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, 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, the semi-myopic algorithm proved to generate a satisfactory test strategy, with little computational burden.

keywords: diagnostic test selection, probabilistic network, semi-myopia

1 Introduction

Since over the last decades researchers have come to understand more and more of diseases and their management, it nowadays is hard, even for medical specialists, to keep up-to-date with medical literature and with new insights of diseases, new drugs and new procedures. On the other hand, the costs of more sophisticated treatments, tests and procedures are increasing. To cope with the increasing complexity of medical practice, decision-support systems can assist physicians in their complex problem-solving tasks, by providing support that is tailored to individual patients.

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 its user with information about which diagnostic tests should be performed. These diagnostic tests then are performed to reduce the uncertainty about a patient's true condition. In this paper, we focus on algorithms for test selection in view of probabilistic networks, where a probabilistic network is a concise representation of a joint probability distribution Pr on a set of stochastic variables.

An automated test-selection facility for a decision-support system is composed of three basic components: an information measure for capturing the uncertainty about the overall diagnosis, an actual test-selection loop, and an associated criterion for deciding when to stop gathering further information. In the context of a probabilistic network, the information measure is typically defined on the probability distribution over the main diagnostic variable. For test-selection purposes, both the Shannon entropy and the

Gini index are suitable information measures for use in a test-selection facility [5]. With respect to the test-selection loop, most algorithms in use in practical decision-support systems serve to select diagnostic tests so-called myopically [2]. In each iteration of the loop, the most informative test variable is selected from among all possible test variables to indicate the next test to perform. The user is then prompted for the value of the selected variable. The result is entered into the probabilistic network of the system 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 into the network.

We feel that the test-selection strategy that is induced by a myopic algorithm is an oversimplification of the experts' problem-solving strategy in many medical fields of application. Based upon a series of interviews we had with two experts in the domain of oesophageal cancer, we identified several aspects why myopic test selection does not fit the daily test-selection routine of our experts. In the strategy of our experts, different subgoals can be identified that are addressed sequentially, such as establishing the characteristics of the primary tumour and deciding upon the absence or presence of metastases. We feel that a more involved test-selection facility should be able to take such subgoals into account. In addition, multiple test variables serve to model the results of a single physical test. It will be evident that physicians do not order stochastic variables but physical tests. Moreover, our experts have been found to 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 test-selection algorithm would be required in which in each step multiple tests can be selected. A fully non-myopic algorithm is computationally very demanding and, in fact, may easily prove to be too much demanding for practical purposes. With respect to the first consideration for non-myopia, we observe that an algorithm that takes a fixed clustering of test variables into account, would suffice. Such an algorithm would retain some of the idea of non-myopia, yet stay computationally feasible.

A test-selection facility should result in a strategy in which not too many tests are selected. Even more importantly, it should prevent the test-selection process from halting too soon. To prevent overtesting, a test-selection algorithm is typically extended with a stopping criterion. With such a criterion the algorithm computes if performing more diagnostic tests is necessary or that the physician can safely stop testing.

In this paper, we address test selection in probabilistic networks. Section 2 describes myopic test selection in general. Section 3 presents three new algorithms for test selection that employ subgoals and restricted concepts of non-myopia. In Section 4 we describe the experiments that we conducted with our new algorithms. This chapter ends with our conclusions in Section 5.

2   Preliminaries

Before presenting our new algorithms for test selection in probabilistic networks, we briefly review the basic myopic algorithm [1, 3, 4]. This algorithm takes for its input a set $\mathscr{T}$ of test variables. For its output, the algorithm sequentially prompts the user to

supply a value for a selected variable $T_i \in \mathcal{T}$. The value entered by the user then is propagated through the probabilistic network under consideration before the next variable is selected and presented to the user. The basic myopic algorithm now amounts to the following in pseudo-code:

Myopic test selection

input: $\mathcal{T}$: list of test variables T_i

Stop $=$ *false*
while $\mathcal{T} \neq \varnothing$ and *Stop* $\neq$ *true* **do**
      compute most informative $T_i \in \mathcal{T}$
      prompt for evidence for T_i and propagate
      remove T_i from $\mathcal{T}$
      compute *Stop*
od

We assume from now on that the algorithm employs the Gini index of the probability distribution over the disease variable. Other information measures can be used, however. The *Gini index G* 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 performing a diagnostic test T_i is defined as the expected value of the Gini index where the expectation is taken over all possible test results:

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

The best test to perform is a test that is expected to result in the largest decrease of diagnostic uncertainty, that is, a test that maximises $\widetilde{G}(T_i) = G(\Pr(D)) - G(\Pr(D \mid T_i))$.

From a computational point of view, the most expensive step in the algorithm is the step in which the most informative test variable is selected. In this step, the Gini index $G(\Pr(D \mid T_i))$ has to be computed for each test variable T_i. Straightforward computation of all conditional probability distributions $\Pr(D \mid T_i)$ would require a single propagation for each value of each test variable, which would clearly be too much demanding for practical purposes. We note that the probability distributions $\Pr(D \mid T_i)$ required for this purpose can be computed efficiently using Bayes' rule:

$$\Pr(D \mid T_i) = \Pr(T_i \mid D) \cdot \frac{\Pr(D)}{\Pr(T_i)}$$

The probability distributions $\Pr(D)$ and $\Pr(T_i)$ for each test variable T_i are already available after the last test result had been entered and therefore do not require any further propagations. The probability distributions $\Pr(T_i \mid D)$ for all test variables can be established by propagating the various possible values for the disease variable D. The number of propagations required to establish these distributions for all test variables simultaneously thus 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 over-simplification of the test-selection routines that are found in many medical fields of application. From the domain of oesophageal cancer, for example, we found that they order tests for specific subgoals. They start gathering general information about the patient and the tumour. After gathering this information, they focus on the presence or absence of distant metastases and order tests for this specific purpose. 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*. Note that myopic test selection would consider these seperate variables and select only one of these.

To arrive at a test-selection facility that fits in more closely with daily practice, we now first enhance the basic myopic algorithm to take a sequence of different subgoals into consideration. To this end, the algorithm is extended to take a list $\mathscr{S}$ of subgoals S_i as part of its input. The algorithm now performs test selection per subgoal, that is, for each current subgoal it focuses on the test variables that provide information about that particular subgoal. For this purpose, the algorithm has to be provided with information about which variables provide information about which subgoals. Since this type of information is rather easily given by domain experts, we decided to provide it as part of the input to the algorithm. For every subgoal S_i, all test variables that provide information about that specific subgoal, are included in the subset $\mathscr{T}(S_i)$ of $\mathscr{T}$. We would like to note that the subsets $\mathscr{T}(S_i)$ do not have to be mutually disjoint: test variables may provide information about various subgoals.

In the presence of subgoals, we further need to reconsider the stopping criterion that is to be used by the algorithm. We recall that such a criterion is included to prevent overtesting. The stopping criterion should now take the subgoals into consideration as well as the overall goal. While selecting test variables for a specific subgoal, for example, the algorithm should be allowed to move to the next subgoal when no further information is needed with regard to the first subgoal even though there still are some tests left for that goal. Also, if for a specific subgoal the stopping criterion indicates that the test-selection process should continue, the algorithm should be allowed to stop selecting tests altogether if the value of the overall goal indicates that further testing is not necessary. We will use two separate variables in our algorithms to provide for a stopping criterion that takes the subgoals and the overall goal into consideration.

Our first algorithm now computes the most informative test to be performed at the level of 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:$\mathscr{S}$: list of subgoals S_i;
$\mathscr{T}$: list of test variables T_j, organised in sublists $\mathscr{T}(S_i)$ per subgoal S_i

Stop-subgoal(S_i)=false
Stop-overall=false

while $\mathscr{S} \neq \varnothing$ and *Stop-overall≠true* **do**
      select next S_i from $\mathscr{S}$
      remove S_i from $\mathscr{S}$
      while $\mathscr{T}(S_i) \neq \varnothing$, *Stop-subgoal($S_i$)≠true* and *Stop-overall≠true* **do**
            compute most informative $T_j \in \mathscr{T}(S_i)$
            prompt for evidence for T_j and propagate
            remove T_j from $\mathscr{T}$
            compute *Stop-subgoal(S_i)*
            compute *Stop-overall*
      od
od

The algorithm proceeds as follows. It selects the first subgoal from the list $\mathscr{S}$ of subgoals. It then selects the set of all test variables that provide information about the selected subgoal. From among these test variables, it selects the variable that is expected to yield the largest decrease in diagnostic uncertainty. The user then is prompted to enter evidence for the selected variable T_j. The entered evidence is propagated throughout the network, resulting in the posterior probability distribution over all variables given the entered evidence. Note that the evidence may have changed the probability distributions for the main diagnostic variable and for the variable representing the subgoal under study. The process of selecting test variables continues until the stopping criterion for the subgoal under study or that for the overall goal has been met, or all tests for the subgoal have been performed. When the stopping criterion for the subgoal is satisfied, or the set of test variables has been exhausted, the algorithm selects the next subgoal. As soon as the overall stopping criterion is satisfied, the test-selection process is halted.

Algorithm A_1 is still strictly myopic in its test-selection strategy: test variables are selected sequentially on a one-by-one basis and the next test 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. The clustering of the test variables in the network is given as part of the input to the algorithm. The algorithm equals

Algorithm A_2: non-myopic test selection with subgoals

input:$\mathscr{S}$: list of subgoals S_i;
$\mathscr{T}$: list of clusters C_j of test variables, organised in sublists $\mathscr{T}(S_i)$ per subgoal

Stop-subgoal(S_i)=false
Stop-overall=false
while $\mathscr{S} \neq \varnothing$ and *Stop-overall≠true* **do**
      select next S_i from $\mathscr{S}$
      remove S_i from $\mathscr{S}$
      while $\mathscr{T}(S_i) \neq \varnothing$, *Stop-subgoal($S_i$)≠true* and *Stop-overall≠true* **do**
            compute most informative cluster $C_j \in \mathscr{T}(S_i)$
            prompt for evidence for all $T_k \in C_j$ and propagate
            remove C_j from $\mathscr{T}$

compute *Stop-subgoal*(S_i)
　　　　compute *Stop-overall*
　　od
od

We would like to note that algorithm A_2 is much more computationally demanding than algorithm A_1. The increase of computation time stems from the computation of the most informative cluster of test variables. We recall from Section 2 that the expected Gini index $G(\Pr(D \mid T_j))$ for a single test variable T_j is defined as the expected value of the Gini index where the expectation is taken over all possible test results t_j^k:

$$G(\Pr(D \mid T_j)) = \sum_{k=1,\ldots,m_j} G(\Pr(D \mid T_j = t_j^k)) \cdot \Pr(T_j = t_j^k)$$

In algorithm A_2, we need to compute the expected Gini index with respect to a cluster C_j of test variables. To this end, the Gini index $G(\Pr(D \mid C_j = c))$ has to be computed for each combination of values $c = t_1^{k_1}, \ldots, t_n^{k_n}$ of the test variables $T_1, \ldots, T_n \in C_j$. In addition, the joint probability $\Pr(C_j = c)$ has to be established for each such combination. The probability distribution $\Pr(D \mid C_j = c)$ for a specific combination of values c can be computed by a single propagation throughout the network. The computation of all distributions $\Pr(D \mid C_j = c)$ now requires at most 2^n propagations for n binary test variables; note that using the concept of d-separation, this number of propagations can often be reduced in practice. The computation of each separate probability $\Pr(C_j = c)$ already requires n propagations, building upon the chain rule $\Pr(t_1^{k_1}, \ldots, t_n^{k_n}) = \Pr(t_1^{k_1} \mid t_2^k, \ldots, t_n^k) \cdot \ldots \cdot \Pr(t_{n-1}^{k_{n-1}} \mid t_n^{k_n}) \cdot \Pr(t_n^{k_n})$. The computation of the entire joint distribution $\Pr(C_j)$ thus requires at most $n \cdot 2^n$ propagations. The computation of the expected Gini index with respect to clusters of test variables takes an exponential number of propagations, while the computation of the expected index with respect to all variables separately takes just a constant number of propagations. We conclude that algorithm A_2 rapidly becomes infeasible as the number of test variables per cluster increases.

　　Algorithm A_2 in essence is non-myopic in its test-selection strategy. As we argued above, this non-myopic algorithm might become computationally too demanding, especially if a meaningful clustering of test variables results 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. This semi-myopic algorithm equals

Algorithm A_3: semi-myopic test selection with subgoals

input:$\mathscr{S}$: list of subgoals S_i;
$\mathscr{T}$: list of clusters C_j of test variables, organised in sublists $\mathscr{T}(S_i)$ per subgoal

Stop-subgoal(S_i)=*false*
Stop-overall=*false*
while $\mathscr{S} \neq \varnothing$ and *Stop-overall*$\neq$*true* **do**
　　select S_i from $\mathscr{S}$
　　remove S_i from $\mathscr{S}$

```
    while 𝒯(S_i) ≠ ∅, Stop-subgoal(S_i) ≠ true  and  Stop-overall ≠ true  do
         compute most informative T_j ∈ 𝒯(S_i)
         prompt for evidence for T_j and for all T_k ∈ C_m with C_m such that T_j ∈ C_m,
         and propagate
         remove C_m from 𝒯
         compute  Stop-subgoal(S_i)
         compute  Stop-overall
    od
od
```

The algorithm very much resembles the myopic algorithm A_1 presented above. The main difference is that algorithm A_3 prompts not just for the result of the selected test variable T_j, but for the results of all test variables T_k that belong to the same cluster as the selected variable. Entering evidence for physical tests rather than for just one test variable seems to fit in more closely with the daily practice of the physicians, since in daily practice they are inclined to think in terms of physical tests even though they may be interested mainly in the value of a single variable. After performing the test, therefore, it seems logical to enter not only the result that is currently of interest, but all other results obtained from the same physical test as well.

4 Test-selection Experiment

To perform a first informal evaluation and compare the performance of the three algorithms for test selection described above, we conducted an experimental study in the context of the oesophageal cancer network. From the results of our test-selection experiments described in [5], we observed that the test-selection sequences do not differ that much among the patient records for which we ran the experiment. This appears to imply that specific test variables are highly informative for many patients, while other variables provide less valuable information. Since some test variables are highly informative, the physical test they pertain to is likely to be quite informative as well. We thus do not expect large differences in the test-selection sequences resulting from the three different algorithms. Note that the three algorithms may in general behave rather differently. Using the oesophageal cancer network we compared the sequences of test variables selected by the different algorithms using a specific second stopping criterion. We briefly evaluated the overall test-selection strategy implied by the myopic and semi-myopic algorithms with our two domain experts. The semi-myopic test-selection algorithm proved to result in the most preferred sequence of test variables.

We implemented the three test-selection algorithms described in Section 3. In our experimental study we used an extended version of the original oesophageal cancer network. Since our algorithms were designed to perform the selection of test variables with respect to the various subgoals we elicited during the interviews, we decided to extend the original oesophageal cancer network with these (sub)goal variables. As described earlier, our new test-selection algorithms need extra information for their input in addition to the probabilistic network under consideration. The extra information in essence consists of three lists: a list of the subgoals for the test-selection process, a list that describes for each subgoal which test variables provide information about that particular subgoal, and a list for each physical test that describes which test variables serve

to capture the results of that test. We provided the system with this extra information obtained from the domain experts.

We ran experiments for eight patient cases that we had designed for our interviews with the experts. These cases were carefully designed to ensure that the patients differed as much as possible with respect to their characteristics and the tests that should be performed. During the interviews we asked the experts to describe the test-selection procedure they would employ for these patients. We were interested in whether or not the various algorithms would result in sequences of test variables that fit the test-selection strategy employed by our experts;

We recall that algorithm A_1 is fully myopic in nature, selecting the diagnostic tests to be performed based upon the informative values of the separate variables: the expected value of the Gini index of the probability distribution over the subgoal after performing a diagnostic test is computed for single test variables. After selecting the most informative test variable, the user is prompted only for the result of this variable. Algorithm A_3 also performs the selection of tests at the level of the separate variables. After selecting the most informative test variable, however, the user is prompted not only for the result of this variable, but also for the results of all test variables for which a result is obtained from the same physical test. In contrast with the algorithms A_1 and A_3, algorithm A_2 selects diagnostic tests at the level of physical tests: the expected value of the Gini index of the posterior probability distribution over the subgoal is computed given all possible combinations of results for the various different test variables that are clustered with respect to a single physical test. We would like to note that not all test variables for which a physical test will yield a result, pertain to the same subgoal. For computing the expected value of the Gini index after performing the test, we decided to take only those test variables into account that pertain to the subgoal under study. Upon prompting the user for the test's results, however, the algorithm requests the results for all variables concerned, also for those that belong to a different subgoal, since the test has now been performed and its results are available.

For all eight patient cases, we ran the three different algorithms. We will now discuss the results obtained for a particular patient and in doing so focus on the subgoal *M1* which was selected by all algorithms. Algorithm A_1 selected the variable *CT-liver* as the most informative test variable with regard to the subgoal under study. The user was prompted for its results. The value *CT-liver=no* was entered and was subsequently propagated to establish the posterior distribution for the subgoal *M1* given the new information. The stopping criterion for subgoal *M1* nor the overall criterion were met. The algorithm therefore proceeded. The now selected test variable was *CT-truncus*. The result *CT-truncus=no* was entered into the network and propagated. Again both stopping criteria were not met yet. The next test variable selected by the algorithm was *Endosono-truncus*. The result *Endosonotruncus=no* was entered into the network and propagated. Again the stopping criterion was not met. Also the overall criterion was not satisfied. The next test variable for which the user was prompted for a result, was *X-lungs*. The result *X-lungs=yes* was entered into the network and propagated. Next, the result of the CT-scan of the lungs was prompted for. The result *CT-lungs=yes* was entered into the network. After propagation, the stopping criterion was now satisfied for the subgoal *M1*. The stopping criterion for the overall goal *Therapy* was met at this

Table 1. The sequence of tests selected for the three different algorithms.

goal: M1	A_1	A_2	A_3
	CT-liver	CT-abdomen	*CT-liver*; CT-abdomen
	CT-truncus	Endosonography	*Endosono-truncus*;
			Endosonography
	Endosono-truncus	X-thorax	*X-lungs*; X-thorax
	X-lungs	CT-thorax	*CT-lungs*; CT-thorax
	CT-lungs		

stage as well. The test-selection process was thus halted. The selection of test variables is summarised in the first column of Table 1.

Algorithm A_2 selected the following sequence of physical tests as the most informative before the stopping criterion was met: the CT scan of the patient's upper abdomen, an endosonography, an X-ray of the patient's thorax and a CT-scan of the thorax. After propagating the information of the CT-scan of the thorax, the stopping criterion was now satisfied for the subgoal *M1*. The stopping criterion was also met for the overall goal *Therapy*. The test-selection process was thus halted. The selection of test variables is summarised in the second column of Table 1. We would like to note that the sequence of physical tests selected by algorithm A_2 equals the sequence generated by algorithm A_1. However, algorithm A_2 prompts the user for twice as many test results.

Algorithm A_3 equally showed a similar behaviour. The algorithm again selected the test variable *CT-liver* to be the most informative. Since the physical test with which a value for *CT-liver* is found also yields a value for *CT-truncus*, the user was prompted for the values of both variables. The test variable *Endosono-truncus* was found to be the next most informative test variable. All results for the variables beloging to the test *Endosonography* were entered. The test variable *X-lungs* was found to be the next most informative test variable. Next, the result of the test variable *CT-lungs* and the variables belonging to the same physical test *CT-thorax* were prompted for. The stopping criterion for the overall goal *Therapy* was also met. The test-selection process was then halted. The selection of test variables is summarised in the right most column of Table 1. We note that the sequence of physical tests selected by algorithm A_3 equals the sequence generated by algorithms A_1 and A_2. Algorithms A_2 and A_3 prompted the user for equally many test results.

We would like to note that all three algorithms resulted in rather similar sequences of tests not just for the patient reviewed above, but also for the other seven patient cases with which we conducted our experiment. Our findings may be explained as follows. When a single test variable is expected to result in the largest decrease of diagnostic uncertainty, then it is likely that the test to which it pertains will be the most informative test. To this end we recall that the most informative test is computed by using all combinations of possible test results for the various different variables belonging to a physical test. Now if, for instance, the variable that models the result of the CT-scan of the liver is very informative, then the CT-scan of the abdomen is likely to be informative as well. We would like to note that this property does not necessarily hold in general. Examples can be readily constructed for which the most informative test variable does not belong to the most informative physical test. We note that the algorithms A_2 and A_3 especially

showed the same behaviour in all patient cases. Since algorithm A_2 is computationally much more demanding than algorithm A_3, we would like to recommend the use of the semi-myopic test-selection algorithm.

To conclude, we presented the sequences of tests constructed by the algorithms A_1 and A_3 for patient 2 as discussed above, to our domain experts. We asked the experts which sequence of tests they felt most comfortable with. We also asked them whether they would prefer to order more tests after the results of the CT-scan of the patient's lungs had become available, or would order fewer tests. The experts indicated that they felt more comfortable with the sequence of test variables generated by the semi-myopic algorithm A_3. They indicated that the sequence generated by algorithm A_1 appeared unnatural, referring to for instance the CT-scan of the liver and the CT-scan of the truncus coeliacus not being performed at the same time.

5 Conclusions

Most test-selection algorithms currently in use with probabilistic networks select variables myopically, that is, 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 presented 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.

From our experiments with the oesophageal cancer network, we found that the three algorithms behave comparably. The semi-myopic and non-myopic algorithms in fact resulted in the same sequences of selected tests. Of these two algorithms, the semi-myopic algorithm is to be preferred, since it is computationally much less demanding. We presented the sequences of test variables, and the moments at which the test-selection process halted, as generated by the myopic and semi-myopic algorithms, to our domain experts. The experts indicated that they felt more comfortable with the results of the semi-myopic. With our test-selection algorithms, the experts are supported in working far more patient specific, which is likely to induce a decrease in the amount of money spent and an increase in the quality of their management.

References

1. S. Andreassen. Planning of therapy and tests in causal probabilistic networks. *Artifical Intelligence in Medicine*, 4:227 – 241, 1992.
2. M. Ben-Bassat. Myopic policies in sequential classification. *IEEE Transactions on Computers*, 27(2):170 – 174, 1978.
3. P. Doubilet. A mathematical approach to interpretation and selection of diagnostic tests. *Medical Decision Making*, 3(2):177 – 195, 1983.
4. P. Glasziou and J. Hilden. Test selection measures. *Medical Decision Making*, 9:133 – 141, 1989.
5. D. Sent and L.C. van der Gaag. On the behaviour of information measures for test selection. *submitted for publication*, 2007.