Verifying Monotonicity of Bayesian Networks with Domain Experts

Linda C. van der Gaag

Hermina J.M. Tabachneck-Schijf

Petra L. Geenen

Technical Report UU-CS-2008-014
June 2008

Department of Information and Computing Sciences
Utrecht University, Utrecht, The Netherlands
www.cs.uu.nl

ISSN: 0924-3275

Department of Information and Computing Sciences
Utrecht University
P.O. Box 80.089
3508 TB Utrecht
The Netherlands

Verifying Monotonicity of Bayesian Networks with Domain Experts

Linda C. van der Gaag, Hermina J.M. Tabachneck-Schijf, Petra L. Geenen
Department of Information and Computing Sciences, Utrecht University,
P.O. Box 80.089, 3508 TB Utrecht, the Netherlands.
e-mail: {linda,hermi,petrag}@cs.uu.nl

Abstract

In many realistic problem domains, the main variable of interest behaves monotonically in the observable variables, in the sense that higher values for the variable of interest become more likely with higher-ordered observations. This type of knowledge appears to naturally emerge from experts during knowledge elicitation, without explicit prompting from the knowledge engineer. The experts' concept of monotonicity, however, may not correspond to the mathematical concept of monotonicity in Bayesian networks. We present a method that provides both for verifying whether or not a network exhibits the properties of monotonicity suggested by the experts and for studying the violated properties with the experts. We illustrate the application of our method for a real Bayesian network in veterinary science.

1 Introduction

For many diagnostic problems, the relation between the output variable of interest and the observable input variables is isotone in the sense that higher values for the input variables give rise to a higher-ordered output for the variable of interest. In a medical diagnostic application, for example, observing symptoms and signs that are more severe will result in a more severe disease becoming a more likely explanation. Such sensations of isotonicity are commonly shared in a domain of application. Upon developing models to support experts in their diagnostic reasoning tasks, it is important that the resulting model reflects any commonly acknowledged sensations of isotonicity. If a model violates one of these sensations, it will exhibit a reasoning behaviour that is counterintuitive to the experts, which is likely to result in a dip in acceptance of the model in daily practice. For many types of model, therefore, techniques have been developed for studying properties of isotonicity, for example for neural networks [1], for classification trees [11], and for regression models [12]. Also for Bayesian networks a mathematical concept of isotonicity has been formulated [4]. A network is said to be isotone in its observable input variables if the probability distribution computed for the output variable given specific observations is stochastically dominated by any such distribution given higher-ordered observations.

Our experiences with developing Bayesian networks in many fields of medicine and veterinary science show that, if sensations of isotonicity are commonly acknowledged in a domain, then experts will naturally produce statements during knowledge elicitation that suggest properties of isotonicity. They in fact will do so without explicit prompting. These experiences corroborate results from educational research which suggest the existence of an intuitive reasoning rule that underlies people's tendency to recognise isotonicity, namely 'more of the input implies more of the output' [13]. As the experts' sense of isotonicity is likely to come from experiential knowledge, however, it is probably heuristic and may not hold for all situations that can possibly be encountered. As a consequence, the mathematical translation of their concept of isotonicity almost inherently differs from the mathematical concept of isotonicity formulated for Bayesian networks. Although the experts' statements appear to imply properties of isotonicity, therefore, any such properties have to be carefully verified before they can be exploited in the engineering of a network.

Based upon the above considerations, we developed a method for verifying isotonicity of Bayesian networks with domain experts. Given the properties of isotonicity that have been suggested by the experts

during knowledge elicitation, the method focuses on a relevant subset of the observable variables. For these variables, a lattice of all possible joint value assignments is constructed, which subsequently is enhanced with probabilistic information about the effects of these assignments on the probability distribution over the output variable of interest. The enhanced lattice then is used for identifying any violations of the suggested properties of isotonicity within the network at hand. The experts subsequently are presented with these violations by means of pairs of vignettes stated in their domain's terminology, and are asked to carefully study the properties of isotonicity that are not matched by the network. Our method thereby provides both for verifying any implied isotonicities in the network and for verifying the experts' statements suggesting isotonicity.

We applied our method for verifying monotonicity to a Bayesian network in veterinary science. In recent years, we developed a network for the early detection of classical swine fever in individual animals. Both the network's structure and its associated probabilities were elicited from two domain experts. During the elicitation interviews, the experts had produced several statements that reflected their sensations of isotonicity. We verified the isotonicities that were suggested for some of the observable input variables with our method. We found a relatively small number of violations of the implied properties of isotonicity in our network and presented these violations to two independent veterinarians. The results from the interviews showed that the network should indeed have been isotone in the observable variables under consideration and that the identified violations were indicative of modelling inadequacies.

The paper is organised as follows. In Section 2, we briefly describe our Bayesian network for classical swine fever. In Section 3, we review the mathematical concept of monotonicity defined for Bayesian networks. In Section 4, we present our method for verifying properties of monotonicity with domain experts. We report on the application of our method in Section 5. The paper ends with our concluding observations in Section 6.

2 A Bayesian network for classical swine fever

In close collaboration with two experts from the Central Institute of Animal Disease Control in the Netherlands, we are developing a Bayesian network for the early detection of classical swine fever in individual pigs. Classical swine fever is an infectious disease of pigs, which has serious socio-economical consequences upon an outbreak. As the disease has a potential for rapid spread, it is imperative that its occurrence is detected in the early stages. The Bayesian network under construction is aimed at supporting veterinary practitioners in the diagnosis of the disease when visiting pig farms because of disease problems of unknown cause.

Classical swine fever is a viral disease. The virus causing the disease is transmitted mainly by direct contact between pigs, yet transmission by farmers is also known to occur. When a pig is infected, the virus first invades the lymphatic system. It subsequently affects the blood vessels and the immune system, which may give rise to haemorrhaging and diminished resistance to secondary infections. The virus will ultimately affect several organs and the pig will die. As a consequence of the infection, a pig will show different disease symptoms, among which are fever, neurological disorders, and skin haemorrhages. Clinical symptoms seen by the farmer or by the veterinarian are usually the first indications of the presence of classical swine fever in a herd. The disease is hard to detect, however, since its early symptoms are rather atypical and are shared to a large extent by common respiratory and gastro-intestinal infections. The disease moreover has a low incidence.

Our Bayesian network for classical swine fever currently includes 42 variables for which over 2400 parameter probabilities have been assessed. The variables in the network model the risk factors and the pathogenesis of the disease. More specifically, the network also models the clinical signs observed in a pig, to provide for diagnosis at a farm site. For the construction of the network, we held one unstructured interview in which the experts were asked to describe the domain, and 11 structured interviews in which the experts were asked detailed questions. In six of these structured interviews, the probabilities required for the network were obtained using standardised forms with questions accompanied by a probability scale containing verbal and numerical anchors [6]. Both experts were present at all interview sessions and consensus was always reached. An initial version of our network has now been completed. The graphical structure of this network is shown in Figure 1; in the sequel, we will refer to this network as the CSF

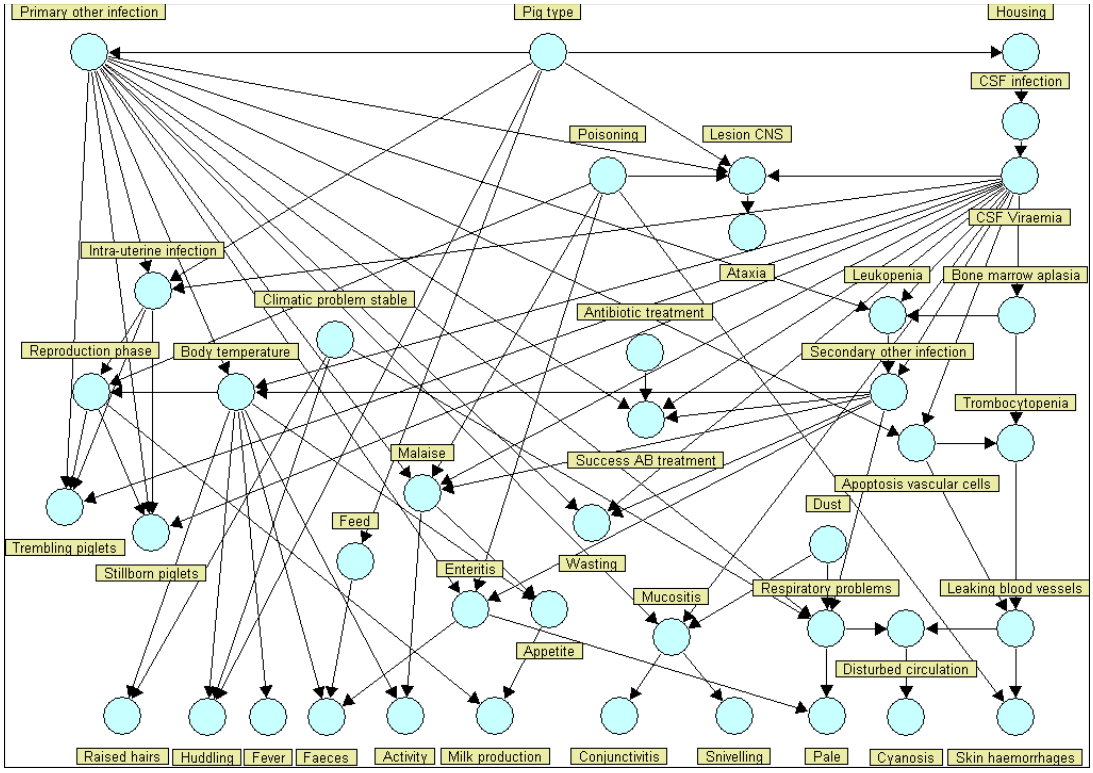

Figure 1: The graphical structure of our Bayesian network for the early detection of classical swine fever in individual pigs.

network. We currently are in the process of studying the network's performance, both by evaluating its output given real data gathered in a field test and by analysing reasoning patterns with artificial data.

3 The concept of monotonicity

Upon reviewing the mathematical concept of monotonicity for Bayesian networks, we assume that the variables of a network have different roles. We assume more specifically that the network includes a single output variable C and one or more observable input variables E; in addition, it may include an arbitrary number of intermediate variables which serve to correctly model the domain's knowledge yet cannot be observed in practice. In the CSF network, for example, the output variable models whether or not a pig has a viraemia of classical swine fever, that is, whether or not the virus has entered into the pig's system of blood vessels. Among other clinical symptoms and signs, the observable input variables include the presence or absence of skin haemorrhages and of a fever; note that these symptoms can be directly observed or measured at a farm site. An example intermediate variable, whose value cannot be observed directly on location, is the variable that models whether or not the pig suffers from bone marrow depletion. Bayesian networks with multiple input variables and a single output variable are typically found in any type of diagnostic, or classification, application. Given a joint value assignment e to their set of observable input variables, these networks are used for computing the posterior probability distribution $\Pr(C \mid e)$ for their main variable of interest C. The CSF network, for example, will be used by a veterinarian to compute the probability of a viraemia of classical swine fever for a pig showing a particular combination of symptoms and signs.

In a Bayesian network, each of the variables can adopt one of a set of discrete values. We assume that there exists a total ordering $\leq$ on such a set of values. This ordering can for example be from less

to more, or from low to high. For the purpose of studying monotonicity, the ordering to be used is best chosen to conform with the natural ordering used by experts in reasoning about their domain. In the CSF network, the orderings of the values of the variables have been chosen to range from the value that will be found in a healthy pig to the value that is the most indicative of some disease being present. The values *yes* and *no* for the variable modelling whether or not a pig has cyanosis, for example, are ordered *no* $\leq$ *yes*. The total orderings per variable are taken to induce a partial ordering $\preceq$ on the set of all joint value assignments to any subset of the network's variables. In the CSF network, for example, the combination of values *Cyanosis = yes, Abn faeces = no* is ordered higher than the combination *Cyanosis = no, Abn faeces = no*. Note that the ordering $\preceq$ on the set of joint assignments is not a total ordering. For example, in the CSF network the combination of values *Cyanosis = yes, Abn faeces = no* cannot be ordered compared to *Cyanosis = no, Abn faeces = yes*.

The concept of monotonicity for Bayesian networks now builds upon the posterior probability distributions over the output variable given the possible joint value assignments to the observable variables [4]. It is defined in terms of stochastic dominance. For a probability distribution $\Pr(C)$ over the output variable C, the cumulative distribution function $F_{\Pr}$ is defined as $F_{\Pr}(c) = \Pr(C \leq c)$ for all values c of C. For two distributions $\Pr(C)$ and $\Pr'(C)$ over C, associated with $F_{\Pr}(C)$ and $F_{\Pr'}(C)$ respectively, we say that $\Pr'(C)$ is stochastically dominant over $\Pr(C)$, denoted $\Pr(C) \leq \Pr'(C)$, if $F_{\Pr'}(c) \leq F_{\Pr}(c)$ for all values c. We now say that a Bayesian network is isotone in its observable variables E if

$$e \preceq e' \;\rightarrow\; \Pr(C \mid e) \leq \Pr(C \mid e')$$

for all joint value assignments e, e' to E. However, if

$$e \preceq e' \;\rightarrow\; \Pr(C \mid e) \geq \Pr(C \mid e')$$

for all e, e', then the network is said to be antitone in E. Informally speaking, we have that a Bayesian network is isotone in its observable input variables if entering a higher-ordered value assignment to these variables cannot make higher-ordered values of the output variable less likely.

We would like to note that, given the above definition, a network is isotone in its observable variables given the orderings $\leq$ on their sets of values if and only if the network is antitone given the reversed orderings. Although antitonicity thus is (reversely) equivalent to isotonicity, we decided to explicitly distinguish between the two types of monotonicity since a domain of application may exhibit an intricate combination of isotonicity and antitonicity for interrelated observable variables. Using orderings for the various variables that differ from the natural orderings used by the domain experts then is very likely to result in confusion in the process of verifying monotonicity.

4 A method for verifying monotonicity

As we have argued in our introduction, upon developing a Bayesian network it is important that the resulting model reflects the commonly acknowledged sensations of monotonicity in its domain of application. These sensations typically are brought up during knowledge elicitation. In our experiences, the domain experts will naturally produce statements that suggest monotonicity, even without explicit prompting from the knowledge engineer. The experts will not explicitly use the word 'monotone', though. An example statement suggesting isotonicity in our domain of application is

> *"Observing skin haemorrhages always makes classical swine fever more likely."*

We have further argued that, although the experts' statements appear to imply properties of monotonicity, the mathematical translation of their concept of monotonicity is quite likely to deviate from the mathematical concept. A method for verifying monotonicity of Bayesian networks with domain experts thus involves two verification tasks: on the one hand the statements produced by the experts have to be verified in terms of the mathematical concept of monotonicity and on the other hand the suggested properties of monotonicity have to be verified in the network.

Based upon the mathematical concept reviewed in the previous section, we designed a method for verifying monotonicity of Bayesian networks with domain experts. Given the properties of monotonicity that have been suggested by the experts during knowledge elicitation, the method focuses on a specific subset of the observable input variables. For these variables, the method constructs a lattice of all possible joint value assignments, which subsequently is enhanced with probabilistic information computed from the Bayesian network under study. From the enhanced lattice, the method identifies all violations of the properties of monotonicity of the network's output. These violations then are presented to the experts by means of pairs of vignettes for their careful consideration. The basic idea of our verification method thus is to first verify the suggested properties of monotonicity in the network and to then verify just the violated properties with the domain experts.

4.1 The assignment lattice and its use for studying monotonicity

In detailing the assignment lattice and its use, we restrict the discussion to binary variables, each of which adopts one of the values *true* and *false*; our concept of assignment lattice and the technical details of its use, however, are readily generalised to non-binary variables [5]. If a variable V has adopted the value *true*, we will write v; we use $\bar{v}$ to denote $V =$ *false*. We further take the total ordering $\leq$ with *false* $\leq$ *true* on the two values.

The *assignment lattice* for a set X of n observable input variables in essence encodes all joint value assignments to X, along with their partial ordering. For each joint value assignment x to X, we construct a set $L(x) \subseteq X$ such that $X_i \in L(x)$ if and only if $X_i =$ *true* occurs in x. From the 2^n possible value assignments to X, 2^n subsets of X are thus obtained, which with each other constitute the power set of X. From these subsets, we construct a (standard) lattice: the elements of the lattice are the various subsets of X and the links in the lattice capture the set-inclusion relation between them. We say that a set $L(x)$ *directly precedes* a set $L(x')$ in the lattice if $L(x) \subset L(x')$ and there is no set $L(x'')$ with $L(x) \subset L(x'')$ and $L(x'') \subset L(x')$, where $\subset$ is used to indicate a proper subset. Note that the set-inclusion relation of the lattice coincides with the partial ordering $\preceq$ on the joint value assignments to X. The bottom of the assignment lattice is the empty set, denoting the joint value assignment to X in which all variables have adopted the value *false*; the top of the lattice equals the set X, encoding the joint value assignment in which all observable variables have adopted the value *true*. Figure 2 depicts, as an example, the assignment lattice that is constructed for the five observable variables *Abn faeces*, *Ataxia*, *Fever*, *Malaise*, and *Skin haemorrhages* from our Bayesian network for classical swine fever. In this lattice, the element {*Abn faeces, Skin haemorrhages*}, for example, encodes a pig showing a combination of findings that indicate the presence of abnormal faeces and of pin-point bleedings of the skin; the pig does not have a fever or ataxia and also does not show the clinical picture of malaise. The element {*Abn faeces, Skin haemorrhages*} directly precedes the element {*Abn faeces, Skin haemorrhages, Fever*}, for example, and is directly preceded by the elements {*Abn faeces*} and {*Skin haemorrhages*} in the lattice.

The assignment lattice for the full set of observable input variables E of a Bayesian network captures all possible joint value assignments to E along with the partial ordering between them. To describe the effects of the various assignments on the probability distribution over the output variable C, we enhance the lattice with probabilistic information. For each element $L(e)$ of the lattice, the conditional probability $\Pr(c \mid e)$ is computed from the Bayesian network under study; this probability is associated with the element $L(e)$ in the lattice. For the assignment lattice from Figure 2, for example, we compute from the network the posterior probability $\Pr($*Viraemia = yes* $\mid$ *Abn faeces = yes, Ataxia = no, Fever = no, Malaise = no, Skin haemorrhages = yes*$) = 0.001$ to be associated with the element {*Abn faeces, Skin haemorrhages*}. For its direct successor {*Abn faeces, Skin haemorrhages, Fever*} in the lattice, we compute $\Pr($*Viraemia = yes* $\mid$ *Abn faeces = yes, Ataxia = no, Fever = yes, Malaise = = no, Skin haemorrhages = yes*$) = 0.017$.

From the definition of isotonicity, we now recall that establishing whether or not a Bayesian network is isotone in its set of observable input variables E amounts to verifying that entering a higher-ordered value assignment to E results in a stochastically dominant probability distribution over the main variable of interest or, for the binary variable C, in a higher probability of C being *true*. Since the partial ordering $\preceq$ on the value assignments coincides with the set-inclusion relation of the assignment lattice for E, we have that the lattice explicitly enumerates all pairs of probabilities to be compared for establishing isotonicity. More specifically, the network is isotone in E if $\Pr(c \mid e) \leq \Pr(c \mid e')$ for each pair of elements $L(e)$

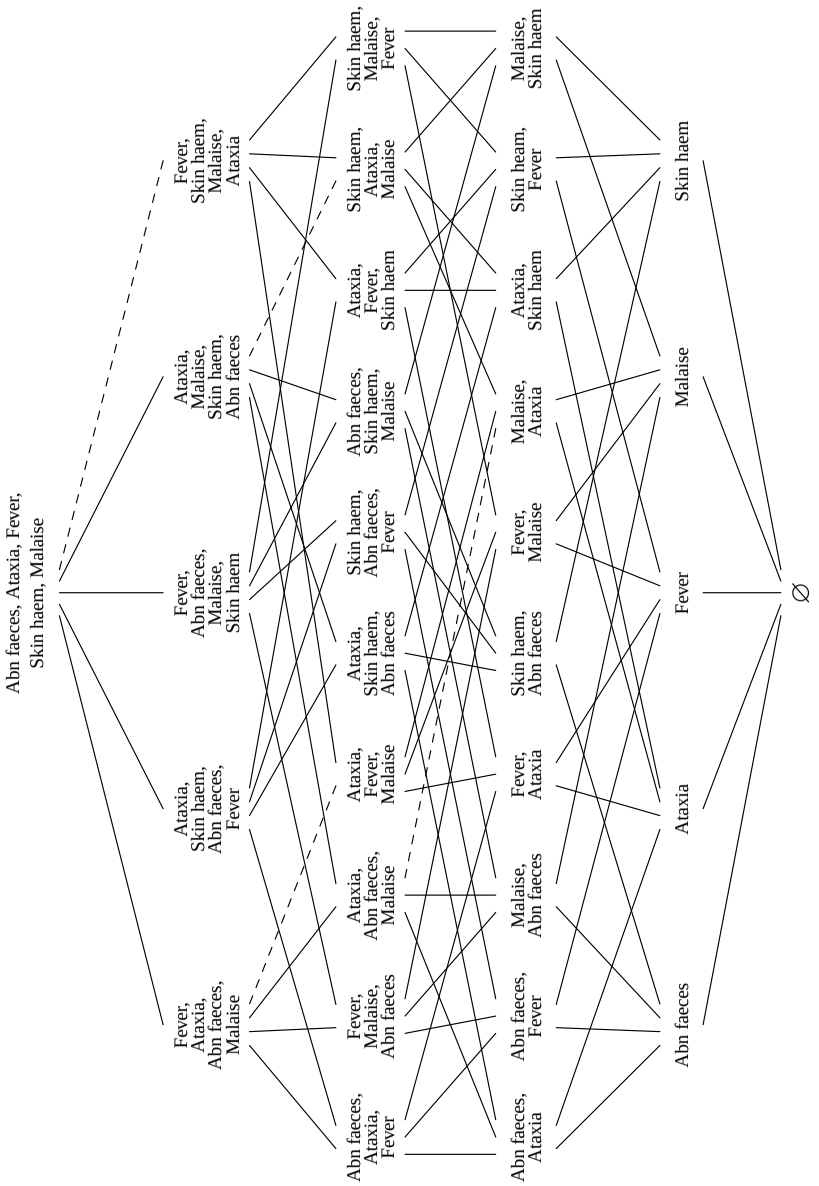

Figure 2: An assignment lattice for our Bayesian network for classical swine fever; the violated monotonicity properties are indicated by dashed lines. *Abn faeces* indicates an abnormal consistency of the pig's faeces; *Ataxia* model unsteadiness of gait; *Fever* indicates a body temperature above normal; *Skin haemorrhages* models the presence of pinpoint bleedings in the pig's skin; *Malaise* indicates a general clinical picture of illness.

and $L(e')$ in the lattice where $L(e)$ directly precedes $L(e')$. We say that the pair of probabilities $\Pr(c \mid e)$ and $\Pr(c \mid e')$ *violates* the properties of isotonicity if $\Pr(c \mid e) > \Pr(c \mid e')$. Similar observations hold for antitonicity in the set of observable variables. From the assignment lattice from Figure 2, for example, we thus need to compare the posterior probabilities associated with the element {*Abn faeces, Skin haemorrhages*} and its direct successor {*Abn faeces, Skin haemorrhages, Fever*}. Note that, since numerical ordering is transitive, we have to study the probabilities of directly linked pairs of elements in the lattice only to decide upon monotonicity. We therefore do not need to compare the probability of {*Abn faeces, Skin haemorrhages*} with that of the element {*Abn faeces, Skin haemorrhages, Fever, Ataxia*}, for example.

From the above considerations, we have that the enhanced assignment lattice can be exploited directly for establishing whether or not a Bayesian network is isotone in its observable input variables. We now observe that the lattice encodes an exponential number of value assignments to the set of input variables. Constructing the lattice and computing the various probabilities to be associated with its elements, therefore, takes exponential time and will be prohibitive for most real networks. We would like to note, however, that establishing monotonicity in a straightforward manner by directly exploiting the definition of the concept of monotonicity also takes exponential time since essentially the same probabilities need to be computed and compared. Unfortunately, the problem of establishing monotonicity of a Bayesian network is highly intractable in general [4], which renders it very unlikely that polynomial-time algorithms to this end will be found. It is not to be expected, therefore, that much more efficient methods can be constructed than using the assignment lattice as described above.

The exponential complexity of our method forestalls its use for large networks with many observable input variables. We observe, however, that the assignment lattice also provides for studying monotonicity properties of a network for a particular *subset* of the set of input variables. To this end, a relevant subset X of observable variables is selected from the network. An assignment lattice then is constructed from these variables as described above. The probabilities associated with the elements of this lattice are conditioned on a *fixed* joint value assignment e^- to the variables $E^- = E \setminus X$, that is, with each element $L(x)$ of the lattice is associated the conditional probability $\Pr(c \mid x, e^-)$. The lattice now provides for studying monotonicity for the set X *in the context of* the assignment e^-. Both the subset X and the assignment e^- for which the properties of monotonicity are to be verified, are dependent upon the domain under study and should be chosen in close consultation with the experts. For many diagnostic problems in which the observable input variables model symptoms of aberrant behaviour, for example, a suitable context assignment may be an assignment in which all context variables have adopted the value *false*. Alternatively, the graphical structure of the Bayesian network at hand may allow for identifying sets of observable input variables which are loosely interrelated and hence can be studied more or less independently.

4.2 Verifying expert statements of monotonicity

We argued in our introduction that, if sensations of monotonicity are commonly acknowledged in a domain, experts will naturally produce statements suggesting monotonicity during knowledge acquisition, even without explicit prompting from the knowledge engineer. Since the experts in making these statements may have a different conception of the idea of monotonicity than is captured by the mathematical concept reviewed above, these statements have to be carefully verified before they can be further used for engineering of the Bayesian network at hand. Now, if an expert offers statements of monotonicity during knowledge elicitation, the properties of monotonicity that are thus suggested can in essence be verified directly. To this end, all pairs of directly linked elements from the constructed assignment lattice are put to the expert for comparison. We will briefly describe how vignettes can be used for this purpose. We will then argue that the task to be performed by the expert during this verification process, although not demanding from a cognitive perspective, is quite time consuming and moreover tends to be annoying. To reduce the time that is required from the expert and to forestall irritation, we therefore propose to not verify the various properties of monotonicity directly, but to study instead only the violations of these properties identified from the network using the assignment lattice.

We begin by observing that the knowledge to be acquired to confirm the mathematical properties of the monotonicities suggested by an expert, concerns orderings of conditional probabilities. We consider again a set X of observable input variables and an output variable C. For any two joint value assignments x and

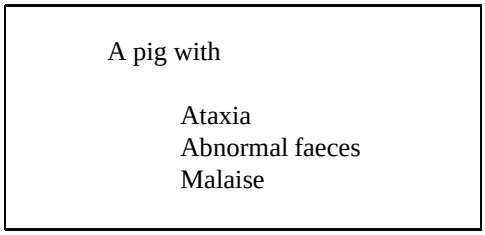

Figure 3: A vignette for the domain of classical swine fever.

x' with $x \preceq x'$, the expert more specifically will have to indicate which of the two probabilities $\Pr(c \mid x)$ and $\Pr(c \mid x')$ is the largest. By building upon the associated assignment lattice and the transitivity of the numerical ordering of probabilities, the expert will in essence have to perform this task for all pairs of value assignments x and x' of which the set $L(x)$ directly precedes the set $L(x')$ in the lattice.

To reduce the cognitive effort that is asked of the experts in the comparison task, we propose to present the various probabilities in a way that is easily accessible for them. For each probability, a description of a concrete case is constructed. By doing so, we situate the task to be performed in the experts' working practice, that is, in the situation in which his experiential knowledge is acquired; thus situating the task is more likely to result in relevant and correct information [2, 3]. More specifically, for the probability $\Pr(c \mid x, e^-)$, where e^- is the context assignment introduced above, a case is constructed with the evidence that is described by x, e^-. For each constructed case, a separate *case card* or *vignette* is created for presentation to the expert. These vignettes are stated in the expert's domain language, again to situate the task and to provide as many recognition cues as possible. Figure 3 shows, as an example, a vignette describing a value assignment to the five observable variables *Abn faeces*, *Ataxia*, *Fever*, *Malaise*, and *Skin haemorrhages* from our Bayesian network for classical swine fever. The description on the vignette mentions just the variables for which the higher-ordered value has been observed; the other variables are implicitly taken to have adopted their lower-ordered value. By explicitly stating just the aberrant values, we build upon the assumption that the experts will fill in normal values for the variables that are not mentioned explicitly. We thereby build upon the implicit assumptions humans make in observational reasoning. For each pair of vignettes thus created, the experts are asked to indicate in which of the two described cases the output c is more likely to occur. Note that by thus using vignettes the experts are never asked to actually provide or peruse numerical probabilities.

While human experts tend to feel uncomfortable expressing their knowledge and experience in terms of probabilities and are known to provide imperfectly calibrated assessments, they typically are able to state probabilistic information of a semi-numerical or qualitative nature with relative conviction and clarity, and with less cognitive effort [9]. Experts, for example, can often easily indicate which of two cases is the least likely to occur. In addition to requiring less cognitive effort, such relative judgements tend to be more reliable than direct numerical assessments [10]. Based upon these observations, we feel that the results from the elicitation procedure outlined above will be quite reliable. We would like to mention that in domains in which experts have little experience with reasoning about single cases in terms of probabilities and likelihoods, the constructed cases can be presented using the frequency format [7, 8].

Even though the task of comparing two probabilities is not very demanding on the part of the experts, direct confirmation of all suggested properties of monotonicity as outlined above requires a large number of such comparisons. By building upon the assignment lattice for n observable variables in fact,

$$\sum_{i=0}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) \cdot (n-i) > 2^n$$

pairs of probabilities have to be compared. For five observable variables, for example, the experts will have to perform as many as 80 comparisons. Since statements implying monotonicity tend to come up naturally during the elicitation interviews with the domain experts while constructing a Bayesian network, we feel that verifying all suggested properties of monotonicity explicitly is too time consuming. The elicitation will

moreover give the impression of unnecessary duplication and, as a result, is likely to generate irritation. We therefore propose to apply the method described above only for the pairs of conditional probabilities that are identified from the assignment lattice as violating the monotonicity properties that were implied by the experts in their statements.

5 Application of the verification method

During the elicitation interviews held for the construction of our Bayesian network for classical swine fever, the veterinary experts involved had made various statements that suggested properties of monotonicity for the observable input variables. To study whether or not our network adhered to the mathematical properties thus implied, we used the verification method described in the previous section for various subsets of observable variables. In this section, we review the results that we obtained for one such subset.

We consider the five observable input variables *Abn faeces*, *Ataxia*, *Fever*, *Malaise*, and *Skin haemorrhages* from our network and take for our output variable the variable *CSF Viraemia*. From the five input variables, we constructed an assignment lattice as depicted in Figure 2. The lattice includes $2^5 = 32$ elements to capture all possible joint value assignments to the five variables under study; it further includes 80 direct set-inclusion statements. Before the lattice could be enhanced with posterior probabilities of the presence of a viraemia of classical swine fever, we had to decide upon the context in which the properties of monotonicity would be verified. For this context, we decided to take a suckling piglet in which all remaining observable variables had adopted the lower-ordered value found in healthy pigs. We chose this assignment since the various aberrant clinical signs have a rather small probability of occurrence and moreover it is highly unlikely to find a large number of such signs in a single live pig. Our choice of assignment further had the advantage of fitting in with the mental model of humans which presumes signs that are not mentioned explicitly to be absent. Note that, if we would have chosen a different context, we would have had to adapt the vignettes to include the clinical signs that were presumed to be present in the context. Given the chosen context, we computed the various conditional probabilities to be associated with the elements of the assignment lattice.

For each pair of directly linked elements from the assignment lattice, we compared the computed posterior probabilities of a viraemia of classical swine fever. We found four violations of the monotonicity properties that had been implied by the experts during the elicitation interviews. These violations all pertained to adding the clinical sign of abnormal faeces to a combination of findings including the presence of ataxia and malaise. All violations arose from minor differences in the computed posterior probabilities. The largest difference was found when adding the sign of abnormal faeces to the combination of ataxia and malaise. Contrary to the property implied by the experts, the posterior probability of a viraemia of classical swine fever being present dropped from 0.017 to 0.014 by adding the sign of diarrhoea to the combination of ataxia and malaise. Note that the four violations of the monotonicity properties identified from our network show a clear pattern of regularity, in the sense that once a violation has arisen, adding further signs cannot remove it. We say that the combination is ataxia and malaise is the *context of offence* for the entire set of violations.

We presented the pairs of assignments underlying the identified violations of monotonicity to two independent veterinarians using vignettes as described in the previous section and asked them to indicate the pig that would be more likely to have a viraemia of classical swine fever. Note that we thus asked the veterinarians to perform four comparisons rather than the 80 comparisons that would have been necessary for eliciting the properties of monotonicity among the five variables involved directly. Upon being confronted with the pairs of vignettes, the two veterinarians independently and with conviction indicated that the probability of a viraemia of classical swine fever should increase upon finding the additional sign of abnormal faeces. Both veterinarians mentioned that the combination of ataxia and abnormal faeces especially served to point to classical swine fever; within the scope of our network, they could not think of any other disease that would be more likely to give this combination of signs. Assuming that the properties of monotonicity that were not confirmed explicitly indeed do hold in their domain of expertise, both veterinarians thus indicated, through their orderings, that the network should indeed have been monotone in the mathematical sense for the five variables under study in the absence of any other signs.

To conclude, we would like to note that during the interviews the two veterinarians mentioned that

diagnostic reasoning patterns in the domain of infectious animal diseases are not monotone in general. Both could rather easily generate, from their accumulated knowledge and experience, examples in which the output would be neither isotone nor antitone in the various clinical signs observed. As a side remark, one of the veterinarians moreover suggested that for studying monotonicity it would indeed not be necessary to include more than six observable variables, since with more aberrant clinical signs a pig would be dead.

6 Concluding observations

In this paper, we have presented a method for verifying monotonicity of Bayesian networks with the help of domain experts. The method focuses on a subset of the observable variables of a network and builds upon a lattice of possible joint value assignments to these variables. The lattice is enhanced with probabilistic information about the effects of these assignments on the probability distribution over the network's main variable of interest. The enhanced lattice then is used for identifying any violations of monotonicity. The experts subsequently are presented with these violations by means of pairs of vignettes stated in the domain's terminology, for their careful consideration. The method thus provides both for verifying any implied monotonicities in the network and for confirming statements suggesting monotonicity with the domain experts. The method has further been designed specifically so as to ask little time as well as little cognitive effort from the experts in the verification of their statements of monotonicity. In the paper, we have focused our discussion on binary variables only and on the verification of either isotonicity or antitonicity of a set of observable input variables. Our method, however, has been extended to apply to sets of observable variables of mixed monotonicity and to non-binary variables [5].

The results from applying our method for verifying monotonicity to a real network in veterinary science indicate that it presents a practicable method for studying reasoning patterns in Bayesian networks. We feel in fact that through the availability of our method, it has become worthwhile to devote additional attention to any statements made by experts during knowledge elicitation that appear to imply properties of monotonicity. So far, our method provides just for identifying violations of properties of monotonicity. Such violations may be indicative of modelling inadequacies, however, that need to be resolved upon further engineering of the network under study. In the near future, we hope to be able to further extend our method to include techniques for identifying the parts of a Bayesian network that have to be modified to ensure the required properties of monotonicity.

References

[1] N.P. Archer, S. Wang. Application of the backpropagation neural network algorithm with monotonicity constraints for two-group classification problems. *Decision Sciences* **24**, pp. 60 – 75, 1993.

[2] W.G. Chase (editor). *Cognitive skills and their acquisition*, Erlbaum, Hillsdale, NJ, 1973.

[3] N.J. Cooke. Knowledge elicitation. In: F.T. Durso, R.S. Nickerson, R.W. Schvaneveldt, S.T. Dumais, D.S. Lindsay, M.T.H. Chi (editors). *Handbook of Applied Cognition*, Wiley, London, 1999.

[4] L.C. van der Gaag, H.L. Bodlaender, A. Feelders. Monotonicity in Bayesian networks. In: M. Chickering, J. Halpern (editors). *Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence*, AUAI, Arlington, pp. 569 – 576, 2004.

[5] L.C. van der Gaag, S. Renooij, P.L. Geenen. Lattices for studying monotonicity of Bayesian networks. In: M. Studeny, J. Vomlel (editors). *Proceedings of the Third European Workshop on Probabilistic Graphical Models*, Prague, Czech Republic, pp. 99 – 106, 2006.

[6] L.C. van der Gaag, S. Renooij, C.L.M. Witteman, B. Aleman, B.G. Taal. How to elicit many probabilities. In: K.B. Laskey, H. Prade (editors). *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann Publishers, San Francisco, California, pp. 647 – 654, 1999.

[7] G. Gigerenzer, U. Hoffrage. How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review*, pp. 684 – 704, 1995

[8] E.M. Helsper, L.C. van der Gaag, A.J. Feelders, W.L.A. Loeffen, P.L. Geenen, A.R.W. Elbers. Bringing order into Bayesian-network construction. *Proceedings of the Third International Conference on Knowledge Capture*, ACM Press, New York, pp. 121 – 128, 2005.

[9] M.A. Meyer, J.M. Booker. *Eliciting and Analyzing Expert Judgment: A Practical Guide*. ASA-SIAM Series on Statistics and Applied Probability, 2001.

[10] M.G. Morgan, M. Henrion. *Uncertainty, a Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, Cambridge, 1990.

[11] R. Potharst, A. Feelders. Classification trees for problems with monotonicity constraints. *SIGKDD Explorations*, vol. 4, pp. 1 – 10, 2002.

[12] T. Robertson, F. Wright, R.L. Dykstra. *Order Restricted Statistical Inference*, Wiley, New York, 1988.

[13] R. Stavy, R. Babai, P. Tsamir, D. Tirosh, F.-L. Lin, C. McRobbie (2006). Are intuitive rules universal ? *International Journal of Science and Mathematics Education*, vol. 4, pp. 417 – 436.