Context-specific Sign-propagation in Qualitative Probabilistic Networks

Silja Renooij, Linda C. van der Gaag
Institute of Information and Computing Sciences
Utrecht University, P.O. Box 80.089, 3508 TB Utrecht
{silja, linda}@cs.uu.nl

Simon Parsons
Massachusetts Institute of Technology
Sloan School of Management
3 Cambridge Center, NE20-336, Cambridge, MA 02142
sparsons@mit.edu

Abstract

Qualitative probabilistic networks are qualitative abstractions of probabilistic networks, summarising probabilistic influences by qualitative signs. As qualitative networks model influences at the level of variables, knowledge about probabilistic influences that hold only for specific values cannot be expressed. The results computed from a qualitative network, as a consequence, can be weaker than strictly necessary and may in fact be rather uninformative. We extend the basic formalism of qualitative probabilistic networks by providing for the inclusion of context-specific information about influences and show that exploiting this information upon reasoning has the ability to forestall unnecessarily weak results.

1 Introduction

Probabilistic networks have become widely accepted as practical representations of knowledge for reasoning under uncertainty. Applications are found in such fields as (medical) diagnosis and prognosis, planning, monitoring, vision, information retrieval, natural language processing, and e-commerce. Probabilistic networks combine a graphical representation of the statistical variables in a problem domain and the relations between them, with (conditional) probabilities that represent the uncertainties involved [1]. More specifically, the graphical representation of a network takes the form of a directed graph where each node represents a variable and each arc expresses a possible probabilistic dependence between the connected variables. To capture the strengths of the represented dependences, each variable has associated a set of conditional probability distributions given the possible combinations of values for the variable's predecessors in the digraph. For reasoning with these probabilities in a mathematically correct way, powerful algorithms are available.

Qualitative probabilistic networks are qualitative abstractions of probabilistic networks [2], introduced for probabilistic reasoning in a qualitative way. Just like a probabilistic network, a qualitative network encodes statistical variables and the probabilistic dependences between them in a directed acyclic graph. Each node A in this digraph once again represents a variable. An arc $A \rightarrow B$ again expresses a probabilistic influence of the variable A on the probability distribution of the variable B. Rather than quantified by conditional probabilities as in a probabilistic network, however, the influence is summarised by a qualitative sign. This sign indicates the direction of shift in B's (cumulative) probability distribution that would be occasioned by an observation for A. For example, a positive influence of A on B expresses that observing higher values for A renders higher values for B more likely. The signs of a qualitative network have a well-defined basis in the mathematical concept of stochastic dominance. Building upon this concept, it is possible to reason with qualitative signs in a mathematically correct way. To this end, an efficient algorithm, based upon the idea of propagating and combining signs, is available [3].

Qualitative probabilistic networks can play an important role in the construction of probabilistic networks for real-life application domains. While constructing the digraph of a probabilistic network requires considerable effort, it is generally considered feasible. The assessment of all probabilities required is a much harder task, especially if it has to be performed with the help of human experts. The quantification task is, in fact, often referred to as a major bottleneck in building a probabilistic network [4, 5]. Assessment of the signs for a qualitative probabilistic network tends to require considerably less effort from human experts, however [3]. Now, by eliciting signs from domain experts for the digraph of a probabilistic network under construction, a qualitative probabilistic network is obtained. This qualitative network can be used to study and validate the reasoning behaviour of the network prior to probability assessment. Upon quantifying the network, the acquired signs can then be used as constraints on the probabilities to be assessed [6, 7].

Qualitative networks model the uncertainties involved in an application domain at the high abstraction level of variables, as opposed to probabilistic networks where uncertainties are represented at the level of the variables' values. Due to this coarse level of representation detail, reasoning with a qualitative probabilistic network often leads to results that are weaker than strictly necessary and may in fact be rather uninformative. To be able to fully exploit a qualitative probabilistic network as outlined above, we feel that it should capture and exploit as much qualitative information from the application domain as possible. First introduced by M.P. Wellman [2] and later extended by M. Henrion and M.J. Druzdzel [3, 8, 9], various researchers have refined qualitative probabilistic networks to enhance their expressiveness. S. Parsons [10, 11], for example, has introduced the concept of *qualitative derivative* where the influence of a variable A on a variable B is summarised by a set of signs, one for each value of B; he has also studied the use of other approaches to uncertain reasoning, such as order-of-magnitude reasoning, within qualitative probabilistic networks [12]. S. Renooij and L.C. van der Gaag [13] have enhanced qualitative probabilistic networks by adding a qualitative notion of strength. Renooij *et al.* [14] have further focused on identifying and resolving troublesome parts of a network. In the current paper, we propose adding a notion of *context* as an extension to the basic formalism of qualitative networks in order to enhance its expressive power.

The notion of context has been studied before in quantitative probabilistic networks. The digraph of a probabilistic network in essence captures independences between variables, that is, it models independences that hold for all values of the variables involved. The independences that hold only for specific values are not explicitly represented in the digraph but are instead captured by the conditional probabilities associated with the variables. As knowledge of the latter independences allows further decomposition of conditional probabilities and can be exploited to speed up inference, a notion of *context-specific independence* was introduced [15, 16]. Context-specific independence occurs often enough that some well-known tools for the construction of probabilistic networks have incorporated special mechanisms to allow the user to more easily specify the conditional probability distributions for the variables involved [15].

A qualitative probabilistic network also captures independences between variables by means of its digraph. Since its qualitative influences are specified at the high abstraction level of variables as well, independences that hold only for specific values of the variables involved cannot be represented. In fact, qualitative influences *hide* such context-specific independences: if the influence of a variable A on a variable B is positive in one context, that is, for one specific combination of values for some other influential variables, and zero in all other contexts, which indicates independence, then the influence is captured by a positive sign. We note that a sign may hide not just independences, but also context-specific positive and negative influences: if a variable A has a positive influence on a variable B in some context and a negative influence in another context, then the influence of A on B is modelled as being ambiguous.

As knowledge of context-specific independences basically is qualitative by nature, we feel that it can and should be captured explicitly in a qualitative probabilistic network. For this purpose, we introduce the notion of *context-specific sign*. A context-specific sign is basically a function that defines different signs for different contexts. In a qualitative network, each influence is now associated with such a context-specific sign. Upon reasoning with the thus extended network, for each influence the appropriate sign is determined as the sign that is defined for the context corresponding to the observed variables' values. We show that exploiting this context-specific information upon reasoning can prevent unnecessarily weak results.

The paper is organised as follows. In Section 2, we provide some preliminaries concerning probabilistic networks and qualitative probabilistic networks. We present, in Section 3, two examples of the type of information that can be hidden in qualitative influences. We introduce our extended formalism and associated algorithm for exploiting context-specific information upon reasoning in Section 4. In Section 5, we discuss the context-specific information that is hidden in the qualitative abstractions of two real-life probabilistic networks. The paper ends with our concluding observations in Section 6.

2 Preliminaries

Before introducing qualitative probabilistic networks, we briefly review their quantitative counterparts.

2.1 Probabilistic networks

A probabilistic network $B = (G, \Pr)$ is a concise representation of a joint probability distribution Pr on a set of statistical variables. It encodes the variables concerned, along with their probabilistic interrelationships, in an acyclic directed graph $G = (V(G), A(G))$. Each node $A \in V(G)$ represents one of the statistical variables. Variables will be indicated by capital letters from the beginning of the alphabet; the values of these variables will be denoted by small letters, possibly with a subscript. As there is a one-to-one correspondence between nodes and variables, we will use the terms 'node' and 'variable' interchangeably. The probabilistic relationships between the represented variables are captured by the set of arcs $A(G)$ of the digraph. Informally speaking, we take an arc $A \rightarrow B$ in G to represent an influential relationship between the variables A and B, designating B as the effect of cause A. The absence of an arc between two variables means that they do not influence each other directly. More formally, the set of arcs captures probabilistic independence among the represented variables by means of the d-separation criterion [1]. Two variables are said to be *d-separated* if all chains between them are blocked by the available observations. We say that a chain between two variables is *blocked* if it includes either an observed variable with at least one outgoing arc or an unobserved variable with two incoming arcs and no observed descendants; a chain that is not blocked is called *active*. If two variables are d-separated then they are considered conditionally independent given the available evidence. Associated with each variable $A \in V(G)$ in the digraph is a set of conditional probability distributions $\Pr(A \mid \pi(A))$ that describe the probabilistic relationship of this variable with its (immediate) predecessors $\pi(A)$ in the digraph.

Example 2.1 *We consider the small probabilistic network shown in Figure 1. The network represents a fragment of fictitious and incomplete medical knowledge, pertaining to the effects of administering antibiotics on a patient. Variable A models whether or not a patient has been taking antibiotics; $A = true$, or 'a' for short, represents that a patient has been taking*

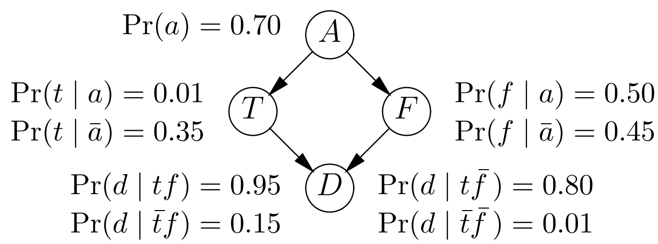

Figure 1: The *Antibiotics* network.

antibiotics, and $A = false$*, or '*$\bar{a}$*', represents that a patient has not been taking antibiotics. Variable* T *models whether or not the patient is suffering from typhoid fever and variable* D *represents presence or absence of diarrhoea in the patient. Variable* F*, to conclude, describes whether or not the composition of the bacterial flora in the patient's gastrointestinal tract has changed. Typhoid fever and a change in the patient's bacterial flora are modelled as the possible causes of diarrhoea. Antibiotics can cure typhoid fever by killing the bacteria that cause the infection. However, antibiotics can also change the composition of the patient's bacterial flora, thereby increasing the risk of diarrhoea.*

The extent to which the variables influence each other is captured by the network's (conditional) probability distributions. The conditional probability distributions of the variable D*, for example, reveal that a patient with typhoid fever almost certainly will be suffering from diarrhoea, regardless of whether or not the composition of his bacterial flora has changed.* □

A probabilistic network $B = (G, \Pr)$ defines a unique joint probability distribution Pr on $V(G)$ with

$$\Pr(V(G)) = \prod_{A \in V(G)} \Pr(A \mid \pi(A))$$

that respects the independences portrayed in the digraph G. Since a probabilistic network captures a unique distribution, it provides for computing any prior or posterior probability of interest over its variables. Although computing such probabilities is known to be NP-hard [17], various powerful algorithms are available that have a polynomial runtime complexity for most realistic networks [1, 18].

2.2 Qualitative probabilistic networks

Qualitative probabilistic networks bear a strong resemblance to their quantitative counterparts. A qualitative probabilistic network $Q = (G, \Delta)$ also comprises an acyclic digraph $G = (V(G), A(G))$ modelling statistical variables and the probabilistic relationships between them. The set of arcs $A(G)$ again models probabilistic independence. Instead of conditional probability distributions, however, a qualitative probabilistic network associates with its digraph a set Δ of qualitative influences and qualitative synergies.

A *qualitative influence* between two variables expresses how the values of one variable influence the probability distribution over the values of the other variable. The direction of the shift in the distribution occasioned is indicated by the *sign* of the influence. A *positive qualitative influence* of a variable A on a variable B, for example, expresses that observing higher values for A makes higher values for B more likely, regardless of any other influences on B. Building upon a total ordering '$>$' on the values per variable, we have that higher values for a variable B are more likely given higher values for a variable A, if the *cumulative* conditional probability distribution $F_{B|a_i}$ of variable B given a_i lies, graphically speaking, below the cumulative conditional probability distribution $F_{B|a_j}$ given a_j, for all values a_i, a_j of A with $a_i > a_j$. When $F_{B|a_i}$ lies below $F_{B|a_j}$ for all values of B, we say that $F_{B|a_i}$ dominates $F_{B|a_j}$ by *first-order stochastic dominance* (FSD) [2]:

$$F_{B|a_i} \; FSD \; F_{B|a_j} \iff F_{B|a_i}(b_i) \leq F_{B|a_j}(b_i) \text{ for all values } b_i \text{ of } B$$

The concept of first-order stochastic dominance underlies the formal definition of qualitative influence.

Definition 2.2 *Let $G = (V(G), A(G))$ be an acyclic digraph and let* Pr *be a joint probability distribution on $V(G)$ that respects the independences in G. Let A, B be variables in G with $A \rightarrow B \in A(G)$, and let $X = \pi(B) \setminus \{A\}$ be the set of all predecessors of B other than A. Then, variable A* positively influences *variable B along arc $A \rightarrow B$, written $S^+(A, B)$, iff for all values b_i of B and a_j, a_k of A, $a_j > a_k$, we have that*

$$\Pr(B \geq b_i \mid a_j x) \;\geq\; \Pr(B \geq b_i \mid a_k x)$$

for any combination of values x for the set X.

A *negative qualitative influence*, denoted by S^-, and a *zero qualitative influence*, denoted by S^0, are defined analogously, replacing $\geq$ in the above formula by $\leq$ and $=$, respectively. If the influence of variable A on variable B is not monotonic or if it is unknown, we say that it is *ambiguous*, denoted $S^?(A, B)$. The '+', '−', '0' and '?' in these definitions are termed the *signs* of the qualitative influences.

In the remainder of this paper, we assume, for ease of exposition, that all variables are binary valued, with a denoting $A = \mathit{true}$, $\bar{a}$ denoting $A = \mathit{false}$, and $a > \bar{a}$ for any binary variable A. For illustrative purposes in examples, binary variables often have different values than *true* and *false*; value statements for these variables however, are again written as a or $\bar{a}$. We note that for binary variables the definition of qualitative influence can be slightly simplified. For a positive qualitative influence of A on B, for example, we now have that

$$\Pr(b \mid ax) - \Pr(b \mid \bar{a}x) \geq 0$$

for any combination of values x for the set $X = \pi(B) \setminus \{A\}$.

A qualitative influence is associated with each arc in the digraph of a qualitative network. Variables, however, not only influence each other directly along arcs, they can also exert indirect influences on one another. The definition of qualitative influence trivially extends to capture such indirect influences along active chains. The signs of these indirect influences are determined by the properties that the set of influences of a qualitative probabilistic network exhibits [2]. The property of *symmetry*, for example, guarantees that, if the network includes the influence $S^\delta(A, B)$, then it also includes the reverse influence $S^\delta(B, A)$ with the same sign δ. The property of *transitivity* asserts that qualitative influences along an active chain without any variables with two incoming arcs on the chain, combine into an indirect influence whose sign equals the product of the signs of the separate influences along the chain; the product of the signs is defined by the $\otimes$-operator from Table 1. The property of *composition* asserts that multiple qualitative influences between two variables along parallel active chains combine into a composite influence with a sign as specified by the $\oplus$-operator.

From Table 1, we observe that combining non-ambiguous qualitative influences with the $\oplus$-operator can yield influences with an ambiguous sign. Such an ambiguity, in fact, results whenever two influences with opposite signs are combined. The two influences in essence are conflicting and represent a *trade-off* in the application domain. The ambiguity that results from

$\otimes$	$+$	$-$	0	$?$
$+$	$+$	$-$	0	$?$
$-$	$-$	$+$	0	$?$
0	0	0	0	0
$?$	$?$	$?$	0	$?$

$\oplus$	$+$	$-$	0	$?$
$+$	$+$	$?$	$+$	$?$
$-$	$?$	$-$	$-$	$?$
0	$+$	$-$	0	$?$
$?$	$?$	$?$	$?$	$?$

Table 1: The $\otimes$- and $\oplus$-operators for combining signs.

combining the two conflicting influences indicates that the trade-off cannot be *resolved* from the information that is represented in the network. In contrast with the $\oplus$-operator, the $\otimes$-operator cannot introduce ambiguities upon combining signs of influences along chains. Note that, once an ambiguous result has arisen, both operators serve to propagate the ambiguity.

In addition to influences, a qualitative probabilistic network includes synergies that model interactions within small sets of variables. We distinguish between *additive synergies* and *product synergies*. As we will not use the additive synergy in the remainder of this paper, we just say that it captures the joint influence of two variables on a common successor [2]. A *product synergy* expresses how the value of one variable influences the probabilities of the values of another variable in view of an observed value for a third variable [8].

Definition 2.3 *Let $G = (V(G), A(G))$ be an acyclic digraph and let* Pr *be a joint probability distribution on $V(G)$ that respects the independences in G. Let A, B, C be variables in G with $A \rightarrow C$, $B \rightarrow C \in A(G)$. Then, variable A exhibits a* negative product synergy *on variable B (and vice versa) given the value c for their common successor C, denoted $X^{-}(\{A,B\},c)$, iff*

$$\Pr(c \mid abx) \cdot \Pr(c \mid \bar{a}\bar{b}x) \; \leq \; \Pr(c \mid a\bar{b}x) \cdot \Pr(c \mid \bar{a}bx)$$

for any combination of values x for the set $X = \pi(C) \setminus \{A,B\}$.

Positive, *zero*, and *ambiguous product synergies* are defined analogously.

Product synergies are of importance for reasoning with a qualitative network since they induce a qualitative influence between the predecessors A and B of a variable C upon its observation. Such an induced influence is coined an *intercausal influence*. The sign of an intercausal influence is determined by the product synergy that served to induce it and may differ for the observations c and $\bar{c}$ for the variable C.

Example 2.4 *The qualitative probabilistic network shown in Figure 2 is the qualitative counterpart of the* Antibiotics *network discussed in Example 2.1. From the conditional probability distributions specified for the variable D, we observe that*

$$\Pr(d \mid tf\,) - \Pr(d \mid \bar{t}f\,) = 0.95 - 0.15 \geq 0 \quad \text{and}$$

$$\Pr(d \mid t\bar{f}\,) - \Pr(d \mid \bar{t}\bar{f}\,) = 0.80 - 0.01 \geq 0$$

and therefore conclude that $S^+(T, D)$. We further find that $S^+(A, F)$, $S^-(A, T)$, and $S^+(F, D)$. The signs of these influences are shown over the corresponding arcs in the qualitative network. In addition, we observe that

$$\Pr(d \mid tf\,) \cdot \Pr(d \mid \bar{t}\bar{f}\,) - \Pr(d \mid t\bar{f}\,) \cdot \Pr(d \mid \bar{t}f\,) = 0.95 \cdot 0.01 - 0.80 \cdot 0.15 \leq 0$$

and

$$\Pr(\bar{d} \mid tf\,) \cdot \Pr(\bar{d} \mid \bar{t}\bar{f}\,) - \Pr(\bar{d} \mid t\bar{f}\,) \cdot \Pr(\bar{d} \mid \bar{t}f\,) = 0.05 \cdot 0.99 - 0.20 \cdot 0.85 \leq 0$$

We conclude that either value for the variable D induces a negative intercausal influence between the variables T and F, that is, $X^-(\{T, F\}, d\,)$ and $X^-(\{T, F\}, \bar{d}\,)$. The intercausal influences that can be induced are represented in the qualitative network by a dotted line, over which the associated signs are shown. □

For reasoning with a qualitative probabilistic network, an elegant algorithm is available from M.J. Druzdzel and M. Henrion [3]; this algorithm, termed the *sign-propagation algorithm*, is summarised in pseudocode in Figure 3. The basic idea of the algorithm is to trace the effect of observing a value for a variable upon the probabilities of the values of all other variables in

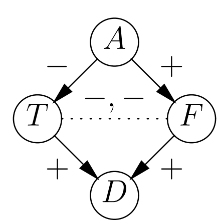

Figure 2: The qualitative *Antibiotics* network.

```
procedure PropagateSign(trail, from, to, messagesign):

    sign[to] ← sign[to] ⊕ messagesign;
    trail ← trail ∪{to};
    for each active neighbour V_i of to
    do linksign ← sign of (induced) influence between to and V_i;
       messagesign ← sign[to] ⊗ linksign;
       if V_i ∉ trail and sign[V_i] ≠ sign[V_i] ⊕ messagesign
       then PropagateSign(trail, to, V_i, messagesign).
```

Figure 3: The sign-propagation procedure for inference in a qualitative network.

the network by message passing between neighbouring variables. In essence, the algorithm computes the sign of influence along all active chains between the newly observed variable and all other variables in the network, using the properties of symmetry, transitivity and composition. For each variable, it summarises the overall influence in a *node sign* that indicates the direction of the shift in the probability distribution of that variable occasioned by the new observation.

The sign-propagation algorithm takes for its input a qualitative probabilistic network, a set of previously observed variables, a variable for which a new observation has become available, and the sign of this observation, that is, either a '+' for the value *true* or a '−' for the value *false*. Prior to the actual propagation of the new observation, for all variables V_i the node sign sign[V_i] is initialised at '0'. For the newly observed variable the appropriate sign is now entered into the network. The observed variable updates its node sign to the sign-sum of its original sign and the entered sign. It thereupon notifies all its (induced) neighbours that its sign has changed, by passing to each of them a message containing a sign. This sign is the sign-product of the variable's current node sign and the sign *linksign* of the influence associated with the arc or intercausal link it traverses. Each message further records its origin in the variable *trail*; this information is used to prevent the passing of messages to variables that were already visited on the same simple chain. Upon receiving a message, a variable *to* updates its node sign to the sign-sum of its current node sign sign[*to*] and the sign *messagesign* from the message it receives. The variable then sends a copy of the message to all its neighbours that need to reconsider their node sign. In doing so, the variable changes the sign in each copy to the appropriate sign and adds itself to *trail* as the origin of the copy. As this process is repeated throughout the network, the chains along which messages have been passed are thus recorded. Note that, as messages travel simple chains only, it is sufficient to just record the variables on these chains.

During sign-propagation, the variables of a qualitative network are only visited if they need a change of node sign. A node sign can change at most twice, once from '0' to '+', '−' or '?' and then only from '+' or '−' to '?'. From this observation we have that no variable is ever activated more than twice upon inference. The algorithm is therefore guaranteed to halt. The time complexity of the algorithm is linear in the number of arcs of the digraph of the network.

Example 2.5 *We illustrate the sign-propagation algorithm by once again considering the qualitative* Antibiotics *network from Figure 2. Suppose that a specific patient is taking antibiotics. This observation is entered into the network by updating the node sign of variable A to a '+'. Variable A thereupon propagates a message with sign $+ \otimes - = -$ towards variable T. T updates its node sign to '$-$' and sends a message with sign $- \otimes + = -$ to variable D. D updates its sign to '$-$'. It does not pass on a sign to variable F, since the chain from A to F through D is blocked. Variable A also sends a message, with sign $+ \otimes + = +$, to F. Variable F updates its node sign accordingly and passes a message with sign $+ \otimes + = +$ to variable D. D thus receives the additional sign '+'. This sign is combined with the previously updated node sign '$-$', which results in the ambiguous sign $- \oplus + = ?$ for D. Note that the ambiguous sign arises from the trade-off that is represented by the two chains from A to D. Also note that if the network would have contained additional variables beyond D, these variables would have all ended up with the sign '?' after inference.* □

3 Context-independent signs

Since qualitative probabilistic networks model knowledge at the abstraction level of variables, context-specific information, that is, information that holds only for specific values of the variables involved, cannot be represented explicitly. This information in essence is hidden in the qualitative influences and synergies of the network. If, for example, the influence of a variable A on a variable B is strictly positive for one combination of values for the set X of B's predecessors other than A, and zero for all other combinations of values for X, then the influence of A on B is positive by definition. The zero influences, indicating context-specific independence, are hidden due to the fact that the inequality in the definition of qualitative influence is not strict. We present an example to illustrate such hidden zeroes.

Example 3.1 *We consider the qualitative probabilistic network from Figure 4, which represents a highly simplified fragment of knowledge in oncology. It pertains to the effects and complications to be expected from treatment of oesophageal cancer. The variable L models the life expectancy of a patient after therapy; the value l indicates that the patient will survive for at least one year and the value $\bar{l}$ expresses that the patient will die within this year. Variable T models the therapy instilled; we consider surgery, modelled by t, and no treatment, modelled by $\bar{t}$, as the only therapeutic alternatives. The effect to be attained from surgery is a complete removal of the tumour, modelled by the variable R. After surgery a life-threatening pulmonary complication, modelled by P, may result; the occurrence of this complication is heavily influenced by whether or not the patient is a smoker, which is modelled by the variable S.*

We consider the conditional probabilities from a quantified network representing the same knowledge. We would like to note that these probabilities serve illustrative purposes only: although not entirely unrealistic, they have not been specified by domain experts. The probability of attaining a complete removal of an oesophageal tumour upon surgery is $\Pr(r \mid t) = 0.45$*; as without surgery there can be no removal of the tumour, we have* $\Pr(r \mid \bar{t}\,) = 0$*. From* $\Pr(r \mid t) \geq \Pr(r \mid \bar{t}\,)$*, we have that the variable T indeed exerts a positive*

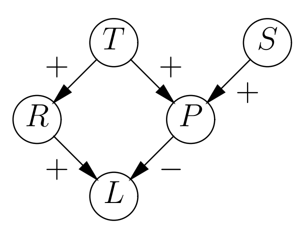

Figure 4: The qualitative *Surgery* network.

qualitative influence on R. The probabilities of a pulmonary complication occurring and of a patient's life expectancy after therapy are, respectively,

$\Pr(p \mid TS)$	s	$\bar{s}$
t	0.75	0.00
$\bar{t}$	0.00	0.00

$\Pr(l \mid RP)$	p	$\bar{p}$
r	0.15	0.95
$\bar{r}$	0.03	0.50

From the rightmost table we observe that $\Pr(l \mid r\,P) \geq \Pr(l \mid \bar{r}\,P)$ *for all values of P and that* $\Pr(l \mid p\,R) \leq \Pr(l \mid \bar{p}\,R)$ *for all values of R. We thus verify that the variable R exerts a positive influence on L, indicating that successful removal of the tumour serves to increase life expectancy, and that the qualitative influence of P on L is negative, indicating that pulmonary complications from surgery are indeed life threatening. From the leftmost table, we observe that* $\Pr(p \mid s\,T) \geq \Pr(p \mid \bar{s}\,T)$ *for all values of T and* $\Pr(p \mid t\,S) \geq \Pr(p \mid \bar{t}\,S)$ *for all values of S. Both T and S thus exert a positive qualitative influence on the variable P, indicating that performing surgery and smoking are risk factors for pulmonary complications. The zeroes in the table reveal that pulmonary complications are likely to occur only in the presence of* both *risk factors. The fact that the influence of T on P is, for example, actually zero in the context of the value $\bar{s}$ for the variable S, however, is not apparent from the sign of the influence. Note that this zero influence does not arise from the probabilities being zero, but rather from their having the same value.* □

The previous example shows that the level of representation detail of a qualitative network can result in *information hiding*. As hidden information cannot be exploited upon reasoning, unnecessarily weak answers may result from inference with the network. Referring to the previous example, for instance, we can derive that performing surgery on a non-smoker has a positive influence on life expectancy: as $\Pr(l \mid t\bar{s}) = 0.70$ and $\Pr(l \mid \bar{t}\bar{s}) = 0.50$[1], we

[1] The values $\Pr(l \mid \bar{t}\bar{s}) = 0.50$ and $\Pr(l \mid t\bar{s}) = 0.70$ are computed by conditioning on the values of R and P and using the fact that T and S are independent of L given R and P. For example,

$$\begin{aligned}
\Pr(l \mid t\bar{s}) &= \Pr(l \mid rp)\cdot\Pr(r \mid t)\cdot\Pr(p \mid t\bar{s}) + \Pr(l \mid \bar{r}p)\cdot\Pr(\bar{r} \mid t)\cdot\Pr(p \mid t\bar{s}) \;+ \\
&\quad \Pr(l \mid r\bar{p})\cdot\Pr(r \mid t)\cdot\Pr(\bar{p} \mid t\bar{s}) + \Pr(l \mid \bar{r}\bar{p})\cdot\Pr(\bar{r} \mid t)\cdot\Pr(\bar{p} \mid t\bar{s}) \\
&= 0.15\cdot 0.45\cdot 0 + 0.03\cdot 0.55\cdot 0 + 0.95\cdot 0.45\cdot 1 + 0.50\cdot 0.55\cdot 1 \\
&= 0.7025
\end{aligned}$$

have that $\Pr(l \mid t\bar{s}) \geq \Pr(l \mid \bar{t}\bar{s})$. In the qualitative network, however, upon entering the observation t for the variable T, in the presence of $\bar{s}$, inference will result in a '?' for L due to the conflicting reasoning chains from T to L. The '?' for the variable L indicates that the resulting influence is unknown. As, from the context $\bar{s}$, we know that the influence of T on P is zero, and hence the influence of T on L via P is zero, the result from qualitative inference is weaker than strictly necessary.

We recall from the definition of qualitative influence that the sign of the influence of a variable A on a variable B must hold for all combinations of values for the set X of predecessors of B other than A. A '?' for the influence may therefore hide the information that A has a positive influence on B for some combination of values for X and a negative influence for another combination. If so, the ambiguous influence of A on B is *non-monotonic* in nature and can in fact be looked upon as specifying different signs for different contexts. We present an example to illustrate this observation.

Example 3.2 *The qualitative network from Figure 5 represents another fragment of knowledge in oncology, this time pertaining to the metastasis of oesophageal cancer. The variable L represents the location of the primary tumour in a patient's oesophagus; the value l models that the tumour resides in the lower two-third of the oesophagus and the value $\bar{l}$ expresses that the tumour is in the oesophagus' upper one-third. An oesophageal tumour upon growth typically gives rise to lymphatic metastases. The extent of such metastases is captured by the variable M. The value $\bar{m}$ of M indicates that just the local and regional lymph nodes are affected; m denotes that distant lymph nodes are affected. Which lymph nodes are local or regional and which are distant depends on the location of the primary tumour in the oesophagus. The lymph nodes in the neck, or cervix, for example, are regional for a tumour in the upper one-third of the oesophagus and distant otherwise. Variable C represents the presence or absence of metastases in the cervical lymph nodes.*

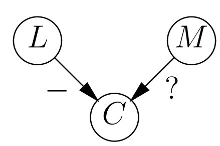

Figure 5: The qualitative *Cervical metastases* network.

We consider the conditional probabilities from a quantified network representing the same knowledge; these probabilities again serve illustrative purposes only. The probabilities of the presence of cervical metastases in a patient are

$\Pr(c \mid ML)$	l	$\bar{l}$
m	0.35	0.95
$\bar{m}$	0.00	1.00

From these probabilities we observe that the variable L indeed has a negative influence on C, indicating that tumours in the lower two-third of the oesophagus are less likely to give rise to lymphatic metastases in the neck than tumours that are located in the upper one-third of the oesophagus. The influence of the variable M on C, however, is non-monotonic:

$$\Pr(c \mid ml) > \Pr(c \mid \bar{m}l), \text{ yet } \Pr(c \mid m\bar{l}\,) < \Pr(c \mid \bar{m}\bar{l}\,)$$

While for tumours in the lower two-third of the oesophagus the lymph nodes in the neck are less likely to be affected when only local and regional metastases are present, they are more likely to be affected for tumours that are located in the upper one-third of the oesophagus. We conclude that the non-monotonic influence of M on C hides a '$+$' for the value l of the variable L and a '$-$' for the context $\bar{l}$. □

With the two examples above we have illustrated that context-specific information about influences that is present in the conditional probabilities of a quantified network cannot be represented explicitly in a qualitative probabilistic network. Upon abstracting the quantified network to the qualitative network, the information is effectively hidden. Of course, in real-life applications of qualitative probabilistic networks, the qualitative network would be built directly with the help of domain experts rather than computed from an already quantified network. During the construction of the qualitative network, however, an expert may express knowledge about non-monotonicities and context-specific independences as discussed above, which cannot be represented explicitly and will thus be hidden in the various signs of the network.

4 Context-specificity and its exploitation

The level of representation detail of a qualitative probabilistic network enforces the signs of influences and synergies to be independent of specific contexts. In this section we present an extension to the basic formalism of qualitative networks that allows for associating context-specific signs with qualitative influences and synergies. In Section 4.1, the extended formalism is introduced; in Section 4.2, we demonstrate, by means of the example networks from the previous section, that exploiting context-specific information can prevent unnecessarily weak results upon inference.

4.1 Context-specific signs

Before introducing context-specific signs, we formally define the notion of context for qualitative probabilistic networks.

Definition 4.1 *Let $G = (V(G), A(G))$ be an acyclic digraph. Let $X \subseteq V(G)$ be a set of variables in G called* context variables. *A* context c_X *for X is a combination of values for a subset $Y \subseteq X$ of the set of context variables. For $Y = \varnothing$ we say that the context is* empty, *denoted ϵ_X. For $Y = X$, we say that the context is* maximal. *The set of all possible contexts for X is called the* context set *for X and is denoted C_X.*

The subscript X for the empty context ϵ will be omitted as long as no confusion is possible. Note that contexts may pertain to arbitrary subsets of variables from a qualitative network. Also note that any combination of values x for a set $X \subseteq V(G)$ can be written as $\mathrm{c}_X x'$, where the context c_X is the combination of values for a set $Y \subseteq X$ and x' is the combination of values for $X \setminus Y$.

Upon inference, we will have to compare different contexts for the same set of context variables. For this purpose, we define a partial order '>' on contexts.

Definition 4.2 *Let $G = (V(G), A(G))$ be an acyclic digraph and let $X \subseteq V(G)$ be a set of context variables. Let c_X and c'_X be combinations of values for the sets $Y \subseteq X$ and $Y' \subseteq X$, respectively. Then, $\mathrm{c}_X > \mathrm{c}'_X$ iff $Y \supset Y'$ and c_X and c'_X specify the same combination of values for Y'.*

From this definition and the notational convention introduced above, we have that if $\mathrm{c}_X > \mathrm{c}'_X$ for two contexts c_X and c'_X, then c_X can be written as $\mathrm{c}'_X x$ for the combination of values x for the appropriate subset of X.

We now define a context-specific sign to be a sign that may vary from context to context. A context-specific sign can basically be looked upon as a function $\delta : C_X \rightarrow \{+, -, 0, ?\}$ from a set of contexts C_X to the set of basic signs introduced in Section 2.

Definition 4.3 *Let $Q = (G, \Delta)$ be a qualitative probabilistic network and let $X \subseteq V(G)$ be a set of context variables. A* context-specific sign *is a function $\delta : C_X \rightarrow \{+, -, 0, ?\}$ such that for any two contexts c_X and c'_X, $\mathrm{c}_X > \mathrm{c}'_X$, the following property holds:*

$$\delta(\mathrm{c}'_X) = \delta_i,\ \ \delta_i \in \{+, -, 0\} \implies$$

$$\delta(\mathrm{c}_X) = \begin{cases} \delta_i & \text{if, for } \mathrm{c}_X = \mathrm{c}'_X x,\ \ \delta(x) \in \{\delta_i, ?\} \\ 0 & \text{otherwise} \end{cases}$$

The definition of context-specific sign in essence states that the sign for a context agrees with the sign for any larger context, in the sense that signs cannot become less constrained for increasing contexts; we say that a '0' is more constrained than a '+' or a '−', which in turn are more constrained than a '?'. More specifically, signs cannot disagree unless they pertain to contexts that cannot occur simultaneously.

We will write $\delta(X)$ to denote the context-specific sign δ that is defined on the context set C_X. To avoid an abundance of braces, we will further write $\delta(A)$ instead of $\delta(\{A\})$ to indicate a context-specific sign for a single context variable A. Note that the basic signs from regular qualitative networks can be looked upon as context-specific signs that are defined by a constant function. By being context-*independent*, they in essence cover all possible contexts.

Having introduced the notion of context-specific sign, we now extend the basic formalism of qualitative networks by allowing context-specific signs for qualitative influences.

Definition 4.4 *Let $G = (V(G), A(G))$ be an acyclic digraph and let* Pr *be a joint probability distribution on $V(G)$ that respects the independences in G. Let A, B be variables in G with $A \rightarrow B \in A(G)$ and let $X = \pi_G(B) \setminus \{A\}$ be the set of predecessors of B other than A. Then, variable A exerts a* qualitative influence of sign $\delta(X)$ *on variable B, denoted $S^{\delta(X)}(A, B)$, iff for each context* c_X *for X we have*

- $\delta(\mathrm{c}_X) = +$ *iff* $\Pr(b \mid a\mathrm{c}_X x') \geq \Pr(b \mid \bar{a}\mathrm{c}_X x')$ *for all combinations of values* $\mathrm{c}_X x'$ *for X;*

- $\delta(\mathrm{c}_X) = -$ *iff* $\Pr(b \mid a\mathrm{c}_X x') \leq \Pr(b \mid \bar{a}\mathrm{c}_X x')$ *for all combinations of values* $\mathrm{c}_X x'$ *for* X*;*
- $\delta(\mathrm{c}_X) = 0$ *iff* $\Pr(b \mid a\mathrm{c}_X x') = \Pr(b \mid \bar{a}\mathrm{c}_X x')$ *for all combinations of values* $\mathrm{c}_X x'$ *for* X*;*
- $\delta(\mathrm{c}_X) = ?$ *otherwise.*

Note that in defining a context-specific influence for an arc between two variables A and B, we have taken the set X of predecessors of B other than A for the set of context variables. This restriction of the set of context variables is not essential, however, and can be lifted whenever desirable. Context-specific qualitative synergies are defined analogously.

A context-specific sign $\delta(X)$ in essence has to specify a basic sign from the set $\{+, -, 0, ?\}$ for each possible combination of values in the context set C_X. From the definition of context-specific sign, however, we have that such a sign cannot specify arbitrary basic signs for the various contexts as some of these signs are fully determined by the basic signs of other contexts. It is therefore not necessary to explicitly specify a basic sign for every context. The following example illustrates this observation.

Example 4.5 *We consider a qualitative influence of a variable A on a variable B with the set of context variables $X = \{D, E\}$. Suppose that the sign $\delta(X)$ of the influence is defined as*

$$\begin{array}{cccc} & \delta(\epsilon) = ?, & & \\ \delta(d) = +, & \delta(\bar{d}) = -, & \delta(e) = ?, & \delta(\bar{e}) = +, \\ \delta(de) = +, & \delta(d\bar{e}) = +, & \delta(\bar{d}e) = -, & \delta(\bar{d}\bar{e}) = 0 \end{array}$$

From the definition of context-specific sign, we have for example that $\delta(d) = +$ enforces $\delta(de)$ and $\delta(d\bar{e})$ to be either '$+$' or '0'. As both contexts de and $d\bar{e}$ induce the same sign basic as context d, the signs $\delta(de)$ and $\delta(d\bar{e})$ reveal that no additional information is hidden by the sign $\delta(d)$. Building upon this observation, the function $\delta(X)$ can be uniquely described by the signs of the smaller contexts whenever the larger contexts are assigned the same basic sign. The function is therefore fully described by the four signs

$$\delta(\epsilon) = ?,\ \delta(d) = +,\ \delta(\bar{d}) = -,\ \delta(\bar{e}) = +$$

The sign for the context $\delta(\bar{d}\bar{e})$, for example, can be easily derived from these signs. As $\delta(\bar{d}) = -$, we have from the definition of context-specific sign that

$\delta(\bar{d}\bar{e})$ can be either '$-$' or '0'. From $\delta(\bar{e}) = +$, we have in addition that $\delta(\bar{d}\bar{e})$ should be either '$+$' or '0'. We conclude that $\delta(\bar{d}\bar{e})$ equals zero. The sign for the context $\bar{d}e$ is derived in much the same way. The unspecified sign $\delta(e)$ equals that of the smaller empty context; the sign for the context e therefore does not pose any restrictions on the sign for $\bar{d}e$. The sign $\delta(\bar{d}) = -$, however, restricts the sign $\delta(\bar{d}e)$ to be either '$-$' or '0'. As no sign has been stated explicitly for the context $\bar{d}e$, it inherits its sign from that for $\bar{d}$: $\delta(\bar{d}e) = -$. □

In order to exploit the above observation, we have to provide for computing the unspecified sign of a larger context from the signs of smaller contexts. For contexts c_X that pertain to a single variable, the unspecified sign $\delta(\mathrm{c}_X)$ is taken to be equal to the sign specified for the empty context ϵ. For contexts c_X that pertain to a set Y of two or more variables, we rewrite c_X as $\mathrm{c}'_X c$, where c is the value assigned by c_X to some variable $C \in Y$ and c'_X assigns the same values to the variables $Y \setminus \{C\}$ as c_X. We then compute the unspecified sign $\delta(\mathrm{c}_X)$ recursively from $\delta(\mathrm{c}_X) = \delta(\mathrm{c}'_X) \oslash \delta(c)$, building upon the *and*-operator from Table 2. If $\delta(\mathrm{c}'_X) = \delta(c)$ then the sign of c_X obviously equals $\delta(c)$. If one of $\delta(\mathrm{c}'_X)$ or $\delta(c)$ equals zero, then $\delta(\mathrm{c}_X)$ should also be zero. If one of $\delta(\mathrm{c}'_X)$ or $\delta(c)$ is a '?', then the strongest of the two signs is taken for $\delta(\mathrm{c}_X)$. If $\delta(\mathrm{c}'_X) = +$ and $\delta(c) = -$, or vice versa, then $\delta(\mathrm{c}_X)$ can only be zero. The procedure for determining basic signs from a partially specified context-specific sign is summarised in pseudocode in Figure 6.

The standard sign-propagation algorithm for probabilistic inference with a qualitative network, as discussed in Section 2.2, is easily extended to handle context-specific signs. The extended algorithm propagates and combines *basic signs* only, as does the standard algorithm. Before a sign is propagated over an influence, however, the currently valid context is determined from the available observations and the basic sign that is either specified or computed for this context is propagated. If none of the context variables have been

$\oslash$	$+$	$-$	0	?
$+$	$+$	0	0	$+$
$-$	0	$-$	0	$-$
0	0	0	0	0
?	$+$	$-$	0	?

Table 2: The $\oslash$-operator for combining signs.

```
function ComputeSign(c_X, δ(X)): {+, −, 0, ?}

  if δ(c_X) is specified
  then return δ(c_X);
  if X is a singleton
  then return δ(ε_X);
  return ComputeSign(c'_X, δ(X)) ⊘ ComputeSign(c, δ(X))
         for c'_X c = c_X.
```

Figure 6: The procedure for computing basic signs from a partially specified context-specific sign.

observed, then the sign specified for the empty context is propagated. The extended sign-propagation algorithm is given in Figure 7. We note that the algorithm handles both context-specific and regular signs.

```
procedure PropagateSign(trail, from, to, messagesign):

  sign[to] ← sign[to] ⊕ messagesign;
  trail ← trail ∪{to};
  for each active neighbour V_i of to
  do linksign ← sign of (induced) influence between to and V_i;
     if linksign = δ(X)
     then determine the current context c_X from the observations;
          linksign ← ComputeSign(c_X, δ(X));
     messagesign ← sign[to] ⊗ linksign;
     if V_i ∉ trail and sign[V_i] ≠ sign[V_i] ⊕ messagesign
     then PropagateSign(trail, to, V_i, messagesign).
```

Figure 7: The extended sign-propagation procedure for handling context-specific signs.

4.2 Exploiting context-specific signs

In Section 3 we have presented two examples showing that the influences of a qualitative probabilistic network can hide context-specific information.

Revealing this hidden information and exploiting it upon inference can be worthwhile. The information that an influence is zero for a certain context can be used, for example, to improve the runtime complexity of the sign-propagation algorithm because propagation of a sign along a chain can be stopped as soon as a zero influence is encountered on that chain. More importantly, however, exploiting context-specific information can prevent conflicting influences arising during inference and can thereby forestall the generation of ambiguous signs. We illustrate this observation by means of an example.

Example 4.6 *We reconsider the qualitative* Surgery *network from Figure 4. Suppose that a non-smoker is undergoing surgery. From Example 3.1 we recall that, in the context of the observation $\bar{s}$ for the variable S, propagation of the observation t for the variable T with the standard sign-propagation algorithm results in the sign '?' for L, that is, the influence of the surgery on the patient's life expectancy is unknown. In essence, there is not enough information present in the network to compute a non-ambiguous sign from the two conflicting reasoning chains between T and L.*

From Example 3.1, we now further recall that the positive qualitative influence of T on P effectively hides a zero influence. With our new notion of context-specific sign, we can associate the sign $\delta(S)$ with

$$\delta(\epsilon) = +,\ \delta(\bar{s}) = 0$$

with the influence of T on P, thereby explicitly representing the information that non-smoking patients are not at risk for pulmonary complications after surgery. The extended network is shown in Figure 8.

We now reconsider our non-smoking patient undergoing surgery. Propagating the observation t for the variable T with the extended sign-propagation algorithm in the context of the observation $\bar{s}$ results in the sign $(+ \otimes +) \oplus (0 \otimes -) = +$ for the variable L. The previously hidden zero influence is exploited upon inference and we find that the surgery indeed is likely to increase the patient's life expectancy. □

In Section 3 we have not only discussed hidden zero influences, but have also argued that positive and negative influences can be hidden by the non-monotonic influences of a qualitative network. As the initial '?'s of these influences tend to spread to major parts of the network upon inference, it is worthwhile to resolve the non-monotonicities involved whenever possible.

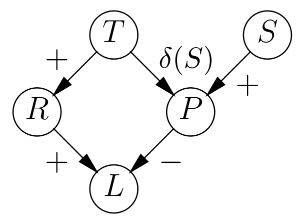

Figure 8: A hidden zero revealed by a context-specific sign.

Our extended formalism of qualitative networks provides for explicitly capturing information about non-monotonicities by context-specific signs. The following example illustrates the basic idea.

Example 4.7 *We reconsider the qualitative* Cervical metastases *network from Figure 5. From Example 3.2, we recall that the influence of the variable M, modelling the extent of lymphatic metastases, on the variable C, which represents the presence or absence of metastases in the lymph nodes in the neck, is non-monotonic. More specifically, we have that*

$$\Pr(c \mid ml) > \Pr(c \mid \bar{m}l) \text{ and } \Pr(c \mid m\bar{l}\,) < \Pr(c \mid \bar{m}\bar{l}\,).$$

In the context of an observation l, that is, for tumours located in the lower two-third of the oesophagus, we have that the influence is positive, while it is negative in the context $\bar{l}$, that is, for tumours higher up in the oesophagus. With our new notion of context-specific sign, we can make this hidden information explicit. In the extended network, shown in Figure 9, the information is captured by the sign $\delta(L)$ with

$$\delta(\epsilon) = ?,\ \ \delta(l) = +,\ \ \delta(\bar{l}\,) = -$$

for the influence of the variable M on C. It will be evident that the now explicitly represented information can be exploited upon inference much in the same way as described in Example 4.6. □

5 Evaluation of context-specificity in real-life networks

To get an impression of the context-specific information that is hidden in real-life qualitative probabilistic networks, we computed qualitative abstractions

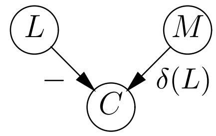

Figure 9: A non-monotonicity captured by a context-specific sign.

of the well-known ALARM-network [19] and of a probabilistic network for oesophageal cancer, called the OESOCA-network [20]. The ALARM-network is reproduced in Figure 10 for ease of reference. It consists of 37, mostly non-binary, variables and 46 arcs; the number of direct qualitative influences in the abstracted network thus equals 46. The OESOCA-network, shown in Figure 11, consists of 42, also mostly non-binary, variables and 59 arcs. In computing the qualitative abstractions of the two networks from the conditional probabilities specified, we have assumed that the values of a variable are ordered from top, the smallest value, to bottom, the largest value, as indicated in the figures. Table 3 summarises for the abstracted networks the numbers of direct influences for the four different basic signs.

The numbers reported in Table 3 pertain to the basic signs of the qualitative influences associated with the arcs in the digraphs of the networks. Each such influence, and hence each associated basic sign, covers a number of maximal contexts. For a qualitative influence associated with an arc $A \rightarrow B$, the number of maximal contexts equals 1 if variable B has no other predecessors than A; the only context then is the empty context. If B does have other predecessors then the number of maximal contexts equals the number of possible combinations of values for this set of predecessors. For the ALARM-network there thus are 218 maximal contexts; for the OESOCA-network, the number of maximal contexts equals 175. For every maximal context in the two networks, we have computed the true context-specific sign from its original quantified versions. Table 4 summarises the numbers of maximal contexts and their

	# *direct influences with sign* δ:				
	+	−	0	?	*total*:
ALARM	17	9	0	20	46
OESOCA	32	12	0	15	59

Table 3: The numbers of direct influences with regular '+', '−', '0' and '?' signs for the qualitative ALARM- and OESOCA- networks.

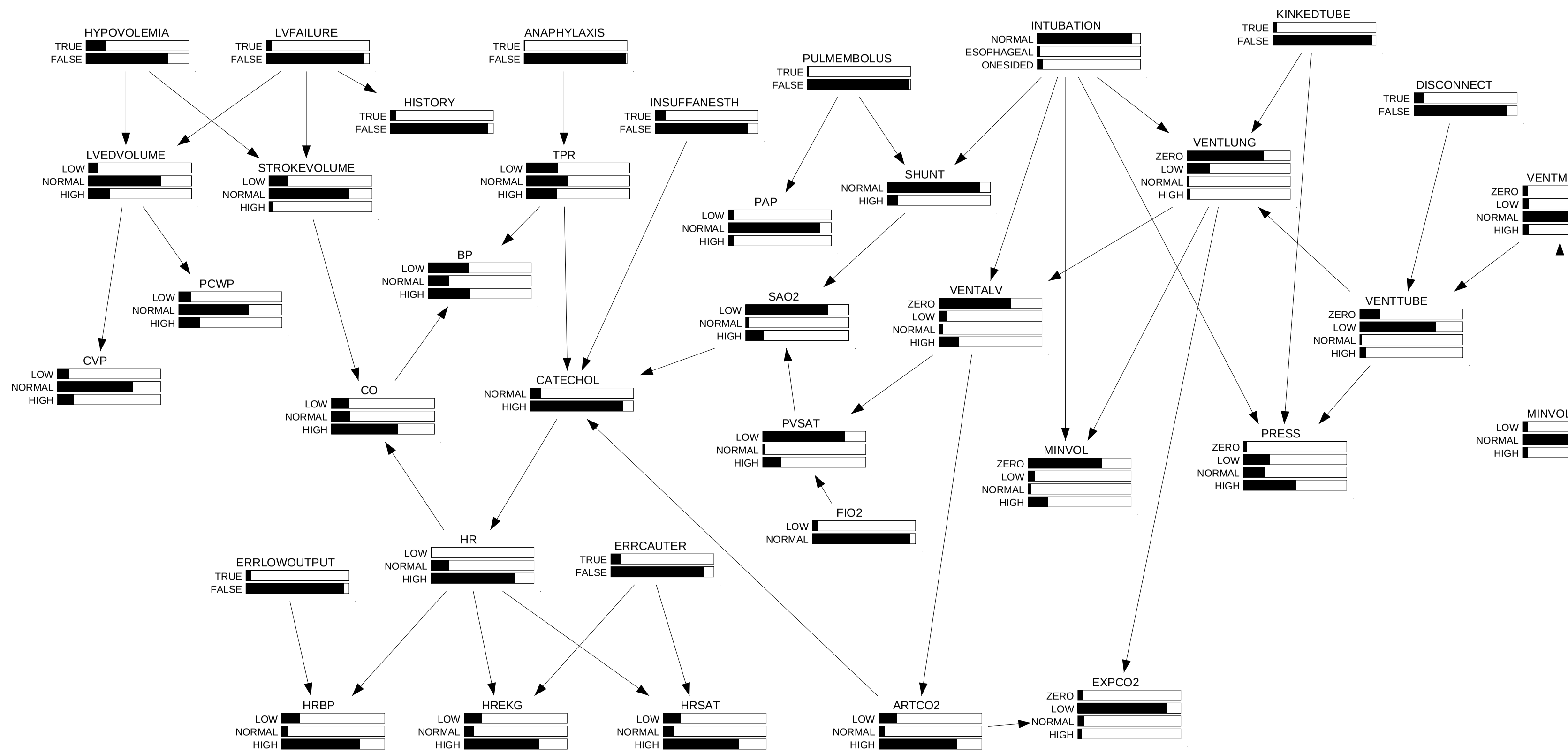

Figure 10: The ALARM-network (with its prior probabilities).

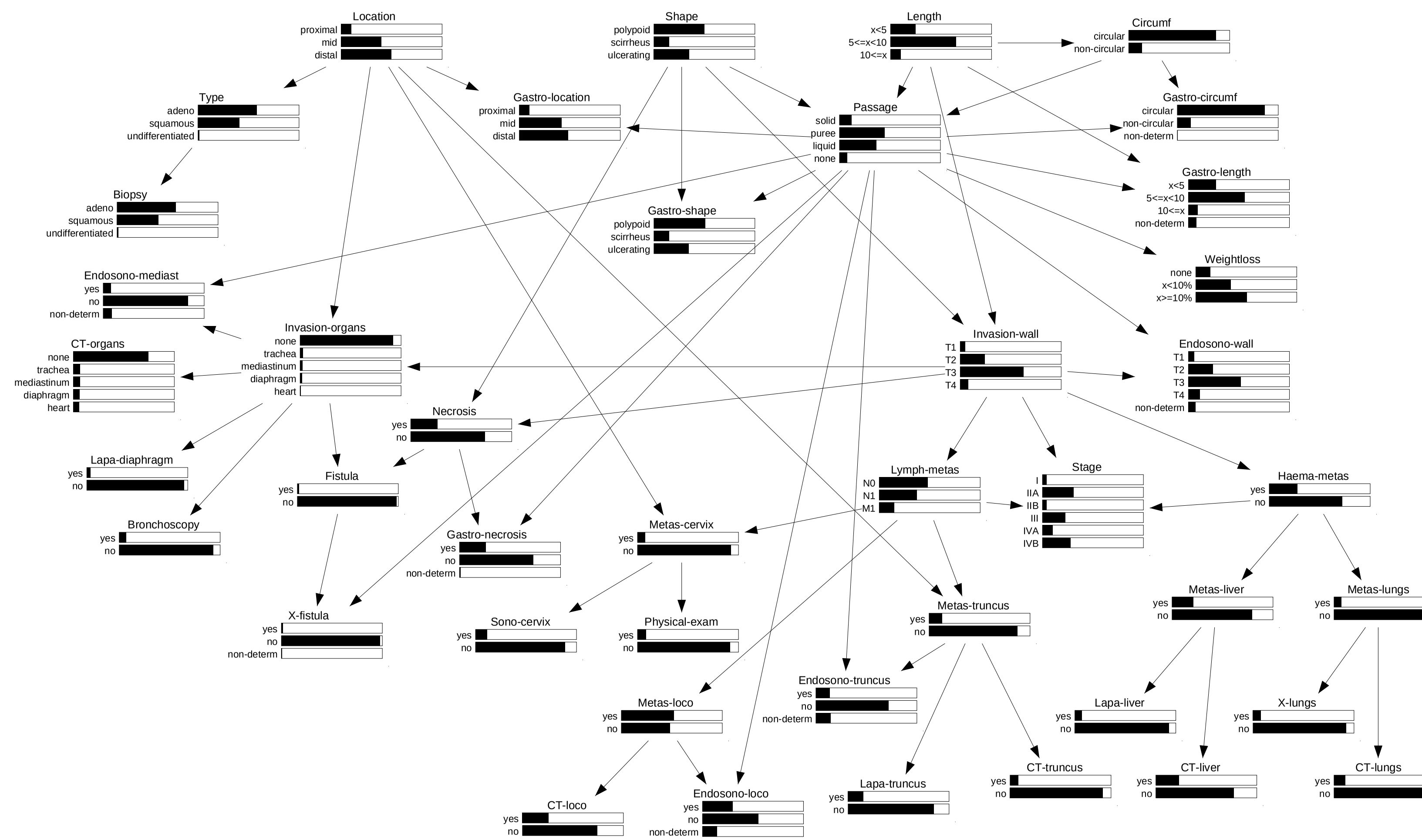

Figure 11: The OESOCA-network (with its prior probabilities).

ALARM		# max. c_X *with sign* δ':				*total*:
		+	−	0	?	
	+	38	–	21	–	59
δ:	−	–	40	11	–	51
	0	–	–	–	–	0
	?	34	24	12	28	108
total:		72	64	44	28	218

OESOCA		# max. c_X *with sign* δ':				*total*:
		+	−	0	?	
	+	74	–	8	–	82
δ:	−	–	36	8	–	44
	0	–	–	–	–	0
	?	6	3	2	38	49
total:		80	39	18	38	175

Table 4: The numbers of maximal contexts c_X covered by the regular '+', '−', '0' and '?' signs (δ) and their associated context-specific signs (δ'), for the qualitative ALARM- and OESOCA- networks.

associated signs, and the way they are covered by the different basic signs in the two abstracted networks. From the table we observe, for example, that the 17 positive qualitative influences from the qualitative ALARM network together cover 59 different maximal contexts. For 38 of these contexts, the influences are indeed positive, but for 21 contexts the positive influences actually hide a zero influence, that is, an independence.

For the qualitative ALARM-network, Table 3 shows that 35% of the direct influences are positive, 17% are negative, and 48% are ambiguous; the network does not include any explicitly specified zero influences. For the extended network, we observe from Table 4 that, in terms of the signs for the maximal contexts, 32% of the influences are positive. Note that 47% of these influences are in fact hidden in the qualitative ALARM-network. 31% of the influences in the extended network are negative, 20% are zero, and 17% remain ambiguous. Note that 65% of the ambiguous influences in the qualitative ALARM-network effectively hide a positive, negative or zero context-specific influence. For the qualitative OESOCA-network, Table 3 shows that 54% of the influences are positive, 21% are negative, and 25% are ambiguous; the network does not include any explicit zero influences. For the extended network, we find that, once again in terms of the signs for the maximal contexts, 46% of the qualitative influences are positive, 22% are negative, 10% are zero, and 22% remain ambiguous. Note that, although the qualitative OESOCA-network also hides context-specific information, its information hiding is less prominent than in the ALARM-network.

We conclude that for both the ALARM- and the OESOCA-network, the use of context-specific signs serves to reveal a considerable number of zero influences and to substantially decrease the number of ambiguous influences. Similar observations have been found for the qualitative abstractions of two other real-life probabilistic networks, pertaining to Wilson's disease [21] and to ventricular septal defect [22], respectively. We feel that by providing for the inclusion of context-specific information about influences, we have effectively extended the expressive power of qualitative probabilistic networks for real-life applications.

6 Conclusions

Qualitative networks model the probabilistic influences involved in an application domain at the high abstraction level of variables, as opposed to quantified probabilistic networks where influences are represented at the level of values of variables. Due to this high level of representation detail, knowledge about probabilistic influences that hold only for specific values of certain variables cannot be expressed in a qualitative network. We have shown that, as a consequence, the results computed from the network can be weaker than strictly necessary. We have argued that some of the knowledge that is hidden in a qualitative network is in fact qualitative in nature and should be represented explicitly to be exploited upon reasoning. To this end, we have extended the formalism of qualitative probabilistic networks with a notion of context-specificity. By doing so, we have provided for a finer level of representation detail and thereby enhanced the expressive power of qualitative networks. While in a regular qualitative network zero influences as well as positive and negative influences can be hidden, in an extended network context-specific signs are used to make these hidden influences explicit. We have shown that these signs can be specified in an efficient way. We have further shown that exploiting context-specific information can forestall unnecessary ambiguous signs during inference.

We have argued that qualitative probabilistic networks can play an important role in the construction of probabilistic networks for real-life application domains. By first obtaining a qualitative network from domain experts, the reasoning behaviour of the projected quantified network can be studied and validated prior to the assessment of the various probabilities required. The elicited signs can further be used as constraints on the probabilities to be assessed. Now, recall that the notion of context-specific independence was introduced before for quantified probabilistic networks as a concept to be exploited to speed up probabilistic inference. To identify the context-specific independences, generally the conditional probability distributions that have been specified for the network have to be inspected [15]. Using context-specific signs in qualitative networks during the construction of a probabilistic network, now brings the additional advantage of context-specific independence information being readily available.

Acknowledgements

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

References

[1] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann Publishers, Palo Alto, CA, 1988).

[2] M.P. Wellman, Fundamental concepts of qualitative probabilistic networks, *Artificial Intelligence* 44 (1990) 257 – 303.

[3] M.J. Druzdzel and M. Henrion, Efficient reasoning in qualitative probabilistic networks, in: *Proceedings of the Eleventh National Conference on Artificial Intelligence* (AAAI Press, Menlo Park, CA, 1993) 548 – 553.

[4] M.J. Druzdzel and L.C. van der Gaag, "Building probabilistic networks: where do the numbers come from?" — Guest editors' introduction, *IEEE Transactions on Knowledge and Data Engineering* 12 (2000) 481 – 486.

[5] A.L. Jensen, Quantification experience of a DSS for mildew management in winter wheat, in: M.J. Druzdzel, L.C. van der Gaag, M. Henrion, and F.V. Jensen, eds., *Working Notes of the IJCAI '95 Workshop on Building Probabilistic Networks: Where Do the Numbers Come From?* (AAAI Press, Menlo Park, CA, 1995) 23 – 31.

[6] M.J. Druzdzel and L.C. van der Gaag, Elicitation of probabilities for belief networks: combining qualitative and quantitative information, in: Ph. Besnard and S. Hanks, eds., *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers, San Francisco, CA, 1995) 141 – 148.

[7] C.-L. Liu and M.P. Wellman, Using qualitative relationships for bounding probability distributions, in: G.F. Cooper and S. Moral, eds., *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers, San Francisco, CA, 1998) 346 – 353.

[8] M.J. Druzdzel and M. Henrion, Intercausal reasoning with uninstantiated ancestor nodes, in: D. Heckerman and A. Mamdani, eds., *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers, San Francisco, CA, 1993) 317 – 325.

[9] M. Henrion and M.J. Druzdzel, Qualitative propagation and scenario-based approaches to explanation in probabilistic reasoning, in: P.P. Bonissone, M. Henrion, L.N. Kanal and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence 6* (Elsevier Science Publishers, North-Holland, 1991) 17 – 32.

[10] S. Parsons and E.H. Mamdani, Reasoning in networks with qualitative uncertainty, in: D. Heckerman and A. Mamdani, eds., *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers, San Francisco, CA, 1993) 435 – 442.

[11] S. Parsons, *Qualitative Methods for Reasoning under Uncertainty* (MIT Press, Cambridge, MA, 2001).

[12] S. Parsons, Refining reasoning in qualitative probabilistic networks, in: Ph. Besnard and S. Hanks, eds., *Proceedings of the Eleventh Conference*

on Uncertainty in Artificial Intelligence (Morgan Kaufmann Publishers, San Francisco, CA, 1995) 427 – 434.

[13] S. Renooij and L.C. van der Gaag, Enhancing QPNs for trade-off resolution, in: K.B. Laskey and H. Prade, eds., *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers, San Francisco, CA, 1999) 559 – 566.

[14] S. Renooij, L.C. van der Gaag, S.D. Parsons and S.D. Green, Pivotal pruning of trade-offs in QPNs, in: G. Boutilier and M. Goldszmidt, eds., *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers, San Francisco, CA, 2000) 515 – 522.

[15] C. Boutilier, N. Friedman, M. Goldszmidt and D. Koller, Context-specific independence in Bayesian networks, in: E. Horvitz and F.V. Jensen, eds., *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers, San Francisco, CA, 1996) 115 – 123.

[16] N.L. Zhang and D. Poole, On the role of context-specific independence in probabilistic inference, in: T. Dean, ed., *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence* (Morgan Kaufmann Publishers, San Francisco, CA, 1999) 1288 – 1293.

[17] G.F. Cooper, The computational complexity of probabilistic inference using belief networks, *Artificial Intelligence* 42 (1990) 393–405.

[18] S.L. Lauritzen and D.J. Spiegelhalter, Local computations with probabilities on graphical structures and their application to expert systems, *Journal of the Royal Statistical Society, Series B* 50 (1988) 157 – 224.

[19] I.A. Beinlich, H.J. Suermondt, R.M. Chavez and G.F. Cooper, The ALARM monitoring system: a case study with two probabilistic inference techniques for belief networks, in: J. Hunter, J. Cookson and J. Wyatt, eds., *Proceedings of the Second Conference on Artificial Intelligence in Medicine* (Springer-Verlag, Berlin, 1989) 247 – 256.

[20] L.C. van der Gaag, S. Renooij, C.L.M. Witteman, B.M.P. Aleman and B.G. Taal, Probabilities for a probabilistic network: A case-study in oesophageal cancer, *Artificial Intelligence in Medicine* (2002) to appear.

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

[22] V.M.H. Coupé, N.B. Peek, J. Ottenkamp and J.D.F. Habbema, Using sensitivity analysis for efficient quantification of a belief network, *Artificial Intelligence in Medicine* 17 (1999) 223 – 247.