Predictive Reasoning in Subjective Bayesian Networks

Magdalena Ivanovska* and Audun Jøsang*

Abstract

Subjective Bayesian networks extend Bayesian networks by substituting the conditional probability distributions with subjective opinions. In that way they enable explicit representation of the uncertainty in the probabilistic information encoded in the network. In this paper we focus on predictive reasoning in subjective Bayesian networks and propose an inference method that is based on the operations of deduction and multiplication of subjective opinions. We demonstrate modelling and inference with subjective Bayesian networks through an example.

1 Introduction

A Bayesian network (BN) is a compact representation of probabilistic information in the form of a directed acyclic graph and probability distributions associated with its nodes. Bayesian network reasoning algorithms provide a way to propagate probabilistic information through the graph, enabling predictive and diagnostic reasoning applicable in risk assessment and decision making. A serious limitation of the Bayesian network reasoning methods is that all the input probabilities must be assigned a precise value in order for the inference algorithms to work and the model to be analysed. This is problematic in many situations where probabilities can not be reliably estimated or are completely missing, while we still want to infer the most accurate conclusions possible. Many different approaches have been proposed for dealing with incomplete Bayesian networks and imprecise probabilistic information in general, like, for example, Bayesian logic (Andersen and Hooker, 1994), credal networks (Cozman, 2000), the probabilistic logics and networks discussed in (Haenni et al., 2011), the logics of likelihood in (Fagin et al., 1990), and imprecise probabilities (Walley, 1991). Conditional reasoning has been an important part of the mentioned theories and has also been analysed in the context of belief theory (Shafer, 1976), (Xu and Smets, 1994).

Subjective Bayesian networks represent second-order uncertainty in BNs, uncertainty about the probabilities, in the form of *subjective opinions*. A *subjective opinion* on a random variable is a composite representation that includes a specific belief assessment of the random variable done by an expert, based on a test, etc; and a base rate distribution obtained from a background statistics about the knowledge domain. (Ivanovska et al., 2015) provides an introduction to subjective networks and briefly discusses several possible reasoning approaches. In Bayesian networks, inference is based on *evidence* in the form of an observation of some of the variables' values. In subjective BNs, the evidence itself can in general have the form of a subjective opinion, which provides a

*Dept. of Informatics, University of Oslo, Norway
This paper was presented at the NIK-2016 conference; see http://www.nik.no/.

way of representing soft evidence like, for example, vague observations. In this paper we propose a method for predictive reasoning in subjective BNs which provides a way of propagating subjective evidence from variables to their descendants in the graph. The proposed method combines the deduction operation for subjective opinions described in (Ivanovska et al., 2016) and the operation for multiplication of opinions introduced in (Jøsang and McAnally, 2004).

In Section 2 we introduce subjective BNs providing the necessary preliminaries for Bayesian networks and subjective opinions, but assuming that the reader is familiar with the basics of probability theory. In Section 3 we review the operations of deduction and multiplication of subjective opinions and propose a method for predictive reasoning in subjective BNs. In Section 4 we provide an example to demonstrate the method. In Section 5 we summarize the results of the paper and discuss topics for future work.

2 Subjective Bayesian Networks

Bayesian Networks

A *Bayesian network* (Pearl, 1988) with n variables is a directed acyclic graph (DAG) with random variables $V = \{X_1, \ldots, X_n\}$ as nodes, and a set of conditional probability distributions $p(X_i|Pa(X_i))$ associated with each node X_i containing one probability distribution $p(X_i|pa(X_i))$ of X_i for every assignment of values $pa(X_i)$ to its parent nodes $Pa(X_i)$.

If the Markov property holds for the given DAG and the joint distribution p of the variables $X_1, \ldots, X_n$ (Every node is conditionally independent of its non-descendant nodes given its parent nodes in the graph, $I(X_i, ND(X_i)|Pa(X_i))$, then p is determined from the input information in the network as follows:

$$p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i|pa(X_i)) \ , \quad (1)$$

where $pa(X_i)$ is the assignment of the parents of X_i that corresponds to the tuple $(x_1, \ldots, x_n)$.

The general *belief update* problem in Bayesian networks is the following: Given evidence in the form of an *observation* of the value of a variable X, to find the probability distribution of another variable Y (X and Y can also be subsets of V). There are different *belief propagation* methods for solving this problem which make use of the conditional independencies embedded in the graph.

Of importance for our later discussion is the graphical criterion for conditional independence in Bayesian networks, called d-separation (see, for example, (Neapolitan, 2003)): For three disjoint sets of nodes X, Y, and Z in a DAG, we say that Z *d-separates* X from Y, if every path between a node from X and a node from Y is *blocked* by Z, meaning: 1) there is a node on the path that delivers an arrow and belongs to Z, or 2) there is a node on the path with converging arrows that is neither in Z nor has a descendant that is in Z. If X and Y are d-separated by Z, then they are independent given Z, $I(X,Y|Z)$.

Subjective Opinions

Let X be a random variable. A *subjective opinion* on X (Jøsang, 2008) is a tuple:

$$\omega_X = (b_X, u_X, a_X) \ , \quad (2)$$

where $b_X : \mathbb{X} \to [0,1]$ is a *belief mass distribution*, $u_X \in [0,1]$ is an *uncertainty mass*, and $a_X : \mathbb{X} \to [0,1]$ is a *base rate distribution* of X, satisfying the following additivity constraints:[1]

$$u_X + \sum_{x \in \mathbb{X}} b_X(x) = 1 , \tag{3}$$

$$\sum_{x \in \mathbb{X}} a_X(x) = 1 . \tag{4}$$

The beliefs and the uncertainty mass are a result of a specific analysis of the random variable by applying expert knowledge, experiments, or a personal judgement. $b_X(x)$ is the belief that X takes the value x expressed as a degree in $[0,1]$. It represents the amount of experimental or analytical evidence in favour of x. u_X is a single value, representing the degree of uncertainty about the belief analysis. It represents lack of evidence that can be due to lack of knowledge or expertise, or insufficient experimental analysis. The base rate a_X is a prior probability distribution of X that reflects domain knowledge relevant to the specific analysis, most usually relevant statistical information. For example, a doctor wants to determine whether a patient suffers from depression. Based on examinations and tests, she concludes that the collected evidence is 10% inconclusive, but is still two times more in support of the diagnosis that the patient suffers from depression than of the opposite one. As a result, the doctor assigns 0.6 belief mass to the patient suffering from depression and 0.3 belief mass to the opposite diagnosis, complemented by 0.1 uncertainty mass. The probability that a random person in the population suffers from depression is 5% and this fact determines the base rates in the doctor's subjective opinion about the condition of the patient.

A subjective opinion in which $u_X = 0$, i.e. an opinion without any uncertainty, is called a *dogmatic opinion*. Dogmatic opinions correspond to probability distributions. A dogmatic opinion for which $b_X(x) = 1$, for some $x \in \mathbb{X}$, is called an *absolute opinion*. Absolute opinions correspond to observations. In contrast, an opinion for which $u_X = 1$ (and consequently $b_X(x) = 0$, for every x) is called a *vacuous opinion*. For a given multinomial opinion ω_X we define its corresponding *projected probability distribution* $\mathrm{P}_X : \mathbb{X} \to [0,1]$ in the following way:

$$\mathrm{P}_X(x) = b_X(x) \; + \; a_X(x) \, u_X \; . \tag{5}$$

$\mathrm{P}_X(x)$ is an estimate for the probability of x which varies from the base rate value, in the case of complete ignorance ($u_X = 1$), to the actual probability in the case $u_X = 0$. In the correspondence between subjective opinions on a random variable and multinomial Dirichlet model (Walley, 1996) of its distribution given in (Jøsang and McAnally, 2004), the belief mass $b_X(x)$ is proportional to the number of observations of x, $n(x_i)$, while u_X is inversely proportional to the total number of observations N:

$$b_X(x_i) = \frac{n(x_i)}{N+s} \; , \; \; u_X = \frac{s}{N+s} \; , \tag{6}$$

where s is the Dirichlet strength. Then P_X corresponds to the mean distribution of the Dirichlet posterior, if a_X is the mean of the Dirichlet prior.

A *joint subjective opinion* on variables $X_1, \ldots, X_n$, $n \geq 2$, is a tuple:

$$\omega_{X_1 \ldots X_n} = (b_{X_1 \ldots X_n}, u_{X_1 \ldots X_n}, a_{X_1 \ldots X_n}) \; , \tag{7}$$

[1] This definition is for a *multinomial subjective opinion*. In general, we can define *hyper opinions*, where $b_X : \mathcal{R}(\mathbb{X}) = 2^{\mathbb{X}} \setminus \{\mathbb{X}, \emptyset\}$, and operate with them through their multinomial projections (see (Ivanovska et al., 2016)).

where $b_{X_1 \ldots X_n} : \mathbb{X}_1 \times \ldots \times \mathbb{X}_n \to [0,1]$ and $u_{X_1 \ldots X_n} \in [0,1]$ satisfy the additivity condition in Eq.(3) and $a_{X_1 \ldots X_n}$ is a joint probability distribution of the variables $X_1, \ldots, X_n$. Given two sets of random variables X and Y, a *conditional opinion* on Y given that X takes the value x is a subjective opinion on Y defined as a tuple:

$$\omega_{Y|x} = (b_{Y|x}, u_{Y|x}, a_{Y|x}) \;, \tag{8}$$

where $b_{Y|x} : \mathbb{Y} \to [0,1]$ and $u_{Y|x} \in [0,1]$ satisfy the condition in Eq.(3) and $a_{Y|x} : \mathbb{Y} \to [0,1]$ is a probability distribution of Y. We use the notation $\omega_{Y|X}$ for a set of conditional opinions on Y, one for each value of X, i.e.:

$$\omega_{Y|X} = \{\omega_{Y|x} \mid x \in \mathbb{X}\} \;. \tag{9}$$

Subjective Bayesian Networks

A *subjective Bayesian network (SBN)* is a generalization of a classical Bayesian network where each probability distribution $p(X|pa(X))$ is represented with an opinion about it, $\omega_{X|pa(X)}$ (Fig.1).

The inference problem of *opinion update* in an SBN can be formulated as follows: Given a subjective opinion on an evidence variable X to derive a subjective opinion on a target variable Y.[2] This means that, unlike in the case of BNs, we allow for *soft evidence*, evidence in the form of a general subjective opinion, not only an absolute one (observation). We denote the derived opinion by $\omega_{Y\|X}$. If X is a root node, then $\omega_{Y\|X}$ is the marginal opinion on Y derived from the network's input. In the case when X is not a root node, we assume that by providing an opinion ω_X, we ignore the opinion on X that would be derived from the network's input, i.e. we "delete" the arrows between X and its parent nodes. A general inference procedure for deriving $\omega_{Y\|X}$ would ideally be performed in the following steps:

1. The projected probability of the resulting opinion is determined from the projected probabilities of the given opinions using standard BNs reasoning methods.
2. The base rate of the resulting opinion is either given *a priori*, or determined from the given ones by Bayesian reasoning or other specified methods.
3. The uncertainty mass and the beliefs are determined by implementing the result of 1. and 2. in Eq.(5), and setting additional constrains that they should satisfy in the specific inference.

While 1. and 2. can be defined in a generic way, it is hard to think of general constrains to be imposed on $u_{Y\|X}$ and $b_{Y\|X}$ that would apply independently of the position of the evidence and target variable in the graph. That is the reason we apply a piece-wise reasoning strategy, which follows the direction of the arrows in the graph and applies one operation (having its own constraints on the resulting opinion) at a time.

3 Predictive Reasoning in SBNs

Predictive reasoning in Bayesian networks (Korb and Nicholson, 2010) is the reasoning along the direction of the arrows in the DAG in the sense that the target variable is a descendant of the evidence variables. Determining the marginal probability

[2]The discussion in the paper is limited to opinion update with one evidence and one target variable.

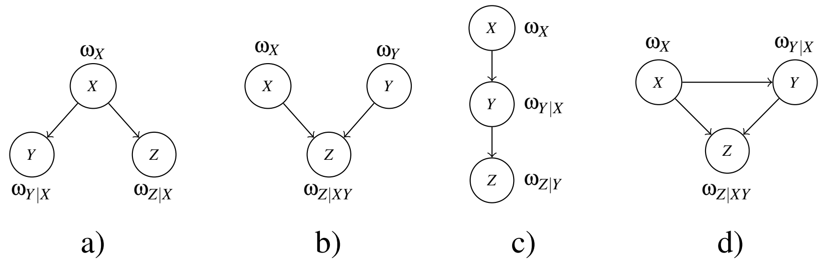


Figure 1: Three-node subjective Bayesian networks

distribution of a variable can be considered a special case of predictive reasoning where the set of evidence variables is empty. We present a method for predictive reasoning in SBNs that combines the subjective logic operations of deduction and multiplication. The deduction operation propagates information from parents to children, while multiplications operation derives the joint opinion from the individual opinions of two or more independent variables. We provide a short review of the deduction and multiplication operation. Then we propose an inference method for singly-connected and a class of multiply-connected networks. The proposed method is an extension of the discussion for predictive reasoning in singly-connected networks provided in (Ivanovska et al., 2016).

Subjective Logic Deduction

Given a set of opinions $\omega_{Y|X}$ and a subjective opinion ω_X, the goal of the operation of deduction is to derive a subjective opinion on Y, $\omega_{Y\|X}$. We denote this by the following expression:

$$\omega_{Y\|X} = \omega_{Y|X} \circledcirc \omega_X \ . \tag{10}$$

First the projected probability distribution of $\omega_{Y\|X}$ is determined as follows:[3]

$$\mathrm{P}(y\|X) = \sum_{x\in\mathbb{X}} \mathrm{P}(x)\mathrm{P}(y|x) \ . \tag{11}$$

The base rate a_Y is determined by an equation similar to Eq.(11) or supplied from another source (statistics). It remains to determine the uncertainty and the beliefs of the deduced opinion.

For the belief masses of the deduced opinion $\omega_{Y\|X}$, we assume the following:

$$b_{y\|X} \geq \min_{x\in\mathbb{X}}\{b_{y|x}\} \ , \tag{12}$$

for every $y \in Y$, which can be found as a *principle of plausible reasoning* in (Pearl, 1990).

Let $\omega_{Y\|\widehat{X}}$ be the deduced opinion from the vacuous opinion $\widehat{\omega}_X$ with a base rate a_X. Then $u_{Y\|\widehat{X}}$ is determined as the maximum possible uncertainty mass value under the conditions imposed by Eq.(12) and Eq.(11) applied to a vacuous opinion. The result is the following expression:

$$u_{Y\|\widehat{X}} = \min_{y\in\mathbb{Y}} \frac{\sum_x \mathrm{P}(y|x)a_x - \min_x b_{y|x}}{a_y} \ . \tag{13}$$

[3]Note that we use simplified notation for the projected probabilities, beliefs, and base rates, for example, $P(x)$ is an abbreviation for $P_X(x)$, $b_{y\|X}$ is an abbreviation for $b_{Y\|X}(y)$.

The uncertainty of the opinion $\omega_{Y\|X}$ deduced from an arbitrary ω_X is then determined as the weighted average of the uncertainty mass $u_{Y\|\widehat{X}}$ and the uncertainty masses of the given conditional opinions:

$$u_{Y\|X} = u_X u_{Y\|\widehat{X}} + \sum_{x\in\mathbb{X}} b_x u_{Y|x} \ . \tag{14}$$

Eq.(14) is the unique transformation that maps $\widehat{\omega}_X$ into $u_{Y\|\widehat{X}}$, and the corresponding absolute opinions on X into $u_{Y|x}$, for $x \in \mathbb{X}$. Once we have the uncertainty mass of the deduced opinion, the beliefs are easily derived as a consequence, applying Eq.(5).[4]

Deduction can be generalized for the case when Y has parents $X_1 \dots X_k$, where $k \geq 2$. Then the input arguments for the deduction operation are: 1) a joint opinion on the parents $\omega_{X_1 \dots X_k}$; and 2) a set of conditional opinions $\omega_{Y|X_1 \dots X_k}$ on Y, one for each combination of values of its parents. While the set of conditional opinions $\omega_{Y|X_1 \dots X_k}$ is a part of the input in a subjective Bayesian network, the joint opinion $\omega_{X_1 \dots X_k}$ would have to be derived. For this purpose we use the multiplication operation described in the next section.

Subjective Logic Multiplication

Given subjective opinions ω_X and ω_Y on two probabilistically independent variables X and Y, the multiplication operation derives an opinion ω_{XY} on their joint distribution. We will denote this by the following expression:

$$\omega_{XY} = \omega_X \cdot \omega_Y \ . \tag{15}$$

Since X and Y are probabilistically independent, the projected probability of the joint opinion satisfies the following:

$$\mathrm{P}(x,y) = \mathrm{P}(x)\mathrm{P}(y) \ . \tag{16}$$

a_{XY} is either obtained by a similar equation or provided separately. We will assume that $a_{XY} = a_X a_Y$ here. Applying Eq.(5) in Eq.(16), we obtain:

$$u_{XY} = \frac{b_x b_y + b_x u_Y a_y + u_X a_x b_y + u_X u_Y a_x a_y - b_{xy}}{a_x a_y} \ . \tag{17}$$

We impose the following requirement on the beliefs:

$$b_{xy} \geq b_x b_y \ . \tag{18}$$

For every pair of values x and y, the maximum value on the right-hand side of Eq.(17) is achieved for the smallest allowable value of b_{xy}, which is $b_{xy} = b_x b_y$. We denote that value with u_{xy}. Applying the latter in Eq.(17), we obtain:

$$u_{xy} = \frac{b_x}{a_x} u_Y + u_X \frac{b_y}{a_y} + u_X u_Y \ . \tag{19}$$

We take u_{XY} to be the minimum of these values, i.e. $u_{XY} = \min_{x,y}\{u_{xy}\}$, to assure that Eq.(18) always holds. This leads to the following expression:

$$u_{XY} = \min_{x,y}\{\frac{b_x}{a_x} u_Y + u_X \frac{b_y}{a_y} + u_X u_Y\} \ . \tag{20}$$

The beliefs b_{xy}, for $x \in \mathbb{X}$ and $y \in \mathbb{Y}$, then follow as a consequence of Eq.(16), Eq.(20), and Eq.(5).

[4]Note that the operation uses only unconditional base rates for Y. This is necessary for the condition in Eq.(12) to hold for the beliefs of $\omega_{Y\|X}$. Relaxing on this constraint, we can consider conditional base rates as well.

Reasoning in Three-Node Structures

In this section we analyse the cases of predictive reasoning that appear in the subjective networks given in Fig.1(a, b, c). The underlying graphs represent the three basic independence structures.

In the *common cause* network given in Fig.1(a), the only case of predictive reasoning is when we are given a subjective opinion on X and want to derive an opinion on Y or Z. Then only the deduction operation is used in deriving the opinions $\omega_{Y\|X}$ and $\omega_{Z\|X}$ from the opinion ω_X and the corresponding sets of conditionals given in the network. In this paper we are generally interested in inference with one evidence and one target variable, but it is worth mentioning that in Fig.1(a), we could easily combine the operations of deduction and multiplication to derive an opinion $\omega_{YZ\|X}$ given ω_X. First we obtain the set $\omega_{YZ|X}$ by a series of multiplications: $\omega_{YZ|x} = \omega_{Y|x} \cdot \omega_{Z|x}$, for every $x \in \mathbb{X}$, and then we obtain $\omega_{YZ\|X}$ applying deduction on $\omega_{YZ|X}$ and ω_X. Multiplication operation can be applied here because the independence relation $I(Y,Z|X)$ holds.

In the *common effect* network in Fig.1(b), the most natural case of predictive reasoning is when we are given opinions on the parents X and Y, and want to derive an opinion on the child Z. Since the variables X and Y in this network are probabilistically independent, we can apply the multiplication operation on the subjective opinions ω_X and ω_Y to obtain the opinion ω_{XY}. Then $\omega_{Z\|XY} = \omega_{Z|XY} \circledcirc \omega_{XY}$. The case that fits the general inference problem we are treating here is when we have a single piece of evidence, i.e. evidence on one of the parents, for example ω_X, and want to derive a subjective opinion on Z. We obtain $\omega_{Z\|X}$ in the same way as $\omega_{Z\|XY}$ here.

In the *chain network* in Fig.1(c), predictive reasoning is used when evidence ω_X is given, and opinion $\omega_{Y\|X}$ is derived by deduction. Applying deduction further on $\omega_{Y\|X}$ and $\omega_{Z|Y}$ would complete a predictive reasoning from X to Z deriving $\omega_{Z\|X}$. Alternatively, we could use two consecutive deductions on the absolute opinions on X, through $\omega_{Y|X}$ and $\omega_{Z|Y}$, to derive a set of opinions $\omega_{Z|X}$, and then use this set to deduce $\omega_{Z\|X}$ from ω_X. The first method is preferable though, since it is more direct, involving less operations. If evidence in the form of a subjective opinion ω_Y is available at Y, then applying deduction on ω_Y and $\omega_{Z|Y}$ we can derive $\omega_{Z\|Y}$. In the latter inference we ignore the input opinion ω_X that is "above" the evidence variable Y, and the opinion $\omega_{Y\|X}$ that can be deduced from it, since we have a new opinion (soft evidence) on the evidence variable Y.

Reasoning in Singly-Connected DAGs

A singly-connected DAG is a graph where there is only one path between any two nodes. Let X be the evidence variable and Y be the target variable in the inference, i.e. we are given a subjective opinion ω_X and want to derive a subjective opinion $\omega_{Y\|X}$. We distinguish between the following two cases:

1. The DAG is in the form of a *tree*. This means that every node has only one parent, so there is an ancestor chain between the evidence and the target: $X_1 \to \cdots \to X_n$ where $X_1 = X$ and $X_n = Y$. Then the reasoning from the evidence to the target is a generalization of the reasoning in the chain network in Fig.1(c), i.e. $\omega_{Y\|X}$ is obtained by $n-1$ consecutive deduction operations. If X is not a root node, then its ancestors in the graph are ignored in the inference.

2. The DAG contains V-structures, i.e. there are nodes that have multiple parents. Suppose Z is a node on the path between X and Y that has multiple parents. Then the parents of Z are probabilistically independent variables according to

the d-separation criterion since the only path between each two of them passes through Z, and Z is a node with converging arrows. This means that we can first derive subjective opinion on each of the parents of Z separately, and then use the multiplication operation to find $\omega_{Pa(Z)\|X}$, which we further propagate to Z. Because the graph is singly-connected, the parents of Z have sets of ancestors that are non-intersecting, hence the deduced opinions on them are derived independently.

Reasoning in Multiply-Connected DAGs

Let $X \to Y$ be an arrow in a given DAG. We call $X \to Y$ a *shortcut* if there is another path in the graph from X to Y containing at least one node Z other than X and Y. Clearly, all the singly-connected graphs are graphs without shortcuts. The simplest example of a graph with a shortcut is the three-node *connected* network given in Fig.1(d). Already in this network, the inference problem becomes complicated due to absence of the necessary independencies. In particular, given an opinion ω_X as evidence, we can propagate the evidence to Y by applying deduction, but we do not have a way of determining the opinion ω_{XY} in order to propagate the evidence to Z. Similar problems would appear in any DAG containing shortcuts.

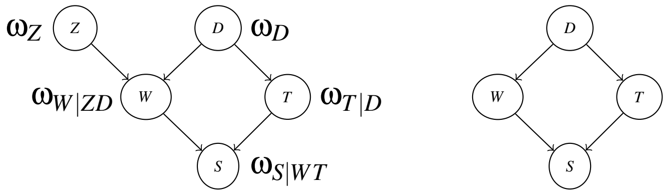
The simplest example of a multiply-connected DAG that does not contain shortcuts is the *directed diamond* structure given in the example in Fig.2(b). We can generalize the directed diamond to a *directed polygon*, which is a DAG in the shape of a polygon with two designated nodes, *start* S and *finish* F, and two directed paths of arbitrary length (greater than 1) from S to F.

If the evidence in the directed polygon is at a node E other than the start node S, then the incoming arrow in E is deleted, and the graph becomes singly-connected. If the evidence is a subjective opinion on the start node S, and the target is not the finish F, we perform chain reasoning. If the evidence is a subjective opinion on the start node S, and the target node is F, we take the following steps:

- $\omega_{P_1|s}$, $\omega_{P_2|s}$, for $P_1, P_2 \in Pa(F)$, for every $s \in \mathbb{S}$, are determined by chain reasoning.
- $\omega_{Pa(F)|s} = \omega_{P_1|s} \cdot \omega_{P_2|s}$, for every s, since $I(P_1, P_2|S)$.
- $\omega_{Pa(F)\|S} = \omega_{Pa(F)|S} \circledcirc \omega_S$
- $\omega_{F\|S} = \omega_{F|Pa(F)} \circledcirc \omega_{Pa(F)\|S}$

If the DAG is consisted of two directed polygons that are connected with the start or finish nodes, and do not have any other nodes in common, then the same reasoning will still work, since the connection points will d-separate any path from a node in one polygon to a node in the other. More precisely, we will have one of the following cases:

- The polygons are connected at the starts, i.e. $S_1 = S_2$. Then the predictive inference problems in them and their solutions are completely disjoint.
- The end of one of the polygons is the start of the other, i.e. $F_1 = S_2$. Then any inference problem with an evidence in the "above" polygon and target in the "below" one is performed by first determining the opinion $\omega_{F_1\|E}$ by reasoning in the first polygon, and then propagating it to the target in the second polygon.
- The polygons are connected with the final nodes, i.e. $F_1 = F_2$. Then the only interesting case is when the target is the common node. The solution in this case



a) b)

Figure 2: a) A graph of a problematic type, b) The graph of the scholarship example

	b_w	$b_{\bar{w}}$	u_W	a_w	p_w
$\omega_{W\|d}$	0.60	0.20	0.20	0.5	0.7
$\omega_{W\|\bar{d}}$	0.30	0.30	0.40	0.5	0.5

	b_t	$b_{\bar{t}}$	u_T	a_t	p_t
$\omega_{T\|d}$	0.20	0.10	0.70	0.20	0.34
$\omega_{T\|\bar{d}}$	0.05	0.85	0.10	0.20	0.07

Figure 3: Conditional opinions on W and T given the values of D

is provided by deducing opinions on the parents of the target node first (which will be a deduction from the new opinion on the evidence node, for the group of parents that are in the same polygon as the evidence node, or a deduction from the opinion on the root (start) node, for the group of parents that are not in the same polygon as the evidence node). Then, the deduced opinions on the two (independent) groups of parents are multiplied and propagated to the target by deduction.

The case of more than two polygons connected in the above described way is dealt with in a similar way. Similar methods apply in a graph structure obtained by connecting a singly-connected graph at one of the end nodes of a directed polygon. Fig.2(a) shows an example of a simple DAG that does not belong to the above discussed categories. The propagation from D to S in this network would require the set of conditional opinions $\omega_{W|D}$, in order to multiply them with the corresponding opinions from $\omega_{T|D}$ $(I(W,T|D))$ and propagate them further to the target, but we do not have a way to determine $\omega_{W|D}$ from the information provided in the graph.

4 Example

Consider the following situation depicted in Fig.2(b): A student can be granted a college scholarship (S) if she wins a race. Using doping (D) would increase the chances of her winning the race (W) at the same time increasing the chances of her testing positive on the doping test (T) after the race. Another student (we will call her "the analyst") competing for the same scholarship wants to predict the racers' chances of receiving the scholarship having an opinion on her use of doping.

We assume all the variables in the example are binary and denote the two states of a variable X by x and $\bar{x}$. The conditional opinions $\omega_{W|D}$ on the influence of doping on the results in the race are given in the left table of Fig.3. The beliefs and uncertainty in these opinions are subjective estimates of the analyst based on common sense and the current situation, while, in the absence of relevant statistics, the base rates are uniform. The analyst's opinions about the accuracy of the doping test are given in the right table of Fig.3 and are based on gathered opinions from experts in the laboratory. We can see that the test is very uncertain when there is doping, but still the chances are double for it being accurate rather than inaccurate. In the case of no doping, it will most certainly give

	b_s	$b_{\bar{s}}$	u_S	a_s	p_s
$\omega_{S\|wt}$	0.1	0.3	0.6	0.5	0.40
$\omega_{S\|w\bar{t}}$	0.8	0.1	0.1	0.5	0.85
$\omega_{S\|\bar{w}t}$	0.0	1.0	0.0	0.5	0.00
$\omega_{S\|\bar{w}\bar{t}}$	0.1	0.8	0.1	0.5	0.15

Figure 4: Conditional opinions on S given W and T

	b_{wt}	$b_{w\bar{t}}$	$b_{\bar{w}t}$	$b_{\bar{w}\bar{t}}$	u_{WT}
$\omega_{WT\|d}$	0.1935	0.2840	0.0575	0.0200	0.4450
$\omega_{WT\|\bar{d}}$	0.0150	0.3850	0.0150	0.3850	0.2000
a_{WT}	0.1000	0.4000	0.1000	0.4000	
$p_{WT\|d}$	0.2380	0.4620	0.1020	0.1980	
$p_{WT\|\bar{d}}$	0.0350	0.4650	0.0350	0.4650	

Figure 5: Conditional opinions on (W,T) given D

the correct results, with small chances of giving false positive result based on presence of similar substances. Statistically, the test has shown positive in 20% of the cases.

Granting the scholarship is not completely guaranteed by winning the race and testing negatively on the doping test, it is also based on availability and competition. Also, there is a little chance that the committee decides to grant the scholarship to the winner in case the test shows positive. On the other hand, it can happen that the racer in question does not win the race, but shows dedication and is clean on the test, and obtains the scholarship. The only deterministic case is when the racer does not win and tests positive on the doping test, which results in not obtaining the scholarship. The corresponding opinions are given in Fig.4. The base rate is non-informative.

Assume the analyst has a subjective opinion on the racer taking doping or not given by $\omega_D = (b_D, u_D, a_D)$, where $b_D = (0.40, 0.10)$, $u_D = 0.50$, and $a_d = a_{\bar{d}} = 0.50$ and resulting $p_d = 0.65$. Based on that, she wants to derive a subjective opinion on the racer winning the scholarship, the opinion $\omega_{S\|D}$. Following the procedure described in Section 3, we first derive the opinions in the set $\omega_{WT|D}$ by multiplication: $\omega_{WT|d} = \omega_{W|d} \cdot \omega_{T|d}$. The results are presented in Fig.5. Then we apply deduction to obtain $\omega_{WT\|D}$ from $\omega_{WT|D}$ and ω_D. At the end we apply one more deduction to obtain $\omega_{S\|D}$ from $\omega_{S|WT}$ and $\omega_{WT\|D}$. The results of these deductions are given in Fig.6.

The resulting opinion suggests belief that is slightly in favor of the racer not winning the scholarship, but we can clearly see that there is a significant amount of uncertainty in this opinion. We can imagine that operating with probabilities only, i.e. excluding the uncertainty factor, is comparable with operating with the projected probabilities here. In this example, it would give the same impression about not obtaining being more probable than obtaining the scholarship, but the uncertainty in this fact (that could be anything from

	b_{wt}	$b_{w\bar{t}}$	$b_{\bar{w}t}$	$b_{\bar{w}\bar{t}}$	u_{WT}
$\omega_{WT\|\|D}$	0.125	0.294	0.036	0.240	0.305
a_{WT}	0.100	0.400	0.100	0.400	
$p_{WT\|\|D}$	0.155	0.416	0.067	0.362	

	b_s	$b_{\bar{s}}$	u_S
$\omega_{S\|\|D}$	0.272	0.331	0.397
a_S	0.500	0.500	
$p_{S\|\|D}$	0.470	0.530	

Figure 6: Deduced opinions

0 to 1) would be ignored.

5 Conclusions and Future Work

In dealing with uncertain probabilistic information, we usually operate with probability estimates where the nature and the amount of evidence they are based on is rarely explicit. The advantage of reasoning with subjective opinions is that it deals with beliefs, uncertainty about them, and prior statistical information at the same time. In that way it enables control over more complex information, returning a more accurate portrait of the modelled situation. The projected probability of a subjective opinion is an estimate for the unknown probability distribution, in which the prior information acts as a backup for the beliefs as much as it is needed (depending on the uncertainty mass).

Subjective Bayesian networks are "uncertain" Bayesian networks whose local modes are subjective opinions instead of probability distributions. In view of the Dirichlet representaton of subjective opinions, SBNs can be seen as credal networks in a hierarchical form (Antonucci et al., 2014), where the hierarchical credal sets $(K(X), \pi)$ that correspond to the local conditional distributions are such that $K(X)$ consists of all the possible distributions of X, and π is a Dirichlet density function over $K(X)$ with parameters determined by the elements of the subjective opinions. Evidence in SBNs in general takes the form of subjective opinions on some of the network's variables, which enables reasoning over partial observations or soft evidence in SBNs. This is, in a sense, a generalization of Jeffrey's updating (Jeffrey, 1983) (where evidence takes the form of a new probability distribution over some of the variables) since subjective opinions are generalization of probability distributions.

We proposed a method for predictive reasoning in SBNs with singly-connected DAGs and multiply-connected DAGs of certain types. The proposed method combines the use of deduction and multiplication operation along the paths from the evidence to the target variable in a way that minimizes the number of applied operations in order to provide a better approximation of the ground truth. The probabilistic inference in BNs, and in particular inferring the marginal probabilities of the nodes, is proven to be NP-hard (Cooper, 1990). Since BNs can be considered as SBNs in which all the subjective opinions are dogmatic (all the uncertainty masses are zero), the problem of determining the marginal probabilities in BNs reduces to a predictive reasoning problem in SBNs and we can conclude that the latter is NP-hard as well.

In future work we want to extend the method to be able to do predictive reasoning in any type of DAGs. Also, evaluation of the procedure as well as a comparison with methods for opinion update based on the Dirichlet representation of subjective opinions (Kaplan and Ivanovska, 2016) will be a part of future work. We are currently working on developing general inference methods for opinion update in SBNs, which will apply to inference problems with any position of the evidence and target variable in the graph, and also to multiple evidence and target variables. In addition to predictive reasoning, this will open the possibilities for diagnostic and combined reasoning.

References

K. A. Andersen and J. N. Hooker. Bayesian logic. *Decis. Support Syst.*, 11(2):191–210, 1994.

A. Antonucci, A. Karlsson, and D. Sundgren. Decision Making with Hierarchical Credal

Sets. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU 2014*, pages 456–465, 2014.

G. Cooper. The Computational Complexity of Probabilistic Inference Using Bayesian Belief Networks. *Artificial Intelligence*, 42(2):393 – 405, 1990.

F. G. Cozman. Credal networks. *Artificial Intelligence*, 120(2):199–233, 2000.

R. Fagin, J. Y. Halpern, and N. Megiddo. A Logic for Reasoning about Probabilities. *Information and Computation*, 87:78–128, 1990.

R. Haenni, J.-W. Romeijn, G. Wheeler, and J. Andrews. Probabilistic Logic and Probabilistic Networks. *Synthese Library*, 350(8), 2011.

M. Ivanovska, A. Jøsang, L. Kaplan, and F. Sambo. Subjective Networks: Perspectives and Challenges. *Fourth International Workshop of Graph Structures for Knowledge Representation and Reasoning (GKR2015)*, pages 107–124, 2015.

M. Ivanovska, A. Jøsang, and F. Sambo. Bayesian Deduction with Subjective Opinions. In *15th International Conference on Principles of Knowledge Representation and Reasoning (KR 2016)*, pages 484–493. AAAI press, 2016.

R. Jeffrey. *The Logic of Decision*. McGraw-Hill, 2nd edition Univ. of Chicago press, 1983.

A. Jøsang. Conditional Reasoning with Subjective Logic. *Journal of Multiple-Valued Logic and Soft Computing*, 15(1):5–38, 2008.

A. Jøsang and D. McAnally. Multiplication and Comultiplication of Beliefs. *International Journal of Approximate Reasoning*, 38(1):19–51, 2004.

L. Kaplan and M. Ivanovska. Efficient Subjective Bayesian Network Belief Propagation for Trees. In *19th International Conference on Information Fusion (FUSION 2016), to appear*, 2016.

K. Korb and A. Nicholson. *Bayesian Artificial Intelligence, Second Edition*. CRC Press, 2010.

R. E. Neapolitan. *Learning Bayesian Networks*. Prentice Hall, 2003.

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

J. Pearl. Reasoning with Belief Functions: An Analysis of Compatibility. *International Journal of Approximate Reasoning*, 4(6):363–389, 1990.

G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, 1991.

P. Walley. Inferences from Multinomial Data: Learning about a Bag of Marbles. *Journal of the Royal Statistical Society*, 58(1):3–57, 1996.

H. Xu and P. Smets. Evidential Reasoning with Conditional Belief Functions. In D. Heckerman et al., editors, *Proceedings of Uncertainty in Artificial Intelligence (UAI94)*, pages 598–606. Morgan Kaufmann, San Mateo, California, 1994.