Enhancing Automated Test Selection in Probabilistic Networks

Danielle Sent[1] and Linda C. van der Gaag[2]

[1] Department of Electrical Engineering, Mathematics and Computer Science, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
danielle.sent@utwente.nl

[2] Department of Information and Computing Sciences, Utrecht University, P.O. Box 80.089, 3508 TB Utrecht, The Netherlands
linda@cs.uu.nl

Abstract. Most test-selection algorithms currently in use with probabilistic networks select variables myopically, that is, test variables are selected sequentially, on a one-by-one basis, based upon expected information gain. While myopic test selection is not realistic for many medical applications, non-myopic test selection, in which information gain would be computed for all combinations of variables, would be too demanding. We present three new test-selection algorithms for probabilistic networks, which all employ knowledge-based clusterings of variables; these are a myopic algorithm, a non-myopic algorithm and a semi-myopic algorithm. In a preliminary evaluation study, the semi-myopic algorithm proved to generate a satisfactory test strategy, with little computational burden.

Keywords: diagnostic test selection, probabilistic networks, semi-myopia.

1 Introduction

To support the entire process of a patient's management, a decision-support system should not only provide information about the most probable diseases or the best suitable therapy, it should also provide information about which diagnostic tests had best be performed to reduce the uncertainty about a patient's condition. In the context of probabilistic networks, an automated test-selection facility is usually composed of an information measure, a test-selection loop, and a criterion for deciding when to stop gathering further information. The information measure is defined on the probability distribution over the main diagnostic variable and essentially captures diagnostic uncertainty. With respect to the actual test-selection loop, most algorithms in use with probabilistic networks serve to select diagnostic tests myopically [2]. In each iteration, the most informative variable is selected from among all possible test variables to indicate the next test to perform. The user is prompted for the value of the selected variable, which is entered into the network and propagated to establish the posterior probabilities for all variables. From the set of test variables still available, the next variable is selected. This process of selecting test variables and propagating their results is continued until a stopping criterion is met or until results for all test variables have been entered.

We feel that the test-selection strategy that is induced by a myopic algorithm is an oversimplification of the problem-solving strategies found in many fields of medicine. Based upon interviews with two experts in the field of oesophageal cancer, we in fact

R. Bellazzi, A. Abu-Hanna, and J. Hunter (Eds.): AIME 2007, LNAI 4594, pp. 331–335, 2007.

identified several aspects where myopic test selection does not match daily routines. In the strategy of our experts, different subgoals are identified that are addressed sequentially, such as discovering the characteristics of the primary tumour and establishing the absence or presence of metastases. We feel that a more involved test-selection facility should take such subgoals into account. Moreover, our experts order tests in packages to reduce the length in time of the diagnostic phase of a patient's management. For the latter purpose, especially, a non-myopic algorithm would be required in which in each step multiple tests can be selected. A fully non-myopic algorithm is computationally very demanding, however, and may easily prove infeasible for practical purposes. Based upon these considerations, we present in this paper three new test-selection algorithms that take a fixed clustering of test variables into account. These algorithms retain some of the idea of non-myopia, yet stay computationally feasible.

The paper is organised as follows. Section 2 reviews the basic test-selection algorithm currently in use with probabilistic networks. Section 3 presents our new algorithms for test selection. In Section 4 we briefly describe the experiments that we conducted with our new algorithms. The paper ends with our conclusions in Section 5.

2 Preliminaries

Before presenting our new algorithms for test selection with probabilistic networks, we briefly review the myopic algorithm in use for this purpose [1,3,4]. This algorithm takes for its input a set $\mathscr{T}$ of test variables. For its output, it sequentially prompts the user to supply a value for a selected variable $T_i \in \mathscr{T}$. The value entered by the user then is propagated through the network at hand before the next variable is selected and presented to the user. The algorithm amounts to the following in pseudo-code:

Myopic test selection
input: $\mathscr{T}$ is a list of test variables T_i
Stop $=$ *false*
while $\mathscr{T} \neq \varnothing$ and *Stop* $\neq$ *true* **do**
    compute most informative $T_i \in \mathscr{T}$ and remove T_i from $\mathscr{T}$
    prompt for evidence for T_i and propagate
    compute *Stop*
od

We assume that the algorithm employs the Gini index of the probability distribution over the disease variable; other information measures can also be used, however. The *Gini index* $G(\Pr(D))$ of the probability distribution Pr over the diagnostic variable D is defined as

$$G(\Pr(D)) = 1 - \sum_{j=1,\ldots,m} \Pr(D = d_j)^2$$

The expected Gini index $G(\Pr(D \mid T_i))$ after obtaining a value for the variable T_i is defined as the expected value of the Gini index where the expectation is taken over all possible values:

$$G(\Pr(D \mid T_i)) = \sum_{k=1,\ldots,m_i} G(\Pr(D \mid T_i = t_i^k)) \cdot \Pr(T_i = t_i^k)$$

The best test variable to select is one that maximises the decrease $G(\Pr(D)) - G(\Pr(D \mid T_i))$ in diagnostic uncertainty. From a computational point of view, the most expensive step in the algorithm is that in which the most informative variable is selected. In this step, the probability distributions $\Pr(D \mid T_i)$ are computed for all test variables T_i. Using Bayes' rule, the number of propagations required equals the number of values of D.

3 Enhanced Test-Selection Algorithms

The test-selection strategy implied by the basic myopic algorithm seems to be an oversimplification of the test-selection routines found in many fields of medicine. In the domain of oesophageal cancer, for example, we found that physicians order tests for specific subgoals. They start gathering general information about the patient and about the tumour. Having this information, they focus on establishing the presence or absence of distant metastases and order tests accordingly. The physicians further order physical tests such as a CT-scan, even though the results of the scan are modelled in the network by multiple variables such as *CT-liver*, *CT-loco*, *CT-lungs*, *CT-organs*, and *CT-truncus*.

To arrive at a test-selection facility that fits in more closely with daily practice, we enhance the basic myopic algorithm to take a list $\mathscr{S}$ of subgoals S_i into consideration. The algorithm performs test selection per subgoal, that is, for each subgoal it focuses on the test variables that provide information about that particular goal. For this purpose, the algorithm is provided with a list of subsets $\mathscr{T}(S_i)$ of $\mathscr{T}$, each of which includes all test variables that pertain to a specific goal. Our first algorithm now computes the most informative test to be performed by investigating single test variables. The user is prompted for just the selected test variable and only the evidence for this variable is propagated throughout the network, before the test-selection process is continued:

Algorithm A_1: myopic test selection with subgoals
input: $\mathscr{S}$ is a list of subgoals S_i,
    $\mathscr{T}$ is a list of test variables T_j, organised in sublists $\mathscr{T}(S_i)$ per subgoal S_i
Stop-subgoal(S_i), *Stop-overall* $=$ *false*
while $\mathscr{S} \neq \varnothing$ and *Stop-overall* $\neq$ *true* **do**
    select next S_i from $\mathscr{S}$ and remove S_i from $\mathscr{S}$
    while $\mathscr{T}(S_i) \neq \varnothing$ and *Stop-subgoal*(S_i), *Stop-overall* $\neq$ *true* **do**
        compute most informative $T_j \in \mathscr{T}(S_i)$ and remove T_j from $\mathscr{T}$
        prompt for evidence for T_j and propagate
        compute *Stop-subgoal*(S_i), *Stop-overall*
    od
od

The algorithm selects a subgoal S_i from the list of subgoals. From the associated set of test variables, it selects the variable T_j that is expected to yield the largest decrease in diagnostic uncertainty. The user is prompted to enter evidence for T_j, which is subsequently propagated through the network. The process of selecting test variables continues until the stopping criterion for the subgoal S_i or that for the overall goal has been met, or all tests for S_i have been performed. When the stopping criterion for S_i is satisfied or its set of test variables has been exhausted, the algorithm selects the next subgoal. As soon as the overall stopping criterion is satisfied, the entire process is halted.

Algorithm A_1 is still strictly myopic: test variables are selected sequentially on a one-by-one basis and the next variable is selected only after the user has entered evidence for the previous one. We have argued above that a myopic test-selection strategy may not be realistic for many applications in medicine. A fully non-myopic algorithm, in which the expected Gini index given every possible subset of test variables is established, on the other hand, may be infeasible for practical purposes. Our second algorithm now is non-myopic in nature, yet uses a predefined clustering of the test variables where each cluster is associated with a single physical test. The clustering of the test variables is given as part of the input to the algorithm:

Algorithm A_2: non-myopic test selection with subgoals
input: $\mathscr{S}$ is a list of subgoals S_i,
  $\mathscr{T}$ is a list of clusters C_j of test variables, organised in sublists $\mathscr{T}(S_i)$ per subgoal
Stop-subgoal(S_i), *Stop-overall* $=$ *false*
while $\mathscr{S} \neq \varnothing$ and *Stop-overall* $\neq$ *true* **do**
  select next S_i from $\mathscr{S}$ and remove S_i from $\mathscr{S}$
  while $\mathscr{T}(S_i) \neq \varnothing$ and *Stop-subgoal*(S_i), *Stop-overall* $\neq$ *true* **do**
    compute most informative cluster $C_j \in \mathscr{T}(S_i)$ and remove C_j from $\mathscr{T}$
    prompt for evidence for all $T_k \in C_j$ and propagate
    compute *Stop-subgoal*(S_i), *Stop-overall*
  od
od

Again driven by subgoals, the algorithm selects the cluster C_j of variables that is expected to yield the largest decrease in uncertainty. The user is prompted for evidence for each separate variable T_k from the cluster. We note that algorithm A_2 is much more computationally demanding than algorithm A_1. The increase in computation time stems from computing the most informative cluster. To this end, the probability distributions $\Pr(D \mid C_j = c)$ and $\Pr(C_j = c)$ are computed for all combinations of values c of the test variables in C_j, which requires an exponential number of propagations.

Algorithm A_2 in essence is non-myopic in its test-selection strategy and may become computationally too demanding if a meaningful clustering would result in clusters of relatively large size. To save computation time yet retain some of the idea of non-myopia, we designed an algorithm that implies a semi-myopic test-selection strategy:

Algorithm A_3: semi-myopic test selection with subgoals
input: $\mathscr{S}$ is a list of subgoals S_i,
  $\mathscr{T}$ is a list of clusters C_j of test variables, organised in sublists $\mathscr{T}(S_i)$ per subgoal
Stop-subgoal(S_i), *Stop-overall* $=$ *false*
while $\mathscr{S} \neq \varnothing$ and *Stop-overall* $\neq$ *true* **do**
  select S_i from $\mathscr{S}$ and remove S_i from $\mathscr{S}$
  while $\mathscr{T}(S_i) \neq \varnothing$ and *Stop-subgoal*(S_i), *Stop-overall* $\neq$ *true* **do**
    compute most informative $T_j \in \mathscr{T}(S_i)$
    prompt for evidence for T_j and for all $T_k \in C_m$ with C_m such that $T_j \in C_m$,
      propagate and remove C_m from $\mathscr{T}$
    compute *Stop-subgoal*(S_i), *Stop-overall*
  od
od

Algorithm A_3 very much resembles the myopic algorithm A_1 presented above. Driven by subgoals, it selects the variable T_j that is expected to yield the largest decrease in diagnostic uncertainty. The main difference is, however, that algorithm A_3 prompts not just for evidence for T_j, but also for evidence for all test variables T_k that belong to the same cluster as T_j. Entering evidence for physical tests rather than for just one test variable fits in more closely with the daily routines of the physicians. Physicians think in terms of physical tests even when they are interested mainly in the value of a single variable. After performing the test, therefore, it seems logical to enter not just the result that is currently of interest, but all other results obtained from the same test as well.

4 Preliminary Experimental Results

To compare the performance of the three algorithms for test selection described above, we conducted a preliminary experimental study in the domain of oesophageal cancer. We found that all three algorithms resulted in rather similar sequences of tests, with just occasional differences. To explain this finding, we observe that, if a single test variable is expected to result in a large decrease in diagnostic uncertainty, then it is likely that the test to which it pertains will be quite informative as well. We presented the sequences of test variables constructed by the algorithms to our domain experts. They indicated that they felt most comfortable with the sequences generated by the semi-myopic algorithm. They indicated more specifically that the sequence generated by the myopic algorithm appeared somewhat unnatural.

5 Conclusions

Most test-selection algorithms currently in use with probabilistic networks select variables myopically. We argued that, while myopic test selection is not realistic for many medical applications, non-myopic test selection would be too demanding. We presented new test-selection algorithms which all employ knowledge-based clusterings of variables. Both from the perspective of fitting in with physicians' daily routines and from a computational perspective, we feel that our semi-myopic algorithm provides an appropriate mean by introducing a concept of restricted non-myopia.

References

1. Andreassen, S.: Planning of therapy and tests in causal probabilistic networks. Artificial Intelligence in Medicine 4, 227–241 (1992)
2. Ben-Bassat, M.: Myopic policies in sequential classification. IEEE Transactions on Computers 27(2), 170–174 (1978)
3. Doubilet, P.: A mathematical approach to interpretation and selection of diagnostic tests. Medical Decision Making 3(2), 177–195 (1983)
4. Glasziou, P., Hilden, J.: Test selection measures. Medical Decision Making 9, 133–141 (1989)