18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Towards Subjective Networks: Extending Conditional Reasoning in Subjective Logic

Lance Kaplan[1], Magdalena Ivanovska[2], Audun Jøsang[2] and Francesco Sambo[3]
[1]US Army Research Lab, Adelphi, MD, US
[2]University of Oslo, Oslo, Norway
[3]University of Padova, Padova, Italy

***Abstract*—This work expands the concept of conditional reasoning in subjective logic between two variables to reasoning over multiple variables as a step to eventually realize efficient means for inference over *subjective networks*, i.e. Bayesian networks with uncertain marginal and conditional probabilities. Three new inference methods are introduced for a three node subjective network. Monte Carlo simulations characterize the accuracy of these methods to represent inferred uncertain knowledge about the ground truth.**

I. Introduction

In many cases, military decisions must be made in highly complex and uncertain environment with limited experience to understand the interactions between all the variables being considered. Unfortunately, the opportunities afforded by Big Data and Deep Learning are limited in these cases because of the lack of relevant prior data to model the situation at hand. While probabilistic graphical models such as Bayesian networks have been successful decision making aids for many applications, their utility relies on large datasets or the extensive experience of subject matter experts to elicit the input conditional probabilities in the network [1].

Very often, small amount of evidence available in the process of elicitation, or lack of expertise, leads to high uncertainty in the input probabilities. For instance, when flipping a coin reveals that heads appears half of the time, one might declare that the coin is fair, i.e., $p = 0.5$. However, the uncertainty behind this assertion is high when the number of coin flips is small. Likewise, in a military scenario where past history of village elder indicates that he will provide support for a given mission with probability $p = 0.7$, the uncertainty of this assertion is high when the past history only provides a small amount of evidence to determine this probability. This type of uncertainty about probabilities is typically called *second order uncertainty* [2].

Many extensions of probability theory exist to accommodate higher orders of uncertainty like, for example, the theory of imprecise probabilities [3] and Dempster-Shafer Theory of evidence (DST) [4]. Subjective logic combines uncertainty with probabilistic reasoning by building upon the basic belief assignments in DST and establishing a connection between the belief representations and second-order probability representations in the form of Dirichlet distributions [5], [6]. In the sense of a Bayesian network, a subject matter expert (SME) can provide his/her subjective opinion about conditional probabilities that includes beliefs complemented with degree of uncertainty. Subjective logic provides means to map that opinion into a Dirichlet distribution for the possible values of the conditional probabilities. This means that a subjective opinion of a SME can be seen as representing an amount of evidence observed from past experience, e.g., results from a set of coin flipping or dice rolling experiments. Alternatively to using SMEs, subjective logic can also represent the actual posterior distribution for the conditional probabilities using machine learning techniques over historical data.

Standard Bayesian networks do not provide the means to represent second-order uncertainty. Extensions of Bayesian networks for evidential reasoning do exist, e.g., valuation-based system [7]. These extensions build upon alternatives for Bayes' rule in various evidential theories, e.g., DST, possibility theory, etc. Subjective networks distinguish themselves from the other methods by implementing Bayes' rule itself, but over second-order probability distributions.

Subjective logic comprises a number of ever growing operations for reasoning under uncertainty about probabilities [5]. Many of the operations use random variables sampled from Dirichlet distributions, and the opinion for the output of the operation approximates the Dirichlet distribution to the actual distribution of the output random variable. Subjective logic tries to maintain the mean of the actual distribution while maximizing the spread (or uncertainty) under natural constraints for the given operator. The conditional reasoning operations for deductive and abductive reasoning were introduced in [8], [9], but the uncertainty in abduction was known to be too high. Recently, the abduction operation was revised in [10] to reduce the uncertainty. Alternatively, it is possible to define subjective logic operations to better match the spread of the actual distributions of the output random variables, and [11] provides a revised deductive operation via moment matching.

The various conditional reasoning methods introduced in subjective logic so far only consider inferring a subjective opinion on one variable given an opinion on another under availability of the corresponding conditional subjective opin-

Research was sponsored by the U.S. Army Research Laboratory and was accomplished under Agreement Number W911NF-14-1-0608. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

ions. This amounts to doing subjective logic representation and inference in a two-node Bayesian network. This paper considers conditional inference in subjective logic in a special case of a three-node Bayesian network with a graph that is a chain. To this end, three inference methods are proposed of varying computational complexity and accuracy. Monte Carlo simulations evaluate the ability of these methods to properly represent and deal with the actual uncertainty about the ground truth.

It is anticipated that the utility of subjective networks can be determined using the tools and case studies being developed by the Evaluation of Techniques for Uncertainty Representation Working Group (ETURWG). This working group is developing the uncertainty representation and reasoning evaluation framework (URREF) ontology [12]. The ontology can be adapted to evaluate Bayesian networks [13], and we expect to evaluate subjective network using the same framework.

The remainder of the paper is structured as follows. Section II provides a brief overview of subjective logic, and Section III introduces the concept of subjective networks. The three inference methods are derived in Section IV, and Section V evaluates these methods via simulations. Finally, Section VI provides concluding remarks.

II. Subjective Logic

Let us assume a proposition X such as *the village elder will help you during a mission* is true or false for a given mission with probability p_x and $1 - p_x$, respectively. Observations, e.g., whether or not the elder helped during past missions can not always give us the exact probabilities, but can over time lead to a *subjective opinion* about them, expressed in terms of *beliefs* and *uncertainty*. In general, subjective logic [6] defines subjective opinions on random variables with n values ($n \geq 2$). A proposition such as the above can be seen as a binary variable that takes the values *true* or *false*. In this paper, we consider only binary variables X, with values x and $\bar{x}$, and we assume there is a probability distribution P over them, $\mathrm{P}(x) = p_x$, $\mathrm{P}(\bar{x}) = p_{\bar{x}} = 1 - p_x$, representing the ground truth.

A subjective opinion on a binary variable (or a proposition) X, called a *binomial opinion*, represents an individual's knowledge about the value p_x (and subsequently $p_{\bar{x}}$) based upon a number of observations and/or relevant knowledge and expertise. Formally, a binomial opinion on X is a tuple:

$$\omega_X = (b_x, b_{\bar{x}}, u_X), \tag{1}$$

where b_x is the *belief* that X takes the value x (or that the proposition X is true), $b_{\bar{x}}$ is the belief in $\bar{x}$ (the proposition X is false), and u_X is the *uncertainty* about the value of the variable X. These parameters satisfy $b_x, b_{\bar{x}}, u_X \in [0, 1]$ and $b_x + b_{\bar{x}} + u_X = 1$. In addition, we assume there is a *base rate distribution* a_X associated with each opinion, which is an *a priori* probability distribution of X in the absence of evidence. a_X represents domain knowledge and can be obtained from background statistics. It is typically taken to be a uniform distribution when no specific knowledge about the domain of interest is available.

There is a correspondence between subjective opinions and beta probability density functions (pdf) [14] that we exploit in this paper. Namely, a given subjective opinion $\omega_X = (b_x, b_{\bar{x}}, u_X)$ determines a beta pdf with parameters $\alpha = e_x + W a_x$ and $\beta = e_{\bar{x}} + W(1 - a_x)$, where e_x and $e_{\bar{x}}$ determined as:

$$e_x = \frac{W b_x}{u_X} \text{ and } e_{\bar{x}} = \frac{W b_{\bar{x}}}{u_X}, \tag{2}$$

are the number of instantiations of X with x and $\bar{x}$, respectively, in a series of observations, and W is the prior strength that controls how quickly evidence overcomes the base rate distribution a_X. Note that $u_X = 0$ equates to an infinite collection of instantiations of X. Inversely, given observation results $(e_x, e_{\bar{x}})$, a transformation of (2) determines the following subjective opinion on X:

$$\omega_X = \left(\frac{e_x}{e_x + e_{\bar{x}} + W}, \frac{e_{\bar{x}}}{e_x + e_{\bar{x}} + W}, \frac{W}{e_x + e_{\bar{x}} + W} \right). \tag{3}$$

This means that expressing a subjective opinion $\omega_X = (b_x, b_{\bar{x}}, u_X)$ is equivalent to expressing the subject's current knowledge about p_x (and subsequently $p_{\bar{x}}$) by the probability density function:

$$f(p_x|\omega_X) = \frac{(p_x)^{\alpha-1} (1 - p_x)^{\beta-1}}{B(\alpha, \beta)}, \tag{4}$$

were $B(\alpha, \beta)$ is the beta function, and α, and β are defined as above. It is easy to show that the mean value of p_x is then

$$m_x = \frac{e_x + W a_x}{e_x + e_{\bar{x}} + W} = b_x + a_x u_X. \tag{5}$$

The mean value associated to the subjective opinion represents an estimate for the ground truth value p_x and is referred to as the *expected probability*. Likewise, the uncertainty u_X is proportional to the spread of the beta distribution, which can be used to determine a confidence interval around the expected probability.

The sum of positive and negative evidence is akin to the number of coin flips to estimate the probability of heads. It is well known that for N_f coin flips with probabilities of the outcomes p_x and $1 - p_x$, the Cramer-Rao lower bound (CRLB) for each marginal is given by:

$$\text{CRLB} = \frac{p_x(1 - p_x)}{N_f}. \tag{6}$$

Given that the estimate m_x converges to the maximum *a posteriori* estimate that is asymptotically Gaussian, unbiased and efficient, we can approximate the variance of the estimate via

$$\text{VAR} \approx \frac{m_x(1 - m_x)}{s_x}, \tag{7}$$

where the Dirichlet strength $s_x = e_x + e_{\bar{x}} + W$ (see (3)) is the equivalent number of coin flips when including the prior as reflected in the expected probability m_x. Note that the variance for p_x is the same as for $p_{\bar{x}}$.

Subjective logic introduces many operations for operating with subjective opinions. This paper specifically needs to build upon *subjective logic abduction*, which takes as input an opinion on Y, ω_Y, and conditional opinions on Y given particular values of X, $\omega_{Y|x}$ and $\omega_{Y|\bar{x}}$, and produces opinion on X, ω_X, as output [9]. In particular, we make use of the process of *inverting* the given conditional opinions into opinions $\omega_{X|y}$, $\omega_{X|\bar{y}}$. This process is based upon an instance of Bayes' rule:

$$p_{x|y} = \frac{p_{y|x}p_x}{p_{y|x}p_x + p_{y|\bar{x}}(1-p_x)}, \tag{8}$$

where (8) is applied on the corresponding expected probabilities when the opinion on X is *vacuous*, i.e., $\omega_X = (0, 0, 1)$ and, consequently, $m_x = a_x$. Hence, the inversion process approximates the expected probability $m_{x|y}$ as follows:

$$m_{x|y} \approx \frac{m_{y|x}a_x}{m_{y|x}a_x + m_{y|\bar{x}}(1-a_x)}. \tag{9}$$

Likewise, the application of Bayes' rule to the corresponding expected probabilities can determine $m_{x|\bar{y}}$. Originally, subjective logic abduction sets the uncertainty of $\omega_{X|y}$ to the maximum possible value under the constraint that the beliefs are non-negative, which is the following:

$$u_{X|y} = \min\left\{\frac{m_{y|x}}{a_x}, \frac{1-m_{y|x}}{1-a_x}\right\}. \tag{10}$$

However, this inversion process sets the uncertainty too high, and the abduction process has recently been modified to scale this uncertainty [10]. Specifically, the modified process computes a *relevance factor*, e.g. the relevance of X to y is:

$$\Psi_{y|X} = \max\{m_{y|x}, m_{y|\bar{x}}\} - \min\{m_{y|x}, m_{y|\bar{x}}\}, \tag{11}$$

and a weighted uncertainty:

$$u^w_{X|y} = u_{Y|x}a_x + u_{Y|\bar{x}}(1-a_x), \tag{12}$$

to set the uncertainty scaling factor as:

$$\tilde{u}_{X|y} = (1-\Psi_{y|X}) + u^w_{X|y} - (1-\Psi_{y|X})u^w_{X|y}. \tag{13}$$

Finally, the new uncertainty is the scaled version of the original uncertainty:

$$u_{X|y} \leftarrow u_{X|y}\tilde{u}_{X|y}. \tag{14}$$

Given the expected probability distribution determined by $m_{x|y}$ ($m_{\bar{x}|y} = 1 - m_{x|y}$), $u_{X|y}$ and the base rate distribution a_X, the opinion $\omega_{X|y}$ is fully determined. Similarly, one can determine the opinion $\omega_{X|\bar{y}}$. The merits of the uncertainty scaling are discussed in [10].

For the remainder of the paper, $W = 2$ and $a_x = 0.5$, which equates to a uniform prior for p_x that is uninformative.

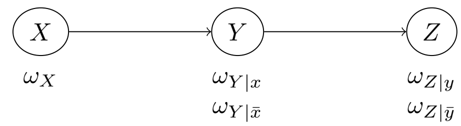


Fig. 1. A three node chain subjective network.

III. Subjective Networks

A Bayesian network [15] is a compact representation of a joint probability distribution of variables in a form of a directed acyclic graph, and conditional probability distributions associated with its nodes. In standard Bayesian reasoning, one is given some *evidence* in the form of values for a subset of the variables, known as *observed variables*. Bayesian inference is then used to compute the probability distribution of the unobserved variables, given the evidence on the observed variables, and the input conditional probabilities in the network.

The subjective logic equivalent to Bayesian network, *subjective network*, substitutes the given conditional probability distributions in the network with conditional opinions. In this framework, the equivalent to standard Bayesian evidence, *subjective evidence*, is a set of *subjective opinions* on the observed variables (which in this case can be only partially observed as well). Such opinions, together with the available conditional opinions, can be used to infer opinions on the unobserved variables.

Subjective evidence, in its more general form, can be any type of subjective opinion. Two subjective logic operations have been recently introduced to propagate evidence in a two-node network: *deduction* [9] which follows the direction of the available conditionals, i.e. given evidence in the form ω_X, and available conditional opinions $\omega_{Y|x}$ and $\omega_{Y|\bar{x}}$, infers ω_Y; and *abduction* [10], [11] that follows the opposite direction, as explained in Section II.

In this framework, the standard Bayesian evidence is just a special case of subjective evidence, where opinions on the variables are *absolute opinions*, *i.e.* opinions with belief equal to 1 for one specific value of a variable and no uncertainty.

The goal of this paper is to extend conditional reasoning over a subjective network with more than two nodes, given standard Bayesian evidence on a subset of the nodes. We focus on a three node subjective network as illustrated in Figure 1. In this network, the variable X is the root node, and the network is parametrized by opinions about the marginal probability distribution $\mathrm{P}(X)$ and the conditional probability distributions $\mathrm{P}(Y|X)$ and $\mathrm{P}(Z|Y)$. According to the Markov property[1], the joint probability distribution of the variables in this network is given by:

$$\mathrm{P}(X, Y, Z) = \mathrm{P}(Z|Y)\mathrm{P}(Y|X)\mathrm{P}(X). \tag{15}$$

For a subjective network of the form given in Figure 1, this paper develops three inference methods for deriving an opinion

[1]Every node is conditionally independent of its non-descendants given its parents in the graph.

on the variable Y, given that the variables X and Z take particular values, and the conditional opinions corresponding to the graph, namely $\omega_{Y|x}$, $\omega_{Y|\bar{x}}$, $\omega_{Z|y}$, and $\omega_{Z|\bar{y}}$, are available. In other words, the inference determines opinion on the posterior distribution $\mathrm{P}(Y|XZ)$, which for a particular value of Y is given as:

$$\mathrm{P}(y|Z,X) = \frac{\mathrm{P}(Z|y)\mathrm{P}(y|X)}{\mathrm{P}(Z|y)\mathrm{P}(y|X) + \mathrm{P}(Z|\bar{y})\mathrm{P}(\bar{y}|X)}. \tag{16}$$

The input conditional probabilities for a Bayesian network can either be provided by SMEs or estimated from a large number of joint instances of the variables. Likewise, the conditional opinions in the subjective network are either provided by SMEs or obtained from training samples.

For the three node network, the conditional opinions can be obtained by observing N instances of the variables: $[X_i, Y_i, Z_i]$, for $i = 1, \ldots, N$. For instance, the opinion:

$$\omega_{Y|x} = \frac{(\ n_{yx},\quad n_{\bar{y}x},\quad 2\ )}{n_{yx} + n_{\bar{y}x} + 2} \text{ and } \omega_{Y|\bar{x}} = \frac{(\ n_{y\bar{x}},\quad n_{\bar{y}\bar{x}},\quad 2\ )}{n_{y\bar{x}} + n_{\bar{y}\bar{x}} + 2} \tag{17}$$

where n_{xy} is the number of instances where $X_i = x$ and $Y_i = y$ out of the N observations. The other three conditional opinions are similarly derived. Note that in the training process for determining the four conditional opinions, the opinions are statistically independent. Furthermore, the conditionals for any edge can be learned independently so that for example in this three node network, one could consider the joint observations of $[X_i, Y_i]$ and $[Y_i, Z_i]$ separately.

IV. Inference

The goal of inference is to determine the opinions about the probability distributions of the unobserved variables conditional on given values for the observed variables. This section presents three approaches to determine opinions for $\mathrm{P}(Y|XZ)$ as given by (16) because the current library of subjective logic operations cannot accommodate this situation.

In what follows we assume we observe the values x for X and z of Z, and based on that and the input opinions in the network, we infer an opinion on Y. We will denote the inferred opinion by $\omega_{Y\|xz}$. The same procedures will apply to all other choices of values for X and Z, i.e. the opinions $\omega_{Y\|\bar{x}z}$, $\omega_{Y\|x\bar{z}}$, and $\omega_{Y\|\bar{x}\bar{z}}$ can be inferred in a similar way.

For fixed values x of X and z of Z, (16) is a function of three independent random variables p_1, p_2, and p_3:

$$\mathrm{P}(y|xz) = \frac{p_1 p_2}{p_1 p_2 + (1-p_1)p_3}, \tag{18}$$

where the distributions for p_1, p_2 and p_3 are are dictated by the opinions $\omega_{Y|x}$, $\omega_{Z|y}$ and $\omega_{Z|\bar{y}}$ respectively.

Because the opinions about the conditionals p_1, p_2 and p_3 are independent, their joint distribution is given by:

$$f(p_1, p_2, p_3) = f(p_1|\omega_1) f(p_2|\omega_2) f(p_3|\omega_3), \tag{19}$$

where $\omega_1 = \omega_{Y|x}$, $\omega_2 = \omega_{Z|y}$, and $\omega_3 = \omega_{Z|\bar{y}}$. By performing a change of variable using (18) for p_1 and integrating out p_2 and p_3, it is possible to derive the inferred distribution for $p_{y|xz}$, which is in general not a beta distribution. The various methods to infer an opinion about Y conditioned on the observed values x and z are actually approximating this inferred distribution by a beta distribution.

A. Method of Moments

The method of moments (MoM) numerically computes the first and second order moments for the inferred distribution as two triple integrals where the expected probability is given by:

$$m_{y|xz} = \int \frac{p_1 p_2}{p_1 p_2 + (1-p_1)p_3} f(p_1|\omega_1) f(p_2|\omega_2) f(p_3|\omega_3) dp_1 dp_2 dp_3, \tag{20}$$

and the non-central second order moment is computed by:

$$v_{y|xz} = \int \left(\frac{p_1 p_2}{p_1 p_2 + (1-p_1)p_3}\right)^2 f(p_1|\omega_1) f(p_2|\omega_2) f(p_3|\omega_3) dp_1 dp_2 dp_3. \tag{21}$$

As shown in [16], the necessary Dirichlet strength for the beta distribution to match the second order moment of the inferred distribution is:

$$s^*_{y|xz} = \frac{m_{y|xz} - v_{y|xz}}{v_{y|xz} - m^2_{y|xz}}. \tag{22}$$

To ensure that the inferred beliefs are non-negative, the Dirichlet strength might need to be increased so the MoM sets:

$$s_{y|xz} = \max\left\{ s^*_{y|xz}, \frac{1}{m_{y|xz}}, \frac{1}{1-m_{y|xz}} \right\}, \tag{23}$$

and thus, the Dirichlet parameters for the beta distribution are:

$$(\alpha, \beta) = s_{y|xz}(m_{y|xz}, 1 - m_{y|xz}). \tag{24}$$

Finally, the inferred opinion maps to

$$\omega_{Y\|xz} = \frac{(\alpha-1, \beta-1, 2)}{s_{y|xz}}. \tag{25}$$

The MoM method forms the best opinion to represent the mean and spread for the inferred conditional probability. However, it is computationally expensive as it requires numerical techniques to compute a triple integral. In this paper, the implementation of MoM employs the `triplequad` function in Matlab.

To demonstrate the effectiveness of moment matching, Figure 2 compares the actual inferred distribution to the approximate beta distribution for a given three node network when the observed values are $\bar{x}$ and $\bar{z}$ and a different number of observations N is used to determine the opinions for the conditionals. For this network, the ground truth conditional probability is $p_{y|\bar{x}\bar{z}} = 0.7231$ as represented by the dotted vertical lines in the figure. The figure shows that as N increases, the beta distribution better approximates the actual distribution that can be inferred from the opinions of the conditional. Furthermore, the spread for these distributions are decreasing as N increases. Interestingly, the mode of these distribution are not close to the ground truth until N is very large, i.e., $N = 1000$. Nevertheless, the ground truth is within the effective support of the distribution because the spread of the distributions are representative of the lack of knowledge about the ground truth. The remainder of the paper only considers the small to moderate sample sizes where $N \leq 100$ and the inferred distributions are not necessarily centered near the ground truth as illustrated in Figure 2.

B. Uncertainty Scaling

When defining the operations, the subjective logic framework tries to precisely match the mean of the output distribution while maximizing the uncertainty under constraints specific to the operation. In fact, the problem of determining the opinion for the inferred conditional given by (16) is a generalization of the inversion process in subjective logic abduction reviewed in section II where the opinion for the p_1 term is vacuous, i.e., $\omega_1 = (0, 0, 1)$. Inspired by the scaling used in the recent inversion process, the Uncertainty Scaling (US) method reduces the maximum possible uncertainty to maintain the desired mean value by scaling terms based upon the uncertainty of the three conditionals in (16). Because the p_1 is not vacuous, the US method does not include the irrelevance factor Ψ.

The first step of US is to determine the expected probability (the mean) of the inferred opinion. Unlike other subjective logic operations such as deduction, the expectation given by (20) has no closed form solution. Following the ideas from [10], US approximates the expected probability as:

$$m_{y|xz} = \frac{m_{y|x} m_{z|y}}{m_{y|x} m_{z|y} + (1 - m_{y|x}) m_{z|\bar{y}}}. \tag{26}$$

Given this mean value, the maximum possible uncertainty such that the beliefs are non-negative is the following:

$$\hat{u}_{Y|xz} = \min\{2m_{y|xz}, 2(1 - m_{y|xz})\}. \tag{27}$$

The scaling factor to reduce the uncertainty is a disjunctive combination of the average uncertainty for the backward conditionals:

$$u^w_{Z|y} = u_{Z|y} m_{y|x} + u_{Z|\bar{y}} (1 - m_{y|x}) \tag{28}$$

and the uncertainty for the forward conditional $u_{Y|x}$. Thus, the scaling factor is

$$\tilde{u}_{Y|xz} = u^w_{Z|y} + u_{Y|x} - u^w_{Z|y} u_{Y|x} \tag{29}$$

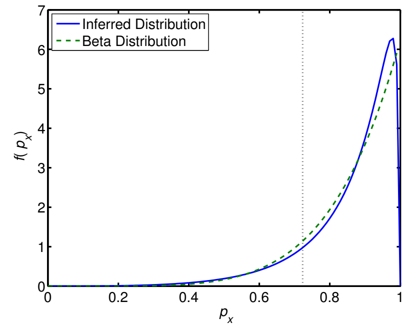
so that the adjusted uncertainty is

$$u_{Y|xz} = \hat{u}_{Y|xz} \tilde{u}_{Y|xz}. \tag{30}$$

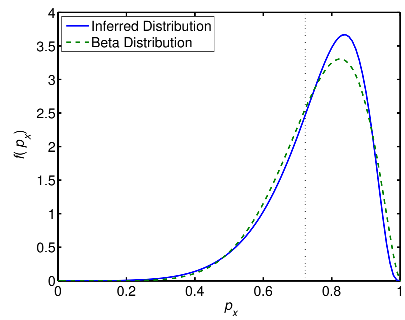
Finally, the inferred option is given by

$$\begin{aligned} \omega_{Y\|xz} = & \left(\; m_{y|xz}, \quad 1 - m_{y|xz}, \quad 0 \; \right) \\ & + u_{Y|xz} \left(\; -\tfrac{1}{2}, \quad -\tfrac{1}{2}, \quad 1 \; \right). \end{aligned} \tag{31}$$

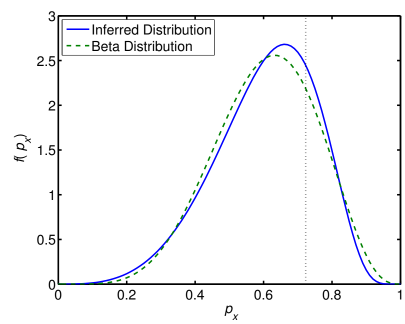
The US method is much faster than the MoM method. However, the selection of the uncertainty is based upon heuristic scaling values. Figure 3 shows how the US method approximates the true distribution using the same observations as in Figure 2. The uncertainty scaling actually causes US to be overly optimistic.



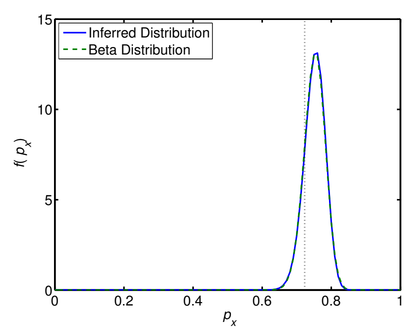
(a)



(b)



(c)



(d)

Fig. 2. Comparison of the actual inferred distribution to the beta approximation with ground truth $p_{y|\bar{x}\bar{z}} = 0.7231$ via moment matching when training a given network using various amounts of joint observations: (a) $N = 10$, (b) $N = 50$, (c) $N = 100$, and (d) $N = 1000$.

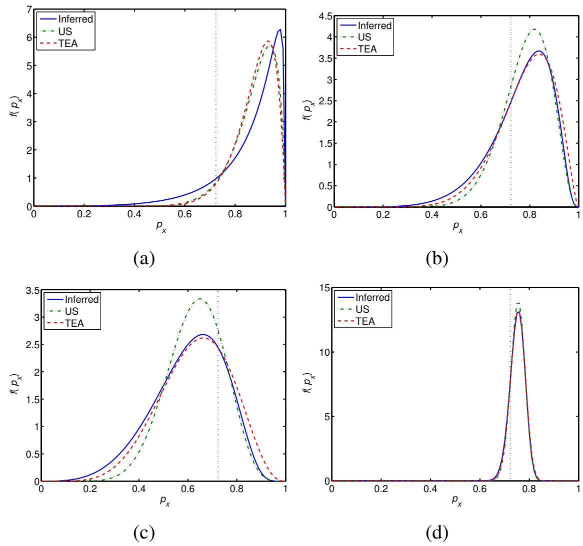


Fig. 3. Comparison of the actual inferred distribution to the beta approximation via US and TEA with ground truth $p_{y|\bar{x}\bar{z}} = 0.7231$ using various amounts of joint observations: (a) $N = 10$, (b) $N = 50$, (c) $N = 100$, and (d) $N = 1000$.

C. Taylor Expansion Approximation

The Taylor Expansion Approximation (TEA) method is derived by taking a first order Taylor series expansion approximation of (16) before taking the expected probabilities in (20) and (21) to approximate the moments. The first order Taylor series expansion is given by:

$$\begin{aligned} p_{y|xz} \approx & \frac{m_{y|x}m_{z|y}}{d} + \frac{m_{z|y}m_{z|\bar{y}}}{d^2}(p_{y|x} - m_{y|x}) \\ & + \frac{m_{y|x}(1-m_{y|x})m_{z|\bar{y}}}{d^2}(p_{z|y} - m_{z|y}) \\ & + \frac{m_{y|x}(1-m_{y|x})m_{z|y}}{d^2}(p_{z|\bar{y}} - m_{z|\bar{y}}), \end{aligned} \tag{32}$$

where $d = m_{y|x}m_{z|y} + (1-m_{y|x})m_{z|\bar{y}}$. It easy to see that inserting (32) into (20) leads to the same mean approximation as the US method given by (26). Inserting the square of the (32) into (21) leads to

$$\begin{aligned} v_{y|xz} = & \frac{m_{y|x}^2 m_{z|y}^2}{d^2} + \frac{m_{z|y}^2 m_{z|\bar{y}}^2}{d^4}\frac{m_{y|x}(1-m_{y|x})}{s_{y|x}+1} \\ & + \frac{m_{y|x}^2(1-m_{y|x})^2 m_{z|\bar{y}}^2}{d^4}\frac{m_{z|y}(1-m_{z|y})}{s_{z|y}+1} \\ & + \frac{m_{y|x}^2(1-m_{y|x}^2)m_{z|y}^2}{d^4}\frac{m_{z|\bar{y}}(1-m_{z|\bar{y}})}{s_{z|\bar{y}}+1}. \end{aligned} \tag{33}$$

The remaining steps of TEA determine the beta distribution that matches these moments using (22)-(25) just as in the MoM method. TEA uses an approximation so that the expected probabilities lead to closed form values. By avoiding numerical integration TEA is significantly less computationally extensive than the MoM method. However, it makes an approximation that provides the same expected probability values as the US method. Unlike the US method, TEA incorporates a more principled approach to determine the Dirichlet strength, and thus the uncertainty. However, the accuracy for which the resulting opinion $\omega_{Y\|xz}$ provides a reasonable representation of the inferred distribution is dependent upon the quality of the initial Taylor series approximation. Figure 3 shows that TEA fits the inferred distribution very well for $N \geq 50$.

V. Evaluation

A. Confidence Interval

As mentioned earlier, a subjective opinion about a variable distinguishes itself from a simple probability estimate in that it also indicates the spread of the probability estimate via the uncertainty. Most papers that define or improve upon a subjective logic operator do not actually validate the uncertainty in the resulting opinion. We believe that subjective logic operators could be evaluated via Monte Carlo simulations where observations are generated by ground truth probabilities, and the spread and mean of the derived opinion could be compared against the ground truth. The evaluation of the quality of the mean is easily accomplished by the root means squared error (RMSE). In [16], the spread was evaluated by confidence intervals assuming a Gaussian distribution. This is meaningful when the number of observations are high so that the beta distribution is approximately Gaussian. For a smaller number of observations, the confidence interval needs to be determined from the underlying beta distribution.

To determine whether or not the subjective opinion is a good representation of the spread of possible probabilities, i.e., the uncertainty (or Dirichlet strength) is properly determined and the beta distribution characterizes the output distribution for the operator, the beta distribution is used to determine upper and lower bounds for a confidence interval at desired level γ via

$$\ell(\omega_{Y\|xz}, \gamma) = q_* \quad \text{and} \quad \nu(\omega_{Y\|xz})(\gamma) = q^* \tag{34}$$

$$\text{such that} \int_{-\infty}^{q_*} f_\beta(p|\omega_{Y\|xz})dp = \frac{1-\gamma}{2} \tag{35}$$

$$\text{and} \int_{q^*}^{\infty} f_\beta(p|\omega_{Y\|xz})dp = \frac{1-\gamma}{2}. \tag{36}$$

The actual (or empirical) level $\tilde{\gamma}$ is determined by counting the number of times the ground truth actually falls in the confidence interval

$$\ell(\gamma, \omega_{Y\|xz}) \leq p_{y|xz} < \nu(\gamma, \omega_{Y\|xz}). \tag{37}$$

A plot of the actual versus desired confidence interval levels reveals the quality of the subjective logic operator. An ideal operator will generate a straight line $\tilde{\gamma}(\gamma) = \gamma$. The degree to which $\tilde{\gamma}(\gamma)$ approximates the ideal line indicates the quality to which the output subjective opinions truly represent the proper distribution for the output probability.

B. Monte Carlo Simulations

To evaluate the three inference methods over the three node network in Figure 1, 100 different instantiations of the network

Method	$N = 10$	$N = 50$	$N = 100$
MoM	0.2039	0.1304	0.1008
US/TEA	0.2055	0.1303	0.1021

TABLE I
RMSE BETWEEN THE EXPECTED PROBABILITY AND GROUND TRUTH FOR THE THREE INFERENCE METHODS AT VARIOUS VALUES OF N.

Method	$N = 10$	$N = 50$	$N = 100$
MoM	0.2176	0.1381	0.1034
US	0.1676	0.1084	0.0778
TEA	0.1837	0.1188	0.0935

TABLE II
ESTIMATION OF THE RMSE USING THE UNCERTAINTY FOR THE THREE INFERENCE METHODS AT VARIOUS VALUES OF N.

were generated by drawing the conditional and marginal probabilities from the uniform distribution over $[0, 1)$. For each network instantiation, the training procedure from Section III was performed 100 times using N observations to generate 10000 instances of trained subjective networks with specified conditional and marginal opinions. For each instances, the three different inference methods derived the four different inferred opinions $\omega_{Y\|xz}$, $\omega_{Y\|\bar{x}z}$, $\omega_{Y\|x\bar{z}}$, and $\omega_{Y\|\bar{x}\bar{z}}$. These opinions are then compared to the ground truth conditional probabilities. The inference methods are evaluated it terms of both accuracy and computational complexity for various values of N.

C. Inference Accuracy

Table I provides the RMSE for the expected probability for the three inference methods for various numbers of training observations. The US and TEA provide the same expected probabilities, and thus, they exhibit the same error. The expected probability of the MoM represents a minimum variance estimator for the ground truth probabilities conditioned on the available observations. Thus, it exhibits the smallest RMSE, but the error for US and TEA is only slightly higher. In fact, for the $N = 50$ case all methods essentially exhibit the same RMSE performance within numerical precision. In short, the mean approximation given by (26) is very reasonable. As expected, the error decreases as the number of training observations grows.

Table II provides the root mean value of expected variance derived from the inferred opinion as given by (7). These values are essentially predicting the RMSE in the absence of ground truth. The MoM prediction tracks the actual RMSE very well by slightly overestimating the values. The US method is consistently under estimating the RMSE, and TEA is providing improving prediction capabilities as N grows.

To determine how well the inferred subjective opinions match the true inferred posterior distribution, the actual confidence interval bound level $\tilde{\gamma}$ was computed over desired levels from zero to one. Figure 4 plots the actual versus the desired levels for the three inference methods for various values of N. Both the MoM and TEA methods better match the desired slope one line as N increases.The MoM is clearly better than TEA for $N = 10$. The US method does not improve with larger N. It is slightly better than TEA for $N = 10$ but is much worse for the other values of N. Apparently, the first order Taylor series expansion requires the spread for the input conditional not to be overly large for the approximation to be effective.

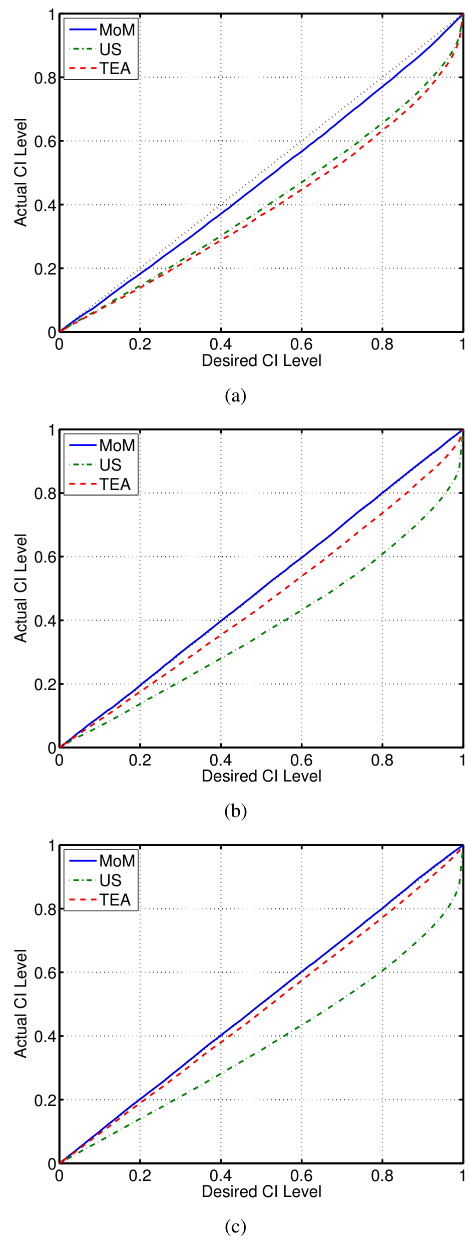


Fig. 4. Actual versus desired confidence interval level derived from the inferred opinion obtained from the various inference methods: (a) $N = 10$, (b) $N = 50$, and (c) $N = 100$.

Method	$N = 10$	$N = 50$	$N = 100$
MoM	0.465s	0.781s	1.17s
US	19.6μs	21.4μs	22.5μs
TEA	22.8μs	34.0μs	32.2μs

TABLE III
AVERAGE EXECUTION TIME FOR THE THREE INFERENCE METHODS OVER VARIOUS VALUES OF N.

D. Computational Complexity

Table III provides the average execution time to estimate the three inference methods for various values of N. The TEA methods is slightly slower than the US, and the execution times for both methods are essentially invariant to N. The MoM method is a factor of 20000 times slower than the US and TEA methods for $N = 10$, and the runtime for the MoM increases as N increases. This is due to the numerical integration of beta distributions with ever increasing sharpness for the MoM as N increases.

VI. CONCLUSIONS

This paper takes steps to expand subjective logic to enable inference over subjective networks. Subjective networks are probabilistic graph models whose conditional probabilities are not known with absolute certainty. Rather, knowledge about the conditional probabilities are expressed as beta distributions in the form of subjective opinions. This paper considers a three node subjective network and introduces three different inference methods of varying computational complexity to form an opinion about the central node (or variable) conditioned on the direction observation of the values of the other two nodes. In this case, the uncertainty is dictated by the uncertainty in the conditional probabilities. The MoM method determines the beta distribution that best characterizes the actual distribution of possible inferred probabilities associated to the central node. The US method approximates the mean by assuming the conditional expected probabilities are absolute and scales the maximal possible uncertainty by the uncertainty of the appropriate conditionals. Finally, the TEA method finds the best beta distribution characterization via a first order Taylor series approximation expansion for the inferred probability so that the moments can be computed analytically in closed form. The expected probabilities that result for the US and TEA methods are equivalent to performing standard Bayesian network inference where the given opinions about the conditionals and marginals are transformed to actual probabilities by (5). In essence, these methods augment the traditional Bayesian network results with a characterization of their uncertainty.

Monte Carlo simulations reveal that MoM best characterizes the actual distribution. TEA is able to characterize the actual distributions for moderate training set sizes. US is the least effective at characterizing the underlying uncertainty because it uses heuristics to determine the uncertainty. Both TEA and US are significantly faster than MoM. It is desirable in future work to determine a computationally fast inference method that works nearly as well as MoM under large uncertainty where the number of training samples is small, e.g., $N = 10$.

The overall goal is to develop a basic set of subjective logic operations to enable effective message passing for inference over subjective networks. Subjective logic deduction enables the flow of information along nodes, and these flows begin at directly observed nodes. This paper investigated inference at nodes where the multiple flows meet and its results could be easily extended to multivariate variables and naïve Bayesian networks, where several nodes are children of a common, root node and at most one edge can go from one of the children to the root. The situation of two or more edges entering a node, known in the Bayesian literature as *v-structure*, will be harder to model and will require a further, non trivial extension of the present operators. Finally, further work will be devoted to extending the general form of subjective evidence, which contemplates non absolute opinions on the observed nodes, from two node networks to multi-node networks.

REFERENCES

[1] D. Heckerman, D. Geiger, and D. M. Chickering, "Learning Bayesian networks: The combination knowledge and statisitcal data," *Machine Learning*, vol. 20, pp. 197–243, 1995.

[2] P. Gärdenfors and N.-E. Sahlin, "Unreliable probabilities, risk taking, and decision making," *Synthese*, vol. 53, no. 3, pp. 361–386, 1982.

[3] P. Walley, *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall, 1991.

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

[5] A. Jøsang, "A logic for uncertain probabilities," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, no. 3, pp. 279–311, Jun. 2001.

[6] ——, "Subjective logic," Jul. 2011, draft book in preparation.

[7] P. P. Shenoy, "A valuation-based language for expert systems," *Int. Journal of Approximate Reasoning*, vol. 3, no. 2, pp. 383–411, 1989.

[8] A. Jøsang, S. Pope, and M. Daniel, "Conditional deduction under uncertainty," in *Proceedings of the 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2005)*, 2005.

[9] A. Jøsang, "Conditional reasoning with subjective logic," *Journal of Multiple-Valued Logic and Soft Computing*, vol. 15, no. 1, pp. 5–38, 2009.

[10] A. Jøsang and F. Sambo, "Inverting conditional opinions in subjective logic," in *Proceedings of the 20th International Conference on Soft Computing (MENDEL 2014)*, Bruno, 2014.

[11] L. M. Kaplan, M. Sensoy, Y. Tang, S. Chakraborty, C. Bisdikian, and G. de Mel, "Reasoning under uncertainty: Variations of subjective logic deduction," in *Proc. of the International Conference on Information Fusion*, Istanbul, Turkey, Jul. 2013.

[12] P. Costa, K. Laskey, E. Blasch, and A. L. Jousselme, "Towards unbiased evaluation of uncertainty reasoning: The URREF ontology," in *Proc. of the 15th Int. Conf. on Information Fusion*, Jul. 2012, pp. 2301–2308.

[13] J. P. de Villiers, G. Pavlin, P. Costa, K. Laskey, and A.-L. Jousselme, "A URREF interpretation of Bayesian network information fusion," in *Proc. of the 17th Int. Conf. on Information Fusion*, Jul. 2014.

[14] A. Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," *Decision Support Systems*, vol. 43, no. 2, pp. 618–644, Mar. 2007.

[15] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufman Publishers, 1988.

[16] L. Kaplan, M. Şensoy, S. Chakraborty, and G. de Mel, "Partial observable update for subjective logic and its application for trust estimation," *Information Fusion*, 2015.