Preprocessing the MAP problem

Janneke H. Bolt and Linda C. van der Gaag
Department of Information and Computing Sciences, Utrecht University
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands
{janneke,linda}@cs.uu.nl

Abstract

The MAP problem for Bayesian networks is the problem of finding for a set of variables an instantiation of highest posterior probability given the available evidence. The problem is known to be computationally infeasible in general. In this paper, we present a method for preprocessing the MAP problem with the aim of reducing the runtime requirements for its solution. Our method exploits the concepts of Markov and MAP blanket for deriving partial information about a solution to the problem. We investigate the practicability of our preprocessing method in combination with an exact algorithm for solving the MAP problem for some real Bayesian networks.

1 Introduction

Upon reasoning with a Bayesian network, often a best explanation is sought for a given set of observations. Given the available evidence, such an explanation is an instantiation of highest probability for some subset of the network's variables. The problem of finding such an instantiation has an unfavourable computational complexity. If the subset of variables for which a most likely instantiation is to be found includes just a single variable, then the problem, which is known as the Pr problem, is NP-hard in general. A similar observation holds for the MPE problem in which an instantiation of highest probability is sought for all unobserved variables. These two problems can be solved in polynomial time, however, for networks of bounded treewidth. If the subset of interest is a non-singleton proper subset of the set of unobserved variables, on the other hand, the problem, which is then known as the MAP problem, remains NP-hard even for networks for which the other two problems can be feasibly solved (Park and Darwiche, 2002).

By performing inference in a Bayesian network under study and establishing the most likely value for each variable of interest separately, an estimate for a solution to the MAP problem may be obtained. There is no guarantee in general, however, that the values in the resulting joint instantiation indeed correspond to the values of the variables in a solution to the MAP problem. In this paper, we now show that, for some of the variables of interest, the computation of marginal posterior probabilities may in fact provide exact information about their value in a solution to the MAP problem. We show more specifically that, by building upon the concept of Markov blanket, some of the variables may be fixed to a particular value; for some of the other variables of interest, moreover, values may be excluded from further consideration. We further introduce the concept of MAP blanket that serves to provide similar information.

Deriving partial information about a solution to the MAP problem by building upon the concepts of Markov and MAP blanket, can be exploited as a preprocessing step before the problem is actually solved with any available algorithm. The derived information in essence serves to reduce the search space for the problem and thereby reduces the algorithm's runtime requirements. We performed an initial study of the practicability of our preprocessing method by solving MAP problems for real networks using an exact branch-and-bound algorithm, and found that preprocessing can be profitable.

The paper is organised as follows. In Section 2, we provide some preliminaries on the MAP problem. In Section 3, we present two propositions that constitute the basis of our preprocessing method. In Section 4, we provide some preliminary results about the practicability of our preprocessing method. The paper is ended in Section 5 with our concluding observations.

2 The MAP problem

Before reviewing the MAP problem, we introduce our notational conventions. A Bayesian network is a model of a joint probability distribution Pr over a set of stochastic variables, consisting of a directed acyclic graph and a set of conditional probability distributions. We denote variables by upper-case letters (A) and their values by (indexed) lower-case letters (a_i); sets of variables are indicated by bold-face upper-case letters ($\mathbf{A}$) and their instantiations by bold-face lower-case letters ($\mathbf{a}$). Each variable is represented by a node in the digraph; (conditional) independence between the variables is encoded by the digraph's set of arcs according to the d-separation criterion (Pearl, 1988). The Markov blanket $\mathbf{B}$ of a variable A consists of its neighbours in the digraph plus the parents of its children. Given its Markov blanket, the variable is independent of all other variables in the network. The strengths of the probabilistic relationships between the variables are captured by conditional probability tables that encode for each variable A the conditional distributions $\Pr(A \mid \mathbf{p}(\mathbf{A}))$ given its parents $\mathbf{p}(\mathbf{A})$.

Upon reasoning with a Bayesian network, often a best explanation is sought for a given set of observations. Given evidence $\mathbf{o}$ for a subset of variables $\mathbf{O}$, such an explanation is an instantiation of highest probability for some subset $\mathbf{M}$ of the network's variables. The set $\mathbf{M}$ is called the *MAP set* for the problem; its elements are called the *MAP variables*. An instantiation $\mathbf{m}$ of highest probability to the set $\mathbf{M}$ is termed a *MAP solution*; the value that is assigned to a MAP variable in a solution $\mathbf{m}$ is called its *MAP value*. Dependent upon the size of the MAP set, we distinguish between three different types of problem. If the MAP set includes just a single variable, the problem of finding the best explanation for a set of observations reduces to establishing the most likely value for this variable from its marginal posterior probability distribution. This problem is called the Pr *problem* as it essentially amounts to performing standard inference (Park and Darwiche, 2001). In the second type of problem, the MAP set includes *all* non-observed variables. This problem is known as the most probable explanation or *MPE problem*. In this paper, we are interested in the third type of problem, called the *MAP problem*, in which the MAP set is a non-singleton proper subset of the set of non-observed variables of the network under study. This problem amounts to finding an instantiation of highest probability for a designated set of variables of interest.

We would like to note that the MAP problem is more complex in essence than the other two problems. The Pr problem and the MPE problem both are NP-hard in general and are solvable in polynomial time for Bayesian networks of bounded treewidth. The MAP problem is NP^{PP}-hard in general and remains NP-hard for these restricted networks (Park, 2002).

3 Fixing MAP values

By performing inference in a Bayesian network and solving the Pr problem for each MAP variable separately, an estimate for a MAP solution may be obtained. There is no guarantee in general, however, that the value with highest marginal probability for a variable corresponds with its value in a MAP solution. We now show that, for some variables, the computation of marginal probabilities may in fact provide exact information about their MAP values.

The first property that we will exploit in the sequel, builds upon the concept of Markov blanket. We consider a MAP variable H and its associated Markov blanket. If a specific value h_i of H has highest probability in the marginal distribution over H for all possible instantiations of the blanket, then h_i will be the value of H in a MAP solution for any MAP problem including H. Alternatively, if some value h_j never has highest marginal probability, then this value cannot be included in any solution.

Proposition 1. *Let H be a MAP variable in a Bayesian network and let $\mathbf{B}$ be its Markov blanket. Let h_i be a specific value of H.*

1. *If $\Pr(h_i \mid \mathbf{b}) \geq \Pr(h_k \mid \mathbf{b})$ for all values h_k of H and all instantiations $\mathbf{b}$ of $\mathbf{B}$, then h_i is the value of H in a MAP solution for any MAP problem that includes H.*

2. *If there exist values h_k of H with $\Pr(h_i \mid \mathbf{b}) < \Pr(h_k \mid \mathbf{b})$ for all instantiations $\mathbf{b}$ of $\mathbf{B}$, then h_i is not the MAP value of H in any solution to a MAP problem that includes H.*

Proof. We prove the first property stated in the proposition; the proof of the second property builds upon similar arguments.

We consider an arbitrary MAP problem with the MAP set $\{H\} \cup \mathbf{M}$ and the evidence $\mathbf{o}$ for the observed variables $\mathbf{O}$. Finding a solution to the problem amounts to finding an instantiation to the MAP set that maximises the posterior probability $\Pr(H, \mathbf{M} \mid \mathbf{o})$. We have that

$$\Pr(h_i, \mathbf{M} \mid \mathbf{o}) = \\ = \sum_{\mathbf{b}} \Pr(h_i \mid \mathbf{b}, \mathbf{o}) \cdot \Pr(\mathbf{M} \mid \mathbf{b}, \mathbf{o}) \cdot \Pr(\mathbf{b} \mid \mathbf{o})$$

For the posterior probability $\Pr(h_k, \mathbf{M} \mid \mathbf{o})$ an analogous expression is found. Now suppose that for the value h_i of H we have that $\Pr(h_i \mid \mathbf{b}) \geq \Pr(h_k \mid \mathbf{b})$ for all values h_k of H and all instantiations $\mathbf{b}$ of $\mathbf{B}$. Since $\mathbf{B}$ is the Markov blanket of H, we have that $\Pr(h_i \mid \mathbf{b}) \geq \Pr(h_k \mid \mathbf{b})$ implies $\Pr(h_i \mid \mathbf{b}, \mathbf{o}) \geq \Pr(h_k \mid \mathbf{b}, \mathbf{o})$ for all values h_k of H and all instantiations $\mathbf{b}$ of $\mathbf{B}$. We conclude that $\Pr(h_i, \mathbf{M} \mid \mathbf{o}) \geq \Pr(h_k, \mathbf{M} \mid \mathbf{o})$ for all h_k. The value h_i of H thus is included in a solution to the MAP problem under study. Since the above considerations are algebraically independent of the MAP variables $\mathbf{M}$ and of the evidence $\mathbf{o}$, this property holds for any MAP problem that includes the variable H. □

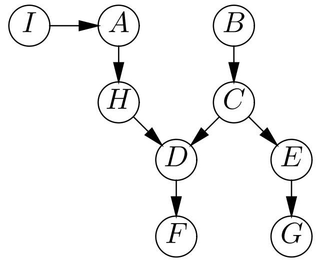

Figure 1: An example directed acyclic graph.

The above proposition provides for preprocessing a MAP problem. Prior to actually solving the problem, some of the MAP variables may be fixed to a particular value using the first property. With the second property, moreover, various values of the MAP variables may be excluded from further consideration. By building upon the proposition, therefore, the search space for the MAP problem is effectively reduced. We illustrate this with an example.

Example 1. We consider a Bayesian network with the graphical structure from Figure 1. Suppose that H is a ternary MAP variable with the values h_1, h_2 and h_3. The Markov blanket $\mathbf{B}$ of H consists of the three variables A, C and D. Now, if for any instantiation $\mathbf{b}$ of these variables we have that $\Pr(h_1 \mid \mathbf{b}) \geq \Pr(h_2 \mid \mathbf{b})$ and $\Pr(h_1 \mid \mathbf{b}) \geq \Pr(h_3 \mid \mathbf{b})$, then h_1 occurs in a MAP solution for any MAP problem that includes H. By fixing the variable H to the value h_1, the search space of any such problem is reduced by a factor 3. We consider, as an example, the MAP set $\{H, E, G, I\}$ of ternary variables. Without preprocessing, the search space includes $3^4 = 81$ possible instantiations. By fixing the variable H to h_1, the search space of the problem reduces to $3^3 = 27$ instantiations. □

Establishing whether or not the properties from Proposition 1 can be used for a specific MAP variable, requires a number of computations that is exponential in the size of the variable's Markov blanket. The computations required, however, are highly local. A single restricted inward propagation for each instantiation of the Markov blanket to the variable of interest suffices. Since the proposition moreover holds for any MAP problem that includes the variable, the computational burden involved is amortised over all future MAP computations.

The second proposition that we will exploit for preprocessing a MAP problem, builds upon the new concept of *MAP blanket*. We consider a problem with the MAP variables $\{H\} \cup \mathbf{M}$. A MAP blanket $\mathbf{K}$ of H now is a minimal subset $\mathbf{K} \subseteq \mathbf{M}$ such that $\mathbf{K}$ d-separates H from $\mathbf{M} \setminus \mathbf{K}$ given the available evidence. Now, if a specific value h_i of H has highest probability in the marginal distribution over H given the evidence for all possible instantiations of $\mathbf{K}$, then h_i will be the MAP value of H in a solution to the MAP problem under study. Alternatively, if some value h_j never has highest marginal probability, then this value cannot be a MAP value in any of the problem's solutions.

Proposition 2. *Let $\{H\} \cup \mathbf{M}$ be the MAP set of a given MAP problem for a Bayesian network and let $\mathbf{o}$ be the evidence that is available for the observed variables $\mathbf{O}$. Let $\mathbf{K}$ be the MAP blanket for the variable H given $\mathbf{o}$, and let h_i be a specific value of H.*

1. *If $\Pr(h_i \mid \mathbf{k}, \mathbf{o}) \geq \Pr(h_k \mid \mathbf{k}, \mathbf{o})$ for all values h_k of H and all instantiations $\mathbf{k}$ to $\mathbf{K}$, then h_i is the MAP value of H in a solution to the given MAP problem.*

2. *If there exist values h_k of H with $\Pr(h_i \mid \mathbf{k}, \mathbf{o}) < \Pr(h_k \mid \mathbf{k}, \mathbf{o})$ for all instantiations $\mathbf{k}$ to $\mathbf{K}$, then h_i is not the MAP value of H in any solution to the given MAP problem.*

The proof of the proposition is relatively straightforward, building upon similar arguments as the proof of Proposition 1.

The above proposition again provides for preprocessing a MAP problem. Prior to actually solving the problem, the values of some variables may be fixed and other values may be excluded for further consideration. The proposition therefore again serves to effectively reduce the search space for the problem under study. While the information derived from Proposition 1 holds for any MAP problem including H, however, Proposition 2 provides information for any problem in which H has a subset of $\mathbf{K}$ for its MAP blanket and with matching evidence for the observed variables that are not d-separated from H by $\mathbf{K}$. The information derived from Proposition 2, therefore, is more restricted in scope than that from Proposition 1.

We illustrate the application of Proposition 2 for our example network.

Example 2. We consider again the MAP set $\{H, E, G, I\}$ for the Bayesian network from Example 1. In the absence of any evidence, the MAP blanket of the variable H includes just the variable I. Now, if for each value i of I we have that $\Pr(h_1 \mid i) \geq \Pr(h_2 \mid i)$ and $\Pr(h_1 \mid i) \geq \Pr(h_3 \mid i)$, then the value h_1 occurs in a solution to the given MAP problem. The search space for actually solving the problem thus again is reduced from 81 to 27. $\square$

Establishing whether or not the properties from Proposition 2 can be used for a specific MAP variable, requires a number of computations that is exponential in the size of the variable's MAP blanket. The size of this blanket is strongly dependent of the network's connectivity and of the location of the various MAP variables and observed variables in the network. The MAP blanket can in fact be larger in size than the Markov blanket of the variable. The computations required, moreover, are less local than those required for Proposition 1 and can involve full inward propagations to the variable of interest. Since the proposition in addition applies to just a restricted class of MAP problems, the computational burden involved in its verification can be amortised over other MAP computations to a lesser extent than that involved in the verification of Proposition 1.

The general idea underlying the two propositions stated above is the same. The idea is to verify whether or not a particular value of H can be fixed or excluded as a MAP value by investigating H's marginal probability distributions given all possible instantiations of a collection of variables surrounding H. Proposition 1 uses for this purpose the Markov blanket of H. By building upon the Markov blanket, which in essence is independent of the MAP set and of the entered evidence, generally applicable statements about the values of H are found. A disadvantage of building upon the Markov blanket, however, is that maximally different distributions for H are examined, which decreases the chances of fixing or excluding values.

By taking a blanket-like collection of variables at a larger distance from the MAP variable H, the marginal distributions examined for H are likely to be less divergent, which serves to increase the chances of fixing or excluding values of H as MAP values. A major disadvantage of such a blanket, however, is that its size tends to grow with the distance from H, which will result in an infeasibly large number of instantiations to be studied. For any blanket-like collection, we observe that the MAP variables of the problem have to either be in the blanket or be d-separated from H by the blanket. Proposition 2 builds upon this observation explicitly and considers only the instantiations of the MAP blanket of the variable. The proposition thereby reduces the computations involved in its application yet retains and even further exploits the advantage of examining less divergent marginal distributions over H. Note that the values that can be fixed or excluded based on the first proposition, will also be fixed or excluded based on the second proposition. It may nevertheless still be worthwhile to exploit the first proposition because, as stated before, with this proposition values can be fixed or excluded in general and more restricted and possibly less computations are required.

So far we have argued that application of the two propositions serves to reduce the search space for a MAP problem by fixing variables to particular values and by excluding other values from further consideration. We would like to mention that by fixing variables the graphical structure of the Bayesian network under study may fall apart into unconnected components, for which the MAP problem can be solved separately. We illustrate the basic idea with our running example.

Example 3. We consider again the MAP set $\{H, E, G, I\}$ for the Bayesian network from Figure 1. Now suppose that the variable H can be fixed to a particular value. Then, by performing evidence absorption of this value, the graphical structure of the network falls apart into the two components $\{A, I, H\}$ and $\{B, C, D, E, F, G\}$, respectively. The MAP problem then decomposes into the problem with the MAP set $\{I\}$ for the first component and the problem with the MAP set $\{E, G\}$ for the second component; both these problems now include the value of H as further evidence. The search space thus is further reduced from 27 to $3 + 9 = 12$ instantiations to be studied. □

4 Experiments

In the previous section, we have introduced a method for preprocessing the MAP problem for Bayesian networks. In this section, we perform a preliminary study of the practicability of our method by solving MAP problems for real networks using an exact algorithm. In Section 4.1 we describe the set-up of the experiments; we review the results in Section 4.2.

4.1 The Experimental Set-up

In our experiments, we study the effects of our preprocessing method on three real Bayesian networks. We first report the percentages of values that are fixed or excluded by exploiting Proposition 1. We then compare the numbers of fixed variables as well as the numbers of network propagations with and without preprocessing, upon solving various MAP problems with a state-of-the-art exact algorithm.

In our experiments, we use three real Bayesian networks with a relatively high connectivity; Table 1 reports the numbers of variables and values for these networks. The *Wilson's disease* network (WD) is a small network in medicine, developed for the diagnosis of Wilson's liver disease (Korver and Lucas, 1993). The *classical swine fever* network (CSF) is a network in veterinary science, currently under development, for the early detection of outbreaks of classical swine fever in pig herds

(Geenen and Van der Gaag, 2005). The extended *oesophageal cancer* network (OESO+) is a moderately-sized network in medicine, which has been developed for the prediction of response to treatment of oesophageal cancer (Aleman et al, 2000). For each network, we compute MAP solutions for randomly generated MAP sets with 25% and 50% of the network's variables, respectively; for each size, five sets are generated. We did not set any evidence.

For solving the various MAP problems in our experiments, we use a basic implementation of the exact branch-and-bound algorithm available from Park and Darwiche (2003). This algorithm solves the MAP problem exactly for most networks for which the Pr and MPE problems are feasible. The algorithm constructs a depth-first search tree by choosing values for subsequent MAP variables, cutting off branches using an upper bound. Since our preprocessing method reduces the search space by fixing variables and excluding values, it essentially serves to decrease the depth of the tree and to diminish its branching factor.

4.2 Experimental results

In the first experiment, we established for each network the number of values that could be fixed or excluded by applying Proposition 1, that is, by studying the marginal distributions per variable given its Markov blanket. For computational reasons, we decided not to investigate variables for which the associated blanket had more than 45 000 different instantiations; this number is arbitrarily chosen. In the WD, CSF and OESO+ networks, there were 0, 8 and 15 of such variables respectively.

The results of the first experiment are presented in Table 1. The table reports, for each network, the total number of variables, the total number of values, the number of variables for which a value could be fixed, and the number of values that could be fixed or excluded; note that if, for example, for a ternary variable a value can be fixed, then also two values can be excluded. We observe that 17.1% to 19.0% of the variables could be fixed to a particular value. The number of values that could be fixed

Table 1: The numbers of fixed and excluded values.

network	*#vars.*	*#vals.*	*#vars.f.*	*#vals.f.+e.*
WD	21	56	4(19.0%)	13(23.2%)
CSF	41	98	7(17.1%)	15(15.3%)
OESO+	67	175	12(17.9%)	27(15.4%)

or excluded ranges between 15.3% and 23.2%. We would like to stress that whenever a variable can be fixed to a particular value, this result is valid for any MAP problem that includes this variable. The computations involved, therefore, have to be performed only once.

In the second experiment, we compared for each network the numbers of variables that could be fixed by the two different preprocessing steps; for Proposition 2, we restricted the number of network propagations per MAP variable to four because of the limited applicability of the resulting information. We further established the numbers of network propagations of the exact branch-and-bound algorithm that were forestalled by the preprocessing.

The results of the second experiment are presented in Table 2. The table reports in the two leftmost columns, for each network, the sizes of the MAP sets used and the average number of network propagations performed without any preprocessing. In the subsequent two columns, it reports the average number of variables that could be fixed to a particular value by using Proposition 1 and the average number of network propagations performed by the branch-and-bound algorithm after this preprocessing step. In the fifth and sixth columns, the table reports the numbers obtained with Proposition 2; the sixth column in addition shows, between parenthesis, the average number of propagations that are required for the application of Proposition 2. In the final two columns of the table, results for the two preprocessing steps combined are reported: the final but one column again mentions the number of variables that could be fixed to a particular value by Proposition 1; it moreover mentions the average number of variables that could be fixed to a particular value by Proposition 2 after Proposition 1 had been used. The rightmost column reports the average number of network propagations required by the branch-and-bound algorithm. We would like to note that the additional computations required for Proposition 2 are restricted network propagations; although the worst-case complexity of these propagations is the same as that of the propagations performed by the branch-and-bound algorithm, their runtime requirements may be considerably less.

From Table 2, we observe that the number of network propagations performed by the branch-and-bound algorithm grows with the number of MAP variables, as expected. For the two smaller networks, we observe in fact that the number of propagations without preprocessing equals the number of MAP variables plus one. For the larger OESO+ network, the number of network propagations performed by the algorithm is much larger than the number of MAP variables. This finding is not unexpected since the MAP problem has a high computational complexity. For the smaller networks, we further observe that each variable that is fixed by one of the preprocessing steps translates directly into a reduction of the number of propagations by one. For the OESO+ network, we find a larger reduction in the number of network propagations per fixed variable. Our experimental results thus indicate that the number of propagations required by the branch-and-bound algorithm indeed is decreased by fixing variables to their MAP value.

With respect to using Proposition 1, we observe that in all networks under study a reasonably number of variables could be fixed to a MAP value. We did not take the number of local propagations required for this proposition into consideration because the computational burden involved is amortised over future MAP computations. With respect to using Proposition 2, we observe that for all networks and all MAP sets the additional computations involved outweigh the number of network computations that are forestalled for the algorithm. This observation applies to using just Proposition 2 as well as to using the proposition after Proposition 1 has been applied. We also observe that fewer variables are fixed in the step based on Proposition 2 than in the step based on Proposition 1. This can be attributed to the limited number of propagations used in the step based on Proposition 2. We conclude that, for the networks under study, it has been quite worthwhile to use Proposition 1 as a preprocessing step before actually solving the various MAP problems. Because of the higher computational burden involved and its relative lack of additional value, the use of Proposition 2 has not been worthwhile for our networks and associated problems.

To conclude, in Section 3 we observed that fixing variables to their MAP value could serve to partition a MAP problem into smaller problems. Such a partition did not occur in our experiments. We would like to note, however, that we studied MAP problems without evidence only. We expect that in the presence of evidence MAP problems will more readily be partitioned into smaller problems.

5 Conclusions and discussion

The MAP problem for Bayesian networks is the problem of finding for a set of variables an instantiation of highest posterior probability given the available evidence. The problem has an unfavourable computational complexity, being $\mathrm{NP}^{\mathrm{PP}}$-hard in general. In this paper, we showed that computation of the marginal posterior probabilities of a variable H given its Markov blanket may provide exact information about its value in a MAP solution. This information is valid for any MAP problem that includes the variable H. We further showed that computation of the marginal probabilities of H given its MAP blanket may also provide exact information about its value. This information is valid, however, for a more restricted class of MAP problems. We argued that these results can be exploited for preprocessing MAP problems before they are actually solved using any state-of-the-art algorithm for this purpose.

We performed a preliminary experimental study of the practicability of the preprocessing steps by solving MAP problems for three different Bayesian networks using an exact branch-

Table 2: The number of network propagations without and with preprocessing using Propositions 1 and 2.

			Prop. 1		Prop. 2		Prop. 1+2	
network	*#MAP*	*#props.*	*#vars. f.*	*#props.*	*#vars. f.*	*#props.* (add.)	*#vars. f.*	*#props.* (add.)
WD	5	6.0	1.2	4.8	0.0	6.0 (0.6)	$1.2 + 0.0$	4.8 (0.6)
	10	11.0	2.4	8.6	1.0	10.0 (6.4)	$2.4 + 0.2$	8.4 (4.0)
CSF	10	11.0	2.0	9.0	1.0	10.0 (1.6)	$2.0 + 0.4$	8.6 (0.4)
	20	21.0	3.4	17.6	1.6	19.4 (8.2)	$3.4 + 0.6$	17.0 (5.0)
OESO+	17	24.8	3.4	18.4	0.8	22.0 (4.2)	$3.4 + 0.0$	18.4 (2.2)
	34	49.8	5.8	40.8	1.8	47.4 (7.8)	$5.8 + 0.4$	40.0 (4.6)

and-bound algorithm. As expected, the number of network propagations required by the algorithm is effectively decreased by fixing MAP variables to their appropriate values. We found that by building upon the concept of Markov blanket for 17.1% to 19.0% of the variables a MAP value could be fixed. Since the results of this preprocessing step are applicable to all future MAP computations and the computational burden involved thus is amortised, we considered it worthwhile to perform this step for the investigated networks. We would like to add that, because the computations involved are highly local, the step may also be feasibly applied to networks that are too large for the exact MAP algorithm used in the experiments. With respect to building upon the concept of MAP blanket, we found that for the investigated networks and associated MAP problems, the computations involved outweighed the reduction in network propagations that was achieved. Since the networks in our study were comparable with respect to size and connectivity, further experiments are necessary before any definitive conclusion can be drawn with respect to this preprocessing step.

In our further research, we will expand the experiments to networks with different numbers of variables, different cardinality and diverging connectivity. More specifically, we will investigate for which types of network preprocessing is most profitable. In our future experiments we will also take the effect of evidence into account. We will further investigate if the class of MAP problems that can be feasibly solved can be extended by our preprocessing method. We hope to report further results in the near future.

Acknowledgements

This research was partly supported by the Netherlands Organisation for Scientific Research (NWO).

References

B.M.P. Aleman, L.C. van der Gaag and B.G. Taal. 2000. *A Decision Model for Oesophageal Cancer – Definition Document* (internal publication in Dutch).

P.L. Geenen and L.C. van der Gaag. 2005. Developing a Bayesian network for clinical diagnosis in veterinary medicine: from the individual to the herd. In *3rd Bayesian Modelling Applications Workshop*, held in conjunction with the 21st Conference on Uncertainty in Artificial Intelligence.

M. Korver and P.J.F. Lucas. 1993. Converting a rule-based expert system into a belief network. *Medical Informatics*, 18:219–241.

J. Park and A. Darwiche. 2001. Approximating MAP using local search. In *17th Conference on Uncertainty in Artificial Intelligence*, pages 403–410.

J. Park. 2002. MAP complexity results and approximation methods. In *18th Conference on Uncertainty in Artificial Intelligence*, pages 388–396.

J. Park and A. Darwiche. 2003. Solving MAP exactly using systematic search. In *19th Conference on Uncertainty in Artificial Intelligence*, pages 459–468.

J. Pearl. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann Publishers, Palo Alto.