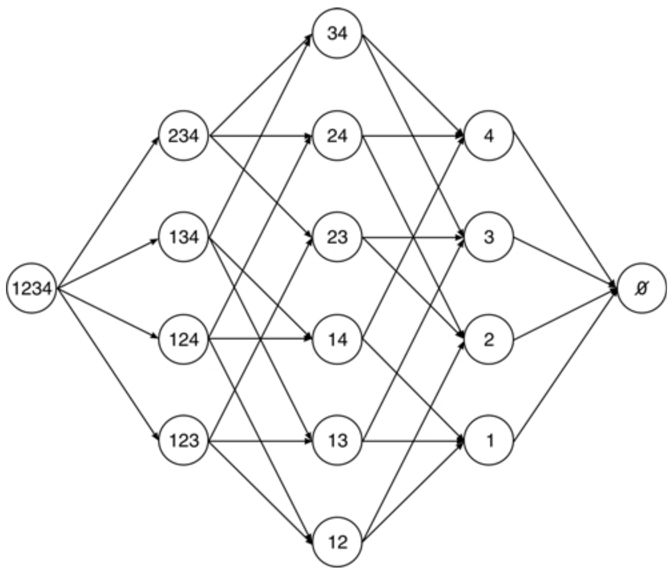

Figure 1. The digraph in the case where there are four sensors. The notation 12, for instance, is used to denote the set of sensors $\{1, 2\}$.

C1 The node containing the set Γ, or simply node Γ, possesses Property D.

C2 Node $\emptyset$ possesses Property $\overline{D}$.

C3 If node A possesses Property D, every node $B \supset A$ also possesses Property D.

C4 If node A possesses Property $\overline{D}$, every node $B \subset A$ also possesses Property $\overline{D}$.

In addition, the following definitions are needed in the subsequent development of the results.

DEFINITION 3.3 *Node B in the digraph is said to be a child of node C in the digraph if and only if there exists a directed path in the digraph from C to B.*

DEFINITION 3.4 *Node B in the digraph is said to be a parent of node C in the digraph if and only if there exists a directed path in the digraph from B to C.*

DEFINITION 3.5 *Let ξ be a set of nodes in the digraph and let $A \in \xi$. A reachable set from ξ, conditioned on the event that A possesses Property D, is a set composed of all nodes in ξ whose cost is strictly less than that of A.*

DEFINITION 3.6 *Let ξ be a set of nodes in the digraph and let $A \in \xi$. A reachable set from ξ, conditioned on the event that A possesses Property $\overline{D}$, is a set composed of all nodes in ξ except node A and its children in ξ.*

DEFINITION 3.7 *A reachable set is a set that results from applying an arbitrary sequence of n tests $\{t_1, t_2, \ldots, t_n\}$ to the set of nodes $2^\Gamma \backslash \{\emptyset, \Gamma\}$, while following the rules in Definitions 3.5 and 3.6.*

The concepts introduced by Definitions 3.1–3.7 as well as the characteristics C1–C4 are subsequently used in the formulation and analysis of Problem P as a MDP. We note that Definitions 3.1–3.7 are valid independently of whether Assumptions A2–A4 are true or not. We illustrate the above concepts with the following example.

Example 3.1: Consider the digraph presented in Figure 1. Let $\xi = 2^\Gamma \backslash \{\emptyset, \Gamma\}$, $A = \{1, 2, 3\}$, and $B = \{1, 2\}$. Assume that the costs of combinations of sensors are ordered as follows

$$\{1\} < \{2\} < \{3\} < \{1,2\} < \{4\} < \{1,3\} < \{2,3\} < \{1,4\} < \{2,4\} < \{1,2,3\} \\ < \{3,4\} < \{1,2,4\} < \{1,3,4\} < \{2,3,4\}$$

We have the following:

- The children of A in ξ are $\{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}$, and $\{2,3\}$.

- The parents of B in ξ are $\{1,2,3\}$ and $\{1,2,4\}$.
- The reachable set from ξ, conditioned on the event that B possesses Property D, is the set $\{\{1\},\{2\},\{3\}\}$.
- The reachable set from ξ, conditioned on the event that A possesses Property $\overline{D}$, is the set $\{\{4\},\{1,4\},\{2,4\},\{3,4\},\{1,2,4\},\{1,3,4\},\{2,3,4\}\}$.

Based on the above preliminaries, we next proceed to formulate Problem P as a MDP.

3.2. Formulation of Problem P as a Markovian Decision Problem

We formulate Problem P as a MDP with perfect observations. We then exploit the characteristics of reachable sets using C3 and C4, and Assumptions A1–A5 to characterize an optimal strategy for the MDP.

To clearly define the MDP we need to define the following five elements (Puterman, 1994): decision epochs, states, actions, transition probabilities, and costs or rewards. The set of decision epochs is $\{1,2,\ldots,K'\}$ since every epoch corresponds to testing a combination of sensors and we know from C1 and C2 that Γ and $\emptyset$ possess Properties D and $\overline{D}$, respectively. The state space of the process is the set of all sets of subsets of Γ (i.e., 2^{2^Γ}), hence a state of the process is a collection of subsets of Γ. The information state (Kumar and Varaiya, 1986) of the process is the set of nodes of the digraph that have not yet been checked for Property D and correspond to potentially least costly combinations of sensors possessing Property D. In this paper we use the terms state and information state interchangeably since we are considering a MDP with perfect observation (see Kumar and Varaiya, 1986). When N is the state of the process, the action space is N and an action $i \in N$ at a given epoch picks up the next subset of sensors to be tested. The transition probabilities are defined as follows:

$$P(N_i^D/N, i \in N \text{ is tested}) = p_{|i|} \tag{6}$$

and

$$P(N_i^{\overline{D}}/N, i \in N \text{ is tested}) = 1 - p_{|i|} \tag{7}$$

where N_i^D ($N_i^{\overline{D}}$) is the new state of the process, i.e., the reachable set from node i, that results when node i possesses Property D ($\overline{D}$), and $|i|$ denotes the cardinality of the set i. Finally the cost for every action $i \in N$ is equal to c. We note here that the individual cost of sensors appear in the optimality equation (10) through the sets N_i^D and $N_i^{\overline{D}}$. The following example illustrates the elements of the MDP.

Example 3.2: Let us revisit Example 3.1. $\Gamma = \{1,2,3,4\}$ and the decision epochs are $\{1,2,\ldots,14\}$. The state space of the process is 2^{2^Γ}, and a state N of the process is of the form

$$N = \{\{4\}, \{1,4\}, \{2,4\}, \{3,4\}, \{1,2,4\}, \{1,3,4\}, \{2,3,4\}\}$$

When the state of the process is N the action space is N, and an action is an element in N. The transition probabilities are defined as follows:

$$P(N_i^D / N, i \in N \ \text{ is tested}) = p_{|i|}, \quad |i| = 1,2,3 \tag{8}$$

and

$$P(N_i^{\overline{D}} / N, i \in N \ \text{ is tested}) = 1 - p_{|i|}, \quad |i| = 1,2,3 \tag{9}$$

Finally the cost of testing is equal to c for every action $i \in N$.

Let $V(N)$ denote the minimum expected (test) cost incurred when the state is N. Then, $V(N)$ satisfies the optimality equation

$$V(N) = \min_{i \in N} \{c + p_{|i|} \cdot V(N_i^D) + (1 - p_{|i|}) \cdot V(N_i^{\overline{D}})\} \tag{10}$$

where $i \in \{A_1, \ldots, A_{K'}\}$, $|i|$ denotes the cardinality of the set i, and $N_i^D (N_i^{\overline{D}})$ is the new state of the process, i.e., the reachable set from node i, that results when node i possesses Property D $(\overline{D})$. These states are determined from N by using C3 and C4, Definitions 3.5 and 3.6, and Assumption A2. The probabilities $p_{|i|}$, $|i| = 1, 2, \ldots, K-1$ remain unchanged because of Assumption A4. With a slight abuse of notation, and for simplicity, we thereafter use p_i instead of $p_{|i|}$ in (10) to denote the *a priori* probability associated with the set of sensors i. Hence, from now on (10) will read as follows

$$V(N) = \min_{i \in N} \{c + p_i \cdot V(N_i^D) + (1 - p_i) \cdot V(N_i^{\overline{D}})\} \tag{11}$$

3.3. Analysis

Equation (11) can be solved by backward induction (Puterman, 1994) under any conditions on the prior probabilities p_i and on the cost of combinations of sensors. This requires a precise way of ordering and enumerating states. The following example describes that procedure applied to Example 3.1.

Example 3.3: We consider Example 3.1. We briefly illustrate how the backward induction procedure works in the context of the example. We proceed in stages.

At the first stage we consider all states that consist of one subset of sensors of Γ. These states are $\{\{1\}\}$, $\{\{2\}\}$, $\{\{3\}\}$, $\{\{4\}\}$, $\{\{1,2\}\}$, $\{\{1,3\}\}$, $\{\{2,3\}\}$, $\{\{1,4\}\}$, $\{\{2,4\}\}$, $\{\{3,4\}\}$, $\{\{1,2,3\}\}$, $\{\{1,2,4\}\}$, $\{\{1,3,4\}\}$, $\{\{2,3,4\}\}$. The cost associated with each of the above states is equal to c, the cost of a single test.

At the second stage we consider all states that consist of two subsets of sensors of Γ. We will not enumerate these states, but will illustrate how to compute the cost of each via the optimality equation (11) by focusing on one such state, say $N = \{\{2,4\},\{1,2,3\}\}$. The possible actions at N are to test either $\{2,4\}$ or $\{1,2,3\}$. The cost associated with testing $\{2,4\}$ is equal to

$$c + (1 - p_2) \cdot V(\{1,2,3\}) = c + (1 - p_2) \cdot c = (2 - p_2) \cdot c \tag{12}$$

since $N^{D}_{\{2,4\}} = \emptyset$ (because of A2), $N^{\overline{D}}_{\{2,4\}} = \{1,2,3\}$, and $V(\{1,2,3\}) = c$ (from the first stage). The cost associated with testing $\{1,2,3\}$ is equal to

$$c + p_3 \cdot V(\{2,4\}) + (1 - p_3) \cdot V(\{2,4\}) = c + p_3 \cdot c + (1 - p_3) \cdot c = 2 \cdot c \tag{13}$$

since $N^{D}_{\{1,2,3\}} = \{2,4\}$ (because of A2), $N^{\overline{D}}_{\{1,2,3\}} = \{2,4\}$, and $V(\{2,4\}) = c$ (from the first stage). From (11), (12), and (13) we conclude that at state $N = \{\{2,4\},\{1,2,3\}\}$ it is optimal to test $\{2,4\}$ and

$$V(\{\{2,4\},\{1,2,3\}\}) = (2 - p_2) \cdot c \tag{14}$$

The cost of all other states N that consist of two subsets of Γ are computed in a similar fashion

At third, fourth, fifth, ... stages we consider all states that consist of 3, 4, 5, ... subsets of Γ, respectively. At each stage, the cost of the corresponding states is computed by considering all possible actions at each state, using the results of the preceding stages to determine the cost of each action, and selecting the minimum of these costs.

The above described procedure determines the cost of each state, and the optimal action (i.e., the optimal test) at each state.

The solution of equation (11) by backward induction as illustrated in Example 3.3 would apply to specific situations. Furthermore, it does not provide any insight into the nature of the solution of Problem P. We are interested in seeking structural results, i.e., in determining properties and characteristics of optimal test strategies for Problem P so that we develop some insight into the nature of the problem. We have isolated one instance of Problem P where, under the set of Assumptions A1–A5, we can explicitly determine an optimal test policy. The following theorem summarizes the solution to Problem P under Assumptions A1–A5 and provides the main result of this paper.

THEOREM 3.1 *If $p_l + p_{l+1} \geq 1$ for $l = 1, \ldots, K-2$, then an optimal test strategy for Problem P is to test combinations of sensors in increasing order of (sensor) cost as defined in Assumption A2.*

Proof of Theorem 3.1: The proof proceeds by induction on the cardinality of the information state N. Let $V_k(N)$ denote the expected cost incurred when we choose to test $k \in N$ when the information state is N and follow the optimal test strategy afterwards, that is

$$V_k(N) = c + p_k \cdot V(N^{D}_{k}) + (1 - p_k) \cdot V(N^{\overline{D}}_{k}) \tag{15}$$

and

$$V(N) = \min_{k \in N} V_k(N) \tag{16}$$

- *Basis of induction* Let $|N| = 2$, i.e., $N = \{i, j\}$ where $i, j \in \{A_1, \ldots, A_{K'}\}$ are subsets of sensors. There are three possibilities for such information states (based on Definitions 3.5 and 3.6 of reachable sets):

  1. $|i| = |j|$ such that the cost attached to i is less than or equal to that attached to j. In this case

  $$V_i(N) = c + p_i \cdot V(N_i^D) + (1 - p_i) \cdot V(N_i^{\overline{D}}) = c + (1 - p_i) \cdot c \tag{17}$$

  since $N_i^D = \emptyset$, i.e., no more tests are needed, and $N_i^{\overline{D}} = \{j\}$. Also,

  $$V_j(N) = c + p_j \cdot V(N_j^D) + (1 - p_j) \cdot V(N_j^{\overline{D}}) = c + p_j \cdot c + (1 - p_j) \cdot c \tag{18}$$

  since $N_j^D = \{i\}$, and $N_j^{\overline{D}} = \{i\}$. From (17) and (18) it follows that it is optimal to first test i and then j since $p_i = p_j$, and $p_j \cdot c$ is always strictly positive.

  2. $|i| = |j| - 1$ with i being a child of j. In this case

  $$V_i(N) = c + p_i \cdot V(N_i^D) + (1 - p_i) \cdot V(N_i^{\overline{D}}) = c + (1 - p_i) \cdot c \tag{19}$$

  since $N_i^D = \emptyset$ (by Assumption A2), i.e., no more tests are needed, and $N_i^{\overline{D}} = \{j\}$. Also,

  $$V_j(N) = c + p_j \cdot V(N_j^D) + (1 - p_j) \cdot V(N_j^{\overline{D}}) = c + p_j \cdot c \tag{20}$$

  since $N_j^D = \{i\}$, and $N_j^{\overline{D}} = \emptyset$. From (19) and (20) we conclude that it is optimal to first test i and then j because by assumption $p_i + p_j \geq 1$, implying that $1 - p_i \leq p_j$.

  3. $|i| = |j| - 1$ such that i is not a child of j. In this case

  $$V_i(N) = c + p_i \cdot V(N_i^D) + (1 - p_i) \cdot V(N_i^{\overline{D}}) = c + (1 - p_i) \cdot c \tag{21}$$

  since $N_i^D = \emptyset$ (by Assumption A2), i.e., no more tests are needed, and $N_i^{\overline{D}} = \{j\}$. Also,

  $$V_j(N) = c + p_j \cdot V(N_j^D) + (1 - p_j) \cdot V(N_j^{\overline{D}}) = c + p_j \cdot c + (1 - p_j) \cdot c = c + c \tag{22}$$

  since $N_j^D = \{i\}$, and $N_j^{\overline{D}} = \{i\}$. From (21) and (22) it follows that it is optimal to first test i and then j since $1 - p_i$ is always strictly less than 1.

- *Induction step* Assume that the assertion of the theorem is true for any set N' with cardinality k, i.e., under Assumptions A1–A5 and $p_l + p_{l+1} \geq 1$ for all l, and for any set of cardinality less than or equal to k an optimal test strategy is to test combinations of sensors by increasing order of their cost. We must prove that for any set N of cardinality $k+1$, the assertion of the theorem is still true. Let i, j be elements of N (that is i, j are subsets of sensors) such that the cost attached to i is the smallest cost attached to any combination of sensors in N. Clearly, $|j| \geq |i|$ by Assumption A2. We need to prove that

$$\begin{aligned} V_i(N) &= c + p_i \cdot V(N_i^D) + (1-p_i) \cdot V(N_i^{\overline{D}}) \leq c + p_j \cdot V(N_j^D) + (1-p_j) \cdot V(N_j^{\overline{D}}) \\ &= V_j(N) \end{aligned} \tag{23}$$

or

$$p_i \cdot V(N_i^D) + (1-p_i) \cdot V(N_i^{\overline{D}}) \leq p_j \cdot V(N_j^D) + (1-p_j) \cdot V(N_j^{\overline{D}}) \tag{24}$$

where

$$\begin{aligned} N_i^D &= \emptyset \\ N_i^{\overline{D}} &= N \backslash \{i\} \\ N_j^D &= \{x \in N : |x| < |j|, \text{or } |x| = |j| \text{ and the cost attached to } x \text{ is less than} \\ &\qquad \text{that attached to } j \text{ and } x \neq j\} \\ N_j^{\overline{D}} &= \{x \in N : x \neq j, \text{and } x \text{ is not a child of } j\} \end{aligned} \tag{25}$$

We can apply the induction hypothesis to the sets $N_i^{\overline{D}}$, N_j^D, and $N_j^{\overline{D}}$ since their cardinalities are less than or equal to k (cf. (25)). By successively applying the induction hypothesis until all elements in $N_i^{\overline{D}}$, N_j^D, and $N_j^{\overline{D}}$ are exhausted, we get

$$\begin{aligned} V(N_i^{\overline{D}}) = c \cdot \Bigg[& \sum_{n=0}^{v_i - 1} (1-p_i)^n + (1-p_i)^{v_i} \cdot \sum_{n=0}^{v_{i+1}-1} (1-p_{i+1})^n \\ & + \cdots + (1-p_i)^{v_i} \cdot (1-p_{i+1})^{v_{i+1}} \cdot \cdots \cdot (1-p_{j-1})^{v_{j-1}} \cdot \sum_{n=0}^{v_j - 1} (1-p_j)^n \\ & + \cdots + (1-p_i)^{v_i} \cdot (1-p_{i+1})^{v_{i+1}} \cdot \cdots \cdot (1-p_{L-1})^{v_{L-1}} \cdot \sum_{n=0}^{v_L - 1} (1-p_L)^n \Bigg] \end{aligned} \tag{26}$$

$$\begin{aligned} V(N_j^D) = c \cdot \Bigg[& \sum_{n=0}^{r_i - 1} (1-p_i)^n + (1-p_i)^{r_i} \cdot \sum_{n=0}^{r_{i+1}-1} (1-p_{i+1})^n \\ & + \cdots + (1-p_i)^{r_i} \cdot (1-p_{i+1})^{r_{i+1}} \cdot \cdots \cdot (1-p_{j-1})^{r_{j-1}} \cdot \sum_{n=0}^{r_j - 1} (1-p_j)^n \Bigg] \end{aligned} \tag{27}$$

and

$$
\begin{aligned}
V(N_j^{\overline{D}}) = c \cdot \Bigg[& \sum_{n=0}^{s_i - 1} (1 - p_i)^n + (1 - p_i)^{s_i} \cdot \sum_{n=0}^{s_{i+1} - 1} (1 - p_{i+1})^n \\
& + \cdots + (1 - p_i)^{s_i} \cdot (1 - p_{i+1})^{s_{i+1}} \cdot \cdots \cdot (1 - p_{L-1})^{s_{L-1}} \cdot \sum_{n=0}^{s_L - 1} (1 - p_L)^n \Bigg]
\end{aligned}
\tag{28}
$$

where v_k denotes the number of combinations of k sensors in $N_i^{\overline{D}}$, r_k denotes the number of combinations of k sensors in N_j^D, s_k denotes the number of combinations of k sensors in $N_j^{\overline{D}}$ and L denotes the highest set cardinality in N with $|j| \leq L \leq K - 1$. In Appendix A we explain in more detail how equations (26)–(28) are obtained. Because of (25)–(28), the inequality we need to prove in (24) can be written as follows (after canceling c from both sides)

$$
\begin{aligned}
(1 - p_i) \cdot \Bigg[& \sum_{n=0}^{v_i - 1} (1 - p_i)^n + (1 - p_i)^{v_i} \cdot \sum_{n=0}^{v_{i+1} - 1} (1 - p_{i+1})^n \\
& + \cdots + (1 - p_i)^{v_i} \cdot (1 - p_{i+1})^{v_{i+1}} \cdot \cdots \cdot (1 - p_{j-1})^{v_{j-1}} \cdot \sum_{n=0}^{v_j - 1} (1 - p_j)^n \\
& + \cdots + (1 - p_i)^{v_i} \cdot (1 - p_{i+1})^{v_{i+1}} \cdot \cdots \cdot (1 - p_{L-1})^{v_{L-1}} \cdot \sum_{n=0}^{v_L - 1} (1 - p_L)^n \Bigg] \\
\leq p_j \cdot \Bigg[& \sum_{n=0}^{r_i - 1} (1 - p_i)^n + (1 - p_i)^{r_i} \cdot \sum_{n=0}^{r_{i+1} - 1} (1 - p_{i+1})^n \\
& + \cdots + (1 - p_i)^{r_i} \cdot (1 - p_{i+1})^{r_{i+1}} \cdot \cdots \cdot (1 - p_{j-1})^{r_{j-1}} \cdot \sum_{n=0}^{r_j - 1} (1 - p_j)^n \Bigg] \\
+ (1 - p_j) \cdot \Bigg[& \sum_{n=0}^{s_i - 1} (1 - p_i)^n + (1 - p_i)^{s_i} \cdot \sum_{n=0}^{s_{i+1} - 1} (1 - p_{i+1})^n \\
& + \cdots + (1 - p_i)^{s_i} \cdot (1 - p_{i+1})^{s_{i+1}} \cdot \cdots \cdot (1 - p_{L-1})^{s_{L-1}} \cdot \sum_{n=0}^{s_L - 1} (1 - p_L)^n \Bigg]
\end{aligned}
\tag{29}
$$

We consider two cases: $|j| > |i|$, and $|j| = |i|$ such that the cost attached to j is higher than that attached to i.

Case 1: $|j| > |i|$. In this case, from the definition of reachable sets (Definitions 3.5 and 3.6), C3 and C4 we have the following relations

$$
\begin{aligned}
& r_j < v_j \\
& r_k = v_k \quad k = |i| + 1, \ldots, |j| - 1 \\
& r_i = v_i + 1 \\
& s_k = v_k \quad k = |j| + 1, \ldots, L \\
& s_j = v_j - 1 \\
& s_k \leq v_k \quad k = |i| + 1, \ldots, |j| - 1 \\
& s_i \leq v_i + 1
\end{aligned} \tag{30}
$$

The set of equalities and inequalities (30) is best explained by referring to (25). For instance, $r_j < v_j$ since the elements in N_i^D having $|j|$ sensors (denoted by v_j) are equal to the elements in N having $|j|$ sensors while the elements in N_j^D having $|j|$ sensors (denoted by r_j) are those elements in $N \backslash \{j\}$ having $|j|$ sensors and a cost strictly less than that of element j. Verification of the remaining equalities and inequalities can be found in Appendix B. Because of (29) and of the relations between the r_k's and v_k's, $k = |i|, \ldots, |j| - 1$, in (30), the inequality we must prove to complete the proof of the induction step in Case 1 is

$$
\begin{aligned}
(1 - p_i) \cdot \Bigg[& \sum_{n=0}^{v_i - 1} (1 - p_i)^n + (1 - p_i)^{v_i} \cdot \sum_{n=0}^{v_{i+1} - 1} (1 - p_{i+1})^n \\
& + \cdots + (1 - p_i)^{v_i} \cdot (1 - p_{i+1})^{v_{i+1}} \cdot \cdots \cdot (1 - p_{j-1})^{v_{j-1}} \cdot \sum_{n=0}^{v_j - 1} (1 - p_j)^n \\
& + \cdots + (1 - p_i)^{v_i} \cdot (1 - p_{i+1})^{v_{i+1}} \cdot \cdots \cdot (1 - p_{L-1})^{v_{L-1}} \cdot \sum_{n=0}^{v_L - 1} (1 - p_L)^n \Bigg] \\
& \leq p_j \cdot \Bigg[\sum_{n=0}^{v_i} (1 - p_i)^n + (1 - p_i)^{v_i + 1} \cdot \sum_{n=0}^{v_{i+1} - 1} (1 - p_{i+1})^n \\
& + \cdots + (1 - p_i)^{v_i + 1} \cdot (1 - p_{i+1})^{v_{i+1}} \cdot \cdots \cdot (1 - p_{j-1})^{v_{j-1}} \cdot \sum_{n=0}^{r_j - 1} (1 - p_j)^n \Bigg] \\
& + (1 - p_j) \cdot \Bigg[\sum_{n=0}^{s_i - 1} (1 - p_i)^n + (1 - p_i)^{s_i} \cdot \sum_{n=0}^{s_{i+1} - 1} (1 - p_{i+1})^n \\
& + \cdots + (1 - p_i)^{s_i} \cdot (1 - p_{i+1})^{s_{i+1}} \cdot \cdots \cdot (1 - p_{L-1})^{s_{L-1}} \cdot \sum_{n=0}^{s_L - 1} (1 - p_L)^n \Bigg]
\end{aligned} \tag{31}
$$

Because of Assumption A3 the condition $p_l + p_{l+1} \geq 1$ for all l of the theorem implies that $1 - p_i \leq p_j$. Moreover, the terms multiplied by p_j on the right hand side of (31) are term by term greater than or equal to the first $v_i + 1 + v_{i+1} + \cdots + v_{j-1} + r_j$ terms multiplied by $(1 - p_i)$ (in between braces) on the left hand side. Hence, because of (31) and the above observation, to complete the proof of the induction step in Case 1 it is sufficient to prove that

$$
\begin{aligned}
&(1-p_i)\cdot(1-p_i)^{v_i}\cdot\cdots\cdot(1-p_{j-1})^{v_{j-1}}\cdot(1-p_j)^{r_j+1}\cdot\left[\sum_{n=0}^{v_j-r_j-2}(1-p_j)^n\right.\\
&\quad+(1-p_j)^{v_j-r_j-1}\cdot\sum_{n=0}^{v_{j+1}-1}(1-p_{j+1})^n\\
&\quad\left.+(1-p_j)^{v_j-r_j-1}\cdot(1-p_{j+1})^{v_{j+1}}\cdot\cdots\cdot(1-p_{L-1})^{v_{L-1}}\cdot\sum_{n=0}^{v_L-1}(1-p_L)^n\right]\\
&\quad\le(1-p_j)\cdot\left[\sum_{n=0}^{s_i-1}(1-p_i)^n+(1-p_i)^{s_i}\cdot\sum_{n=0}^{s_{i+1}-1}(1-p_{i+1})^n\right.\\
&\quad\left.+\cdots+(1-p_i)^{s_i}\cdot(1-p_{i+1})^{s_{i+1}}\cdot\cdots\cdot(1-p_{L-1})^{s_{L-1}}\cdot\sum_{n=0}^{s_L-1}(1-p_L)^n\right]
\end{aligned}\tag{32}
$$

Since $(1-p_i)\cdot(1-p_i)^{v_i}\cdot(1-p_{i+1})^{v_{i+1}}\cdot\cdots\cdot(1-p_{j-1})^{v_{j-1}}\cdot(1-p_j)^{r_j+1}\le(1-p_j)$, to establish (32) it is sufficient to show that

$$
\begin{aligned}
&\sum_{n=0}^{v_j-r_j-2}(1-p_j)^n+(1-p_j)^{v_j-r_j-1}\cdot\sum_{n=0}^{v_{j+1}}(1-p_{j+1})^n+\cdots+(1-p_j)^{v_j-r_j-1}\\
&\cdot(1-p_{j+1})^{v_{j+1}}\cdot\cdots\cdot(1-p_{L-1})^{v_{L-1}}\cdot\sum_{n=0}^{v_L-1}(1-p_L)^n\\
&\le\sum_{n=0}^{s_i-1}(1-p_i)^n+(1-p_i)^{s_i}\cdot\sum_{n=0}^{s_{i+1}-1}(1-p_{i+1})^n+\cdots+(1-p_i)^{s_i}\\
&\cdot(1-p_{i+1})^{s_{i+1}}\cdot\cdots\cdot(1-p_{L-1})^{s_{L-1}}\cdot\sum_{n=0}^{s_L-1}(1-p_L)^n
\end{aligned}\tag{33}
$$

From (30) we have that $s_j=v_j-1>v_j-r_j-1$ since $r_j\ge0$, and $s_k=v_k$, $k=|j|+1,\ldots,L$. Hence, the following is true

$$(v_j-r_j-1)+v_{j+1}+\cdots+v_L\le s_i+s_{i+1}+\cdots+s_L \tag{34}$$

Inequality (34) implies that (33) is true since the terms in the left hand side of (33) are term by term less than or equal to the first $(v_j-r_j-1)+v_{j+1}+\cdots+v_L$ terms on the right hand side of (33) (by noting that since $p_i\le p_j$ then $(1-p_j)\le(1-p_i)$). The proof of the induction step is complete and the assertion of the theorem is true for Case 1.

Case 2: $|j|=|i|$. Such that the cost attached to j is higher than that attached to i. In this case Inequality (29) becomes

$$
\begin{aligned}
&(1-p_i)\cdot\left[\sum_{n=0}^{v_i-1}(1-p_i)^n+(1-p_i)^{v_i}\cdot\sum_{n=0}^{v_{i+1}-1}(1-p_{i+1})^n\right.\\
&\quad\left.+\cdots+(1-p_i)^{v_i}\cdot(1-p_{i+1})^{v_{i+1}}\cdot\cdots\cdot(1-p_{L-1})^{v_{L-1}}\cdot\sum_{n=0}^{v_L-1}(1-p_L)^n\right]\\
&\le p_i\cdot\sum_{n=0}^{r_i-1}(1-p_i)^n+(1-p_i)\cdot\left[\sum_{n=0}^{s_i-1}(1-p_i)^n\right.\\
&\quad+(1-p_i)^{s_i}\cdot\sum_{n=0}^{s_{i+1}-1}(1-p_{i+1})^n+\cdots+(1-p_i)^{s_i}\cdot(1-p_{i+1})^{s_{i+1}}\\
&\quad\left.\cdot\cdots\cdot(1-p_{L-1})^{s_{L-1}}\cdot\sum_{n=0}^{s_L-1}(1-p_L)^n\right]
\end{aligned}
\tag{35}
$$

where v_k, r_k, s_k, and L are as defined in Case 1, and they are related as follows

$$
\begin{aligned}
&r_i \le v_i\\
&s_k = v_k \quad k=|i|,\ldots,L
\end{aligned}
\tag{36}
$$

Observe that $s_k = v_k$ for $k=|i|,\ldots,L$ since the sets $N_i^{\overline{D}}$ and $N_j^{\overline{D}}$ contain the same combinations of $|i|$ sensors or more. The inequality $r_i \le v_i$ is true since the set $N_i^{\overline{D}}$ contains all elements of $|i|$ sensors in N except the element i, while $N_j^{\overline{D}}$ only contains elements with $|i|$ sensors that have a cost strictly less than that of the element j. Hence, to complete the proof of the induction step in Case 2 we must prove that (35) is true. For that matter we observe that the terms multiplied by $(1-p_i)$ on the left hand side of (35) are term by term equal to the terms multiplied by $(1-p_i)$ on the right hand side of (35). Hence the inequality in (35) is always true.

The proof of the induction step is now complete. Therefore, the assertion of the theorem is true. ■

Theorem 3.1 presents conditions sufficient to guarantee that the strategy that tests combinations of sensors in increasing order of cost is optimal for Problem P. This result may be intuitively interpreted as follows: the conditions $p_l + p_{l+1} \ge 1$, $l = 1, 2, \ldots, K-2$, imply that $p_i \ge 1/2$ for $i = 2, \ldots, K-1$. Thus, it is assumed *a priori* that a combination of two or more sensors is equally or more likely to result in a system possessing Property D than a system possessing Property $\overline{D}$. Therefore, given that the cost of testing for Property D is independent of the cardinality of the combination of sensors that is being used, it is plausible that the optimal test will examine combinations of sensors in increasing order of cost.

The proof of Theorem 3.1 is achieved by showing that the test strategy claimed to be optimal performs no worse than any other strategy. The method of proof of the theorem cannot be used to compare any two test strategies, none of which is the one specified by the theorem. Later, in Section 3.4, we show that under a slightly modified set of assumptions it is possible to compare test strategies that satisfy certain conditions even when none of these strategies is an optimal one (see Theorem 3.2).

Using the result of Theorem 3.1, it is possible to determine an analytic expression for the value function $V(N)$ that satisfies equation (11).

COROLLARY 3.1 *Consider Problem P with the conditions* $p_l + p_{l+1} \geq 1,\ l = 1, 2, \ldots, K-2$.

i. Let N be a reachable set. Then the minimum expected cost associated with N is

$$V(N) = c \cdot \left[\sum_{n=0}^{n_l - 1} (1-p_l)^n + (1-p_l)^{n_l} \cdot \sum_{n=0}^{n_{l+1}-1} (1-p_{l+1})^n \right.$$
$$\left. + \cdots + (1-p_l)^{n_l} \cdot (1-p_{l+1})^{n_{l+1}} \cdot \cdots \cdot (1-p_{h-1})^{n_{h-1}} \cdot \sum_{n=0}^{n_h - 1} (1-p_h)^n \right] \quad (37)$$

where l and h denote, respectively, the lowest and highest set cardinalities in N and n_l *and* n_h *denote the number of combinations of sensors in N of cardinality l and h, respectively.*

ii. Let N_1 *and* N_2 *be two reachable sets. If* $N_1 \subseteq N_2$ *then* $V(N_1) \leq V(N_2)$.

Proof of Corollary 3.1: The proof of part (*i*) of this corollary can be found in the proof of Theorem 3.1. Namely, in the proof of Theorem 3.1 we showed that (29) and (35) are always true, hence the left hand side of the inequality (after adding the constant c) represents the cost associated with the optimal policy. That cost can be rewritten as presented in the statement of the corollary. The proof of part *ii* of the corollary is a direct consequence of part *i*. If we write the expansions of $V(N_1)$ and $V(N_2)$ according to (37), with the terms ordered as given in (37), then the $|N_1|$ elements in $V(N_1)$ are term by term less than the first $|N_1|$ elements in $V(N_2)$ since $(1-p_i) > (1-p_{i+1})$ for all i. ■

The following example illustrates in part the results of our analysis. It shows that when the conditions $p_l + p_{l+1} \geq 1, l = 1, 2, \ldots, K-2$, are not satisfied the result of Theorem 3.1 is not, in general, true. It also provides a comparison of the test policy g_1, that tests combinations of sensors in increasing order of cost, with two other specific policies, and demonstrates that when the conditions of Theorem 3.1 are satisfied, g_1 outperforms these two policies.

Example 3.4: Consider the digraph presented in Figure 1 and let the information state $N = 2^{\Gamma} \backslash \{\Gamma, \emptyset\}$. Assume that the combination of sensors {1,2,3,4} achieves Property D, the costs of combinations of sensors are ordered as follows

$$\{1\} < \{2\} < \{3\} < \{4\} < \{1,2\} < \{1,3\} < \{1,4\} < \{2,3\} < \{2,4\} < \{3,4\}$$
$$< \{1,2,3\} < \{1,2,4\} < \{1,3,4\} < \{2,3,4\}$$

$p_1 < p_2 < p_3$, and for any combination of sensors the cost for testing for Property D is equal to unity. Hence Assumptions A1–A3 and A5 are met. We also assume that Assumption A4 is satisfied. Denote by g_1 the policy that always tests combinations of sensors in increasing order of cost. We compare g_1 with two other policies g_2 and g_3 defined as follows:

g_2 Always test combinations of sensors in decreasing order of cardinality with the tie-breaker being the maximum number of children a combination has;

g_3 Always test combinations of sensors whose cardinality equals the median of available cardinalities in the information state with the tie-breaker being the maximum number of children a combination has.

Figure 2 depicts the cost associated with the three policies above as function of p_2 with fixed p_1 and p_3. In the top left graph, the conditions of Theorem 3.1 are satisfied, so g_1 outperforms g_2 and g_3 as expected. In the top right (bottom left) graph the condition $p_1 + p_2 \geq 1$ (all conditions) of Theorem 3.1 is (are) violated; yet g_1 outperforms g_2 and g_3 in both instances. In the bottom right graph, the conditions $p_l + p_{l+1} \geq 1$, $l = 1, 2$, are all violated and g_1 is inferior to g_2 and g_3.

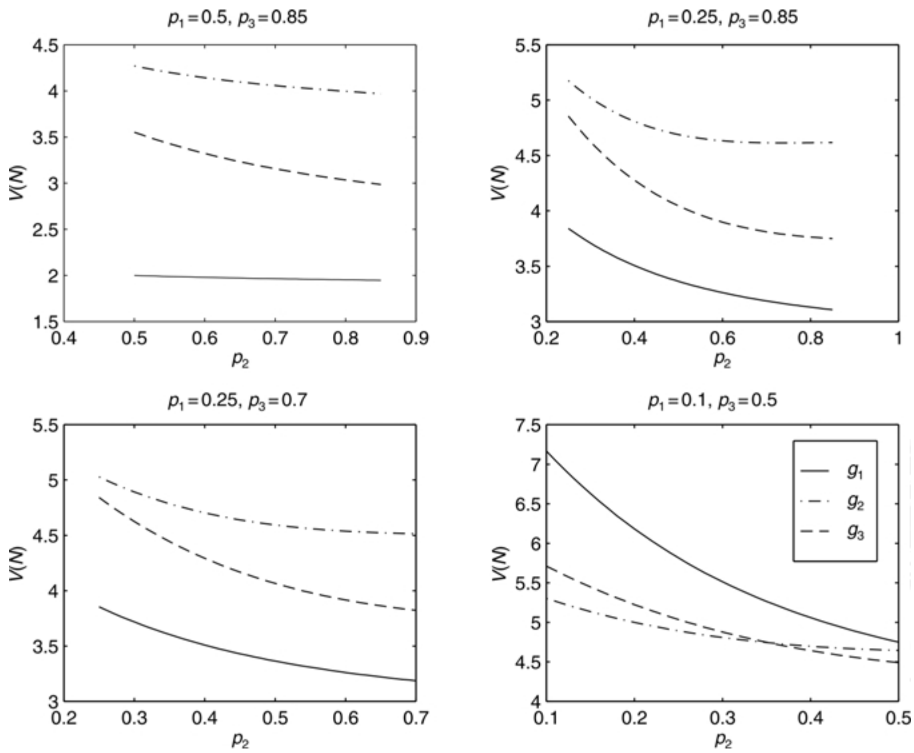

Figure 2. Policy comparisons for Example 3.4.

3.4. Analysis of a Special Instance of Problem P

As discussed in Section 3.3, under Assumptions A1–A5 we can show that the test strategy described in Theorem 3.1 is optimal but the theorem only includes a comparison between any strategy and an optimal strategy. By slightly modifying one assumption we can obtain one instance of Problem P that has more structural properties than those described in Theorem 3.1. Specifically, by replacing Assumption A3 with

A3′ The *a priori* probability that the system possesses Property D when the combination of i sensors is used, $1 < i < K - 1$, is equal to a constant, p,

we achieve the following:

- Comparison of any two test strategies that satisfy certain conditions even when none of these strategies is an optimal one (cf. Theorem 3.2).
- Discovery of an additional property of the value function $V(\cdot)$ (cf. Theorem 3.3).

Assumption A3′ may not always be realistic. It may be justified only in the case where there is no prior experimental data available about the diagnostic properties of any set of sensors; in such a case one might choose $p = p_i = 1/2$ for all i. The result of Theorem 3.2 that follows is valid for all $p \geq 1/2$; Theorem 3.3 is true for all p.

Under Assumptions A1, A2, A3′, A4, and A5, if $p_l + p_{l+1} \geq 1$ or $p \geq 1/2$, Theorem 3.1 is still valid since A3′ implies A3. Furthermore, under A1, A2, A3′, A4, A5, and $p \geq 1/2$ we can compare the expected costs incurred when at any instant of time we test any two distinct combinations of sensors within a reachable set and we follow the optimal test strategy afterwards. The following theorem precisely states the result.

THEOREM 3.2 *Let N be an information state, and let i and j be two elements in N. Under Assumptions A1, A2, A3′, A4, A5, and $p \geq 1/2$ the following are true:*

i. If $|N_i^D| + |N_i^{\overline{D}}| \leq |N_j^D| + |N_j^{\overline{D}}|$ (38)

and $|N_i^D| \leq |N_j^D| \leq |N_i^{\overline{D}}|, \quad |N_i^D| \leq |N_j^{\overline{D}}| \leq |N_i^{\overline{D}}|$ (39)

then $V_i(N) \leq V_j(N)$ (40)

where $V_k(N)$ denotes, as previously, the expected (test) cost incurred when we choose to test $k \in N$ when the information state is N and follow the optimal test strategy afterwards.

ii. An optimal test strategy for the sensor selection problem under consideration is to test combinations of sensors in increasing order of their cost. Moreover, the minimum expected cost associated with N is

$$V(N) = c \cdot \sum_{n=0}^{|N|-1} (1-p)^n \tag{41}$$

Proof of Theorem 3.2: Part *ii* is already proved by Theorem 3.1 and the fact that Assumption A3$'$ implies Assumption A3. Equation (41) is obtained by letting all probabilities equal to p in equation (37). We prove part *i* by induction on the cardinality of the information state N.

- *Basis of induction* Let $|N| = 2$, i.e., $N = \{i, j\}$ where $i, j \in \{A_1, \ldots, A_{K'}\}$ are subsets of sensors. There are three possibilities for such information states (based on Definitions 3.5 and 3.6 of reachable sets):

1,2: $|i| = |j|$ such that the cost attached to i is less than that attached to j, or $|i| = |j| - 1$ where i is not a child of j. In this case

$$V(N_i^D) = \emptyset, V(N_i^{\overline{D}}) = \{j\}, V(N_j^D) = \{i\}, \quad \text{and } V(N_j^{\overline{D}}) = \{i\} \tag{42}$$

Therefore,

$$|N_i^D| + |N_i^{\overline{D}}| \le |N_j^D| + |N_j^{\overline{D}}| \tag{43}$$

and

$$|N_i^D| \le |N_j^D| \le |N_i^{\overline{D}}|, \quad |N_i^D| \le |N_j^{\overline{D}}| \le |N_i^{\overline{D}}| \tag{44}$$

Moreover,

$$V_i(N) = c + p \cdot V(N_i^D) + (1-p) \cdot V(N_i^{\overline{D}}) = c + (1-p) \cdot c = c \cdot (2-p) \tag{45}$$

and

$$V_j(N) = c + p \cdot V(N_j^D) + (1-p) \cdot V(N_j^{\overline{D}}) = c + p \cdot c + (1-p) \cdot c = 2 \cdot c \tag{46}$$

From (45) and (46) it follows that $V_i(N) \le V_j(N)$ and it is optimal to test i next.

3: $|i| = |j| - 1$ with i being a child of j. In this case

$$V(N_i^D) = \emptyset, V(N_i^{\overline{D}}) = \{j\}, V(N_j^D) = \{i\}, \quad \text{and } V(N_j^{\overline{D}}) = \emptyset \tag{47}$$

Consequently,

$$|N_i^D| + |N_i^{\overline{D}}| \le |N_j^D| + |N_j^{\overline{D}}| \tag{48}$$

and

$$|N_i^D| \leq |N_j^D| \leq |N_i^{\overline{D}}|, \quad |N_i^D| \leq |N_j^{\overline{D}}| \leq |N_i^{\overline{D}}| \tag{49}$$

Furthermore,

$$V_i(N) = c + p \cdot V(N_i^D) + (1-p) \cdot V(N_i^{\overline{D}}) = c + (1-p) \cdot c = c \cdot (2-p) \tag{50}$$

and

$$V_j(N) = c + p \cdot V(N_j^D) + (1-p) \cdot V(N_j^{\overline{D}}) = c + p \cdot c = c \cdot (1+p) \tag{51}$$

From (50), (51) and the assumption that $p \geq 1/2$ implying that $(2-p) \leq (1+p)$, we conclude that $V_i(N) \leq V_j(N)$.

- *Induction step* Assume that the assertion of the theorem is true for any set N' with cardinality k, i.e., for any set N' of cardinality less than or equal to k, if i and j subsets of sensors belonging to N' such that (38) and (39) are satisfied and $p \geq 1/2$ then (40) is true and an optimal test strategy is to test combinations of sensors in increasing order of cost. We need to prove that for any set N of cardinality $k+1$, the assertion of the theorem is still true. Let i and j be subsets of sensors belonging to N such that (38) and (39) are satisfied. We need to prove that $V_i(N) \leq V_j(N)$ and an optimal test strategy is to test combinations of sensors in increasing order of cost. To prove that $V_i(N) \leq V_j(N)$ we need to show that

$$\begin{aligned} V_i(N) = {} & c + p \cdot V(N_i^D) + (1-p) \cdot V(N_i^{\overline{D}}) \leq c + p \cdot V(N_j^D) + (1-p) \cdot V(N_j^{\overline{D}}) \\ & = V_j(N) \end{aligned} \tag{52}$$

or

$$p \cdot V(N_i^D) + (1-p) \cdot V(N_i^{\overline{D}}) \leq p \cdot V(N_j^D) + (1-p) \cdot V(N_j^{\overline{D}}) \tag{53}$$

Let u, v, r, and s denote, respectively, the cardinality of the sets $N_i^D, N_i^{\overline{D}}, N_j^D$, and $N_j^{\overline{D}}$. By assumption we have that

$$u + v \leq r + s, \quad \text{and } u \leq r \leq v, u \leq s \leq v \tag{54}$$

Since u, v, r, and s are all less than or equal to k, we obtain, by the induction hypothesis and *ii*

$$V(N_i^D) = c + (1-p) \cdot [c + (1-p) \cdot [c + \cdots] = c \cdot \sum_{n=0}^{u-1} (1-p)^n$$

and similarly,

$$V(N_i^{\overline{D}}) = c \cdot \sum_{n=0}^{v-1} (1-p)^n$$

$$V(N_j^{D}) = c \cdot \sum_{n=0}^{r-1} (1-p)^n$$

$$V(N_i^{\overline{D}}) = c \cdot \sum_{n=0}^{s-1} (1-p)^n \tag{55}$$

Proving Inequality (52) reduces to showing that

$$p \cdot c \cdot \sum_{n=0}^{u-1} (1-p)^n + (1-p) \cdot c \cdot \sum_{n=0}^{v-1} (1-p)^n \leq p \cdot c \cdot \sum_{n=0}^{r-1} (1-p)^n + (1-p) \cdot c \cdot \sum_{n=0}^{s-1} (1-p)^n \tag{56}$$

subject to the constraints (54). By rearranging terms in (56) and using the facts that $u \leq r$ and $s \leq v$, we find that proving (52) is equivalent to showing

$$(1-p) \cdot c \cdot \sum_{n=s}^{v-1} (1-p)^n \leq p \cdot c \cdot \sum_{n=u}^{r-1} (1-p)^n \tag{57}$$

By the assumption that $p \geq 1/2$ we have that $(1-p) \leq p$. Hence, to show that (57) is true it is sufficient to prove that the summation on the left hand side is less than the summation on the right hand side. The left hand side summation of (57) has fewer terms than the right hand side summation since $v - s \leq r - u$ (cf. (54)). Moreover, the terms in the summation on the left hand side are term by term less than the first $v - s$ terms in the summation on the right hand side: $s \geq u$ implies that $(1-p)^s \leq (1-p)^u$, i.e., the first term in the left hand side summation is smaller than the first term in the right hand side summation, and similarly for the remaining $v - s - 1$ terms. Therefore, we conclude that Inequality (57) is always true which implies that (52) is always true and this concludes the proof of the induction step for part *(i)* of the theorem. ■

Remark: For this problem the value function $V(N)$ is upper bounded by c/p. This can be seen as follows. First, it is clear that $V(N)$ is a nondecreasing function of the cardinality of the information state N. Since the a priori probability that the system possesses Property D when any combination of sensors is used is p, as cardinality $(N) \to \infty$ the average number of tests required to determine the least costly combination of sensors is $1/p$, according to the optimal test strategy of Theorem 3.2. Consequently $V(N)$ cannot exceed c/p for any information state N.

The following example illustrates, in part, Theorem 3.2.

Example 3.5: Consider the digraph presented in Figure 1 and let $N = 2^\Gamma \backslash \{\Gamma, \emptyset, \{1\}, \{2\}\}$ be the information state. Assume that the costs of combinations of sensors are ordered as follows

$$\{3\} < \{4\} < \{2,3\} < \{1,3\} < \{1,2\} < \{3,4\} < \{2,4\} < \{1,4\}$$
$$< \{1,2,3\} < \{2,3,4\} < \{1,3,4\} < \{1,2,4\}$$

and let $p = 0.6$. Pick $i = \{1,2\}$ and $j = \{1,3,4\}$. Then

$$\begin{aligned}
N_i^D &= \{\{3\},\{4\},\{2,3\},\{1,3\}\}\\
N_i^{\overline{D}} &= \{\{3\},\{4\},\{2,3\},\{1,3\},\{3,4\},\{2,4\},\{1,4\},\{1,2,3\},\{2,3,4\},\\
&\quad \{1,3,4\},\{1,2,4\}\}\\
N_j^D &= \{\{3\},\{4\},\{2,3\},\{1,3\},\{1,2\},\{3,4\},\{2,4\},\{1,4\},\{1,2,3\},\\
&\quad \{2,3,4\}\}\\
N_j^{\overline{D}} &= \{\{2,3\},\{1,2\},\{2,4\},\{1,2,3\},\{2,3,4\},\{1,2,4\}\}\\
V_i(N) &= c + p \cdot V(N_i^D) + (1-p) \cdot V(N_i^{\overline{D}})\\
\text{and } V_j(N) &= c + p \cdot V(N_j^D) + (1-p) \cdot V(N_j^{\overline{D}})
\end{aligned}$$

Since the above sets satisfy (38) and (39) of Theorem 3.2 part i, we must have $V_i(N) \le V_j(N)$. Indeed, by using (55), we obtain

$$V_i(N) = c + p \cdot c \cdot \sum_{n=0}^{3} (1-p)^n + (1-p) \cdot c \cdot \sum_{n=0}^{10} (1-p)^n = 2.6410 \cdot c$$
$$V_j(N) = c + p \cdot c \cdot \sum_{n=0}^{9} (1-p)^n + (1-p) \cdot c \cdot \sum_{n=0}^{5} (1-p)^n = 2.6689 \cdot c > V_i(N)$$

Both $V_i(N)$ and $V_j(N)$ are greater than the optimal cost which is equal to

$$c \cdot \sum_{n=0}^{11} (1-p)^n = 1.6666 \cdot c(\text{cf.}(41))$$

We conclude this section by showing that Assumption A3′ together with Assumptions A1, A2, A4, and A5 lead to an additional property of the value function $V(N)$ stated in Theorem 3.3 below. This property relates the cost function of two reachable sets, one being a subset of the other. Note that Corollary 3.1 presented a similar property under Assumption A3 provided that the conditions $p_l + p_{l+1} \ge 1$ hold for all l. Under Assumptions A1, A2, A3′, A4, and A5, Theorem 3.3 is valid for any value of p. Since no restrictions are imposed on p, the strategy that tests combinations of sensors in increasing order of cost is not necessarily an optimal test strategy. Therefore, Theorem 3.3 does not follow directly from the results of Theorems 3.1 and 3.2. Moreover, as presented in the proof, the arguments needed to establish Theorem 3.3 are different from those used in the proofs of Theorems 3.1 and 3.2.

THEOREM 3.3 *Consider the sensor selection problem under Assumptions A1, A2, A3′, A4, and A5. Let A and B be two reachable sets in the digraph. If* $A \subseteq B$ *then* $V(A) \leq V(B)$.

Proof of Theorem 3.3: We prove this Theorem by induction on the number of elements in A.

- Basis of induction: Let $A = \{a\}$, $|B| = 1 + k$, $k \geq 0$. We have $V(\{a\}) = c$, and $V(\{B\}) \geq c = V(\{a\}) = V(A)$.
- Induction step: Assume that $V(A) \leq V(B)$ for any two sets A and B such that $A \subseteq B$, $|A| = n$, $|B| = n + k$, for any $k \geq 0$. Pick any A', B' such that $A' \subseteq B'$, $|A'| = n + 1$, $|B'| = n + k + 1$. We need to show that $V(A') \leq V(B')$. Pick an optimal decision at state B', say x. Depending on whether $x \in A'$ or not there are two cases.

Case 1: $x \in A'$. The choice x may or may not be an optimal decision for A'. Therefore,

$$V(A') \leq c + p \cdot V(A_x'^{D}) + (1 - p) \cdot V(A_x'^{\overline{D}}) \tag{58}$$

and

$$V(B') = c + p \cdot V(B_x'^{D}) + (1 - p) \cdot V(B_x'^{\overline{D}}) \tag{59}$$

Furthermore,

$$A_x'^{D} \subseteq B_x'^{D}, \quad |A_x'^{D}| \leq n, \quad |B_x'^{D}| \leq n + k \tag{60}$$

and

$$A_x'^{\overline{D}} \subseteq B_x'^{\overline{D}}, \quad |A_x'^{\overline{D}}| \leq n, \quad |B_x'^{\overline{D}}| \leq n + k \tag{61}$$

by the properties of reachable sets and the fact that $A' \subseteq B'$. Therefore, by the induction hypothesis

$$V(A_x'^{D}) \leq V(B_x'^{D}) \tag{62}$$

and

$$V(A_x'^{\overline{D}}) \leq V(B_x'^{\overline{D}}) \tag{63}$$

which, together with (58)–(59), result in

$$V(A') \leq c + p \cdot V(B_x'^{D}) + (1 - p) \cdot V(B_x'^{\overline{D}}) = V(B') \tag{64}$$

Case 2: $x \notin A'$. Here we have two sub-cases:

Case 2-i: x or one of its parents was tested (before reaching state A') and proved not to possess Property D, i.e., x possesses property $\overline{D}$. Pick for A' a decision element, say x', such that x' is the lowest cost combination of sensors in A'. Since x' may not be the optimal decision for A',

$$V(A') \leq c + p \cdot V(A'^{D}_{x'}) + (1-p) \cdot V(A'^{\overline{D}}_{x'}) \tag{65}$$

furthermore, since x is an optimal decision at B' equation (59) holds. By assumption A' does not contain any child of x, which implies that $A'^{\overline{D}}_{x'}$ does not contain any child of x ($A'^{\overline{D}}_{x'} \subseteq A' \subseteq B'$), but $B'^{\overline{D}}_{x}$ contains all elements in B' except x and its children. Therefore,

$$A'^{\overline{D}}_{x'} \subseteq B'^{\overline{D}}_{x} \tag{66}$$

In addition

$$A'^{D}_{x'} = \emptyset \subseteq B'^{D}_{x} \tag{67}$$

since x' is the lowest cost combination of sensors in A'. By the induction hypothesis, since

$$|A'^{\overline{D}}_{x'}| \leq n, \quad |B'^{\overline{D}}_{x}| \leq n+k, \quad |A'^{D}_{x'}| \leq n, \quad \text{and } |B'^{D}_{x}| \leq n+k \tag{68}$$

we obtain

$$V(A'^{\overline{D}}_{x'}) \leq V(B'^{\overline{D}}_{x}) \tag{69}$$

and

$$V(A'^{D}_{x'}) \leq V(B'^{D}_{x}) \tag{70}$$

From (59), (65), (69), and (70) it follows that

$$V(A') \leq c + p \cdot V(B'^{D}_{x}) + (1-p) \cdot V(B'^{\overline{D}}_{x}) = V(B') \tag{71}$$

Case 2-ii: x was tested (before reaching state A') and proved to possess Property D. This implies that A' only contains elements whose number of sensors is strictly less than that of x in addition to those elements with the same number of sensors as x but with a lower cost. This case may be further divided into two sub-cases.

Case 2-ii-a: A' contains at least one child of x. Pick the child of x in A' having the highest number of sensors as a decision for A', say x''. Since x'' may or may not be an optimal decision at state A',

$$V(A') \leq c + p \cdot V(A'^{D}_{x''}) + (1-p) \cdot V(A'^{\overline{D}}_{x''}) \tag{72}$$

Furthermore, since x is an optimal decision at state B' Equation (59) holds. By the choice of x'' it follows that

$$A'^{D}_{x''} \subseteq B'^{D}_{x} \tag{73}$$

since $A'^{D}_{x''} \subseteq A' \subseteq B'$ and $A'^{D}_{x''}$ has only elements having number of sensors less than or equal that of x'', and consequently strictly less than that of x. By definition,

$$A'^{\overline{D}}_{x''} = A' \backslash \{x'' \text{ and its children}\} \tag{74}$$

and

$$B'^{\overline{D}}_{x} = B' \backslash \{x \text{ and its children}\} \tag{75}$$

and since $A' \subseteq B'$, and x'' is the child of x in A' with the maximum number of sensors,

$$A'^{\overline{D}}_{x''} \subseteq B'^{\overline{D}}_{x} \tag{76}$$

By the induction hypothesis, since

$$|A'^{\overline{D}}_{x''}| \leq n, \quad |B'^{\overline{D}}_{x}| \leq n+k, \quad |A'^{D}_{x''}| \leq n, \quad \text{and } |B'^{D}_{x}| \leq n+k \tag{77}$$

we obtain

$$V(A'^{\overline{D}}_{x''}) \leq V(B'^{\overline{D}}_{x}) \tag{78}$$

and

$$V(A'^{D}_{x''}) \leq V(B'^{D}_{x}) \tag{79}$$

Consequently, from (59), (72), (78), and (79) we conclude that

$$V(A') \leq c + p \cdot V(B'^{D}_{x}) + (1-p) \cdot V(B'^{\overline{D}}_{x}) = V(B') \tag{80}$$

Case 2-ii-b: A' contains no children of x. Pick the element in A' that has the highest cost, say x'''. Since x''' may or may not be an optimal decision,

$$V(A') \leq c + p \cdot V(A'^{D}_{x'''}) + (1-p) \cdot V(A'^{\overline{D}}_{x'''}) \tag{81}$$

Furthermore, since x is an optimal decision at state B equation (59) holds. The choice of x''' implies that

$$A'^{D}_{x'''} = A' \backslash \{x'''\} \tag{82}$$

i.e., $A'^{D}_{x'''}$ only contains elements in A' that have cost strictly lower than that of x'''. Moreover, by assumption x does not belong to A', and it was tested before reaching A' and proved to possess Property D; hence, all elements in A' and consequently all elements in $A'^{D}_{x'''}$ have strictly lower cost than that of x. Furthermore, B'^{D}_{x} has all of its elements as those elements in B' whose cost is strictly less than that of x. Therefore, since $A' \subseteq B'$, it follows that

$$A'^{D}_{x'''} \subseteq B'^{D}_{x} \tag{83}$$

In addition,

$$A'^{\overline{D}}_{x'''} \subseteq B'^{\overline{D}}_{x} \tag{84}$$

since $A'^{\overline{D}}_{x'''} \subseteq A' \subseteq B'$ and the only elements missing from B' in $B'^{\overline{D}}_{x}$ are x and its children, which are not in $A'^{\overline{D}}_{x'''}$ by the assumption that A' contains no children of x. By the induction hypothesis, since

$$|A'^{\overline{D}}_{x'''}| \leq n, \quad |B'^{\overline{D}}_{x}| \leq n+k, \quad |A'^{D}_{x'''}| \leq n, \quad \text{and } |B'^{D}_{x}| \leq n+k \tag{85}$$

we obtain

$$V(A'^{\overline{D}}_{x'''}) \leq V(B'^{\overline{D}}_{x}) \tag{86}$$

and

$$V(A'^{D}_{x'''}) \leq V(B'^{D}_{x}) \tag{87}$$

From (59), (81), (86), and (87) we obtain

$$V(A') \leq c + p \cdot V(B'^{D}_{x}) + (1-p) \cdot V(B'^{\overline{D}}_{x}) = V(B') \tag{88}$$

The proof of the induction step is complete, and this completes the proof of Theorem 3.3. ■

4. Concluding Remarks

We have presented conditions sufficient to guarantee that the strategy that tests combinations of sensors in increasing order of cost is optimal for Problem P. This result is intuitive since it is assumed that a combination of two or more sensors is equally or more likely to result in a system possessing Property D than a system possessing Property $\overline{D}$ and the cost of testing for Property D is independent of the cardinality of the combination of sensors that is being used.

Assumptions A2–A5 as well as $p_l + p_{l+1} \geq 1$, $l = 1, 2, \ldots, M-2$, are crucial in establishing the result of Theorem 3.1. A brief critique of these assumptions was presented in Section 2. We add a few more remarks that further reveal the importance of these assumptions. We have already noted that Assumption A2 is reasonable but it may not always be true. Relaxation of Assumption A2 may drastically change the structure of the optimization problem because ordering of sensor combinations according to their cost is not the same as the ordering according to their cardinality. Therefore, it may be unlikely that the strategy described in Theorem 3.1 will be optimal; even if the same strategy were optimal, a proof of its optimality would require arguments different from those employed in the proof of Theorem 3.1. If the sensors' information gathering ability drastically differs from sensor to sensor, then Assumption A3 may not be true. In this case, an optimal test strategy for the sensor selection problem may be different from that of Theorem 3.1 and arguments significantly different from those presented in Theorem 3.1 may be required to discover the nature of an optimal test strategy. The conditions $p_l + p_{l+1} \geq 1$, $l = 1, 2, \ldots, M-2$ are essential for the validity of the result of Theorem 3.1; the plausibility of the main result under these conditions has already been discussed. Finally, Assumption A3$'$ is stronger than A3. This is why when Assumption A3$'$ is used instead of A3 we obtain more structural results about Problem P. Specifically, under Assumptions A1–A5 we have proved the optimality of the strategy that tests combinations of sensors in increasing order of cost, but have been unable to compare any other policies; on the contrary, under Assumptions A1, A2, A3$'$, A4, and A5 we can compare policies that satisfy the conditions of Theorem 3.2.

We conclude by noting that although the methodology presented in this paper was motivated by the failure diagnosis and hypothesis testing examples, it can be used in other situations as long as the property to be tested exhibits the same characteristics as Property *D*, namely C3 and C4. For example, in the field of discrete event systems, the properties of controllability, observability and normality possess C3 and C4 (see Cassandras and Lafortune, 1999).

Acknowledgments

The authors would like to thank the reviewers for their comments and suggestions which helped improve the presentation and readability of the paper. The first author is grateful for the time provided by General Motors Corporation for the work on revising the manuscript.

Appendices

A. Computation of Equations (26), (27), and (28)

Assume that the three lowest-cost combinations of sensors in the set $N_i^{\overline{D}}$ (cf. (25)) are a, b, and d in that order, and a and b are combinations of i sensors while d is a combination of $i+1$ sensors. Since the cardinality of $N_i^{\overline{D}}$ is less than or equal to k we can apply the induction hypothesis to the set $N_i^{\overline{D}}$, i.e., we test next the element with the smallest cost. Therefore, we have

$$V(N_i^{\overline{D}}) = c + (1 - p_i) \cdot V(N_i^{\overline{D}} \backslash \{a\}) \tag{89}$$

By applying the induction hypothesis to $N_i^{\overline{D}} \backslash \{a\}$ in (89), we have

$$V(N_i^{\overline{D}}) = c + (1 - p_i) \cdot c + (1 - p_i)^2 \cdot V(N_i^{\overline{D}} \backslash \{a, b\}) \tag{90}$$

By applying the induction hypothesis to $N_i^{\overline{D}} \backslash \{a, b\}$ in (90), we have

$$V(N_i^{\overline{D}}) = c + (1 - p_i) \cdot c + (1 - p_i)^2 \cdot c + (1 - p_i)^2 \cdot (1 - p_j) \cdot V(N_i^{\overline{D}} \backslash \{a, b, d\}) \tag{91}$$

Finally, by successively applying the induction hypothesis until exhausting all elements in $N_i^{\overline{D}}$ we get (26). Following a similar procedure, one can verify (27) and (28).

B. *Verification of (30): Properties of Reachable Sets*

The reachable sets from (25) are $N_i^{\overline{D}}$, N_j^D, and $N_j^{\overline{D}}$. The properties of these reachable sets are listed below (reproducing (30))

$$r_j < v_j \tag{92}$$

$$r_k = v_k \quad k = |i| + 1, \ldots, |j| - 1 \tag{93}$$

$$r_i = v_i + 1 \tag{94}$$

$$s_k = v_k \quad k = |j| + 1, \ldots, L \tag{95}$$

$$s_j = v_j - 1 \tag{96}$$

$$s_k \leq v_k \quad k = |i| + 1, \ldots, |j| - 1 \tag{97}$$

$$s_i \leq v_i + 1 \tag{98}$$

where v_k denotes the number of combinations of k sensors in $N_i^{\overline{D}}$, r_k denotes the number of combinations of k sensors in N_j^D, s_k denotes the number of combinations of k sensors in $N_j^{\overline{D}}$, and L denotes the highest set cardinality in N with $|j| \leq L \leq K - 1$. We need to verify (93)–(98) since we verified (92) in Section 3.3.

From (25), the set $N_i^{\overline{D}}$ is equal to the set N without the element i. The set N_j^D includes all element in N except the elements with more than $|j|$ sensors and those with $|j|$ sensors but with strictly higher cost than j. The set $N_j^{\overline{D}}$ includes all elements in N except j and its children. The elements of $N_i^{\overline{D}}$ and N_j^{D} with more than $|i|$ sensors and less than $|j|$ sensors are the same, hence (93) is true. N_j^{D} has all elements in N with $|i|$ sensors while $N_i^{\overline{D}}$ includes the same elements with the exception of the element i, and that verifies (94). The sets $N_i^{\overline{D}}$ and $N_j^{\overline{D}}$ share the same elements with more than $|j|+1$ sensors, and that verifies (95). The only element of $|j|$ sensors missing from $N_j^{\overline{D}}$ and belonging to $N_i^{\overline{D}}$ is the element j, hence (96) is true. Inequality (97) is true since $N_i^{\overline{D}}$ contains all elements in N with more than $|i|$ sensors and less than $|j|$ sensors, while $N_j^{\overline{D}}$ contains the same elements with the exception of those that are children of the element j. The same argument applies to the elements with $|i|$ sensors with the possible addition of the element i, which does not belong to $N_i^{\overline{D}}$, to the set $N_j^{\overline{D}}$, and that verifies (98).

References

Bavishi, S., and Chong, E. 1994. Automated fault diagnosis using a discrete event systems framework. In *Proc. 9th IEEE International Symposium on Intelligent Control*, pp. 213–218.

Cassandras, C. G., and Lafortune, S. 1999. *Introduction to Discrete Event Systems*. Boston, MA: Kluwer Academic Publishers.

Darabi, H., and Jafari, M. A. 1998. Supervisory control and optimal partial observation problem. Preprint, Rutgers University.

Debouk, R., Lafortune, S., and Teneketzis, D. 2000. Coordinated decentralized protocols for failure diagnosis of discrete-event systems. *Journal of Discrete Event Dynamical Systems: Theory and Applications* 10: 33–86.

Haji-Valizadeh, A., and Loparo, K. 1996. Minimizing the cardinality of an events set for supervisors of discrete-event dynamical systems. *IEEE Trans. Automat. Contr.* 41(11): 1579–1593.

Kumar, P. V., and Varaiya, P. 1986. *Stochastic Systems: Estimation, Identification, and Adpative Control*. Englewood Cliffs, NJ: Prentice Hall.

Puterman, M. L. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY: John Wiley & Sons.

Sampath, M., Sengupta, R., Lafortune, S., Sinnamohideen, K., and Teneketzis, D. 1995. Diagnosability of discrete-event systems. *IEEE Trans. Automat. Contr.* 40(9): 1555–1575.

Sampath, M., Sengupta, R., Lafortune, S., Sinnamohideen, K., and Teneketzis, D. 1996. Failure diagnosis using discrete-event models. *IEEE Trans. Contr. Syst. Tech.* 4(2): 105–124.

Sampath, M., Lafortune, S., and Teneketzis, D. 1998. Active diagnosis of discrete-event systems. *IEEE Trans. Automat. Contr.* 43(7): 908–929.