Human-in-the-Loop Situational Understanding via Subjective Bayesian Networks*

Dave Braines
IBM, UK
Cardiff University, UK

Anna Thomas
IBM, UK

Lance Kaplan
Army Research Laboratory, USA

Murat Şensoy
Ozyegin University, Turkey
Cardiff University, UK

Magdalena Ivanovska
University of Oslo, Norway

Alun Preece and **Federico Cerutti**
Cardiff University, UK

Abstract

In this paper we present a methodology to exploit human-machine coalitions for situational understanding. Situational understanding refers to the ability to relate relevant information and form logical conclusions, as well as identifying gaps in information. This process requires the ability to reason inductively, for which we will exploit the machines' ability to 'learn' from data. However, important phenomena are often rare in occurrence, thus severely limiting the availability of instance data for training, and hence the applicability of many machine learning approaches. Therefore, we present the benefits of Subjective Logic Bayesian Networks—i.e. Bayesian Networks with imprecise probabilities—for situational understanding; and the potential role of conversational interfaces for supporting decision makers in the evolution of situational understanding.

1 Introduction

Human situational understanding is filled with inductive reasoning. You just landed at Heathrow Airport in London, UK: the sun is blazing in the sky and a glorious warm temperature of 23 Celsius (74 Fahrenheit) welcomes you in the South of Britain. On the basis of this observation, it is rational to conclude that usually the South of Britain enjoys a lovely weather, especially if the same happens the second day, and the third day, and the fourth day of your visit. From a human perspective general rules are therefore often derived on the basis of scarce data.

The scarcity of data is often not a problem, especially in those cases where we can have access to an *oracle*, mostly an expert in the domain. You might receive a useful piece of information from a friend who lived in the South of Britain for years, or you can access historical data and statistics showing that it is not the case that usually the South of Britain enjoys lovely weather and therefore this apparent normality is in fact an exception. Oracles can help in overcoming scarcity of actual data through access to other information or rules which are relevant to the domain.

As humans we therefore apply analyses and judgements to relevant information "to determine the relationships of the factors present and form logical conclusions concerning threats [. . .], opportunities [. . .], and gaps in information" [Dostal, 2007]. This is *situational understanding*.

Machine learning approaches are potentially powerful allies in situational understanding [Brannon *et al.*, 2009]. Machine learning algorithms are able to efficiently handle large quantities of information, which is extremely useful to support inductive reasoning in situational understanding, as well as deriving logical conclusions. However, they are generally useless for identifying gaps in information as well as in providing insights such as those that could be provided by oracles. Moreover, the best algorithms for machine learning often assume the existence of a large training set: unfortunately this assumption is often unrealistic. The need for less training data is particularly important in situational understanding problems where many important phenomena will be rare in occurrence, severely limiting the availability of instance data and, hence, the applicability of many machine learning approaches, including Bayesian and Deep Learning [LeCun *et al.*, 2015] approaches. Coupled with this, supporting human analysts in terms of more effective communication of uncertain information is also a key issue in situational understanding problems [Dhami *et al.*, 2015].

In this paper we propose a human-machine coalition partnership for real-world situational understanding exploiting the strengths of each member in the coalition. Machines' strengths are linked to data analysis, and we explicitly address the unrealistic assumption of large training sets which could undermine the role of machine agents in such a human-machine coalition. Moreover, human experts are usually considered useful oracles, and we need to provide useful human-machine interfaces in order to support co-design and co-evolution of the coalition for situational understanding. Specifically we consider a system within which the human

*This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-16-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

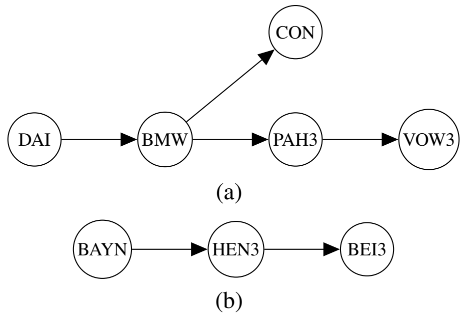


Figure 1: German automotive (a) and cosmetic (b) company dependency networks provided as input

	Company	Comment
BAYN	Bayer	Pharmaceutical company
BEI3	Beiersdorf	Cosmetic company
BMW	BMW	Automotive manufacturer
CON	Continental	Tyre manufacturer
DAI	Daimler	Automotive manufacturer
HEN3	Henkel	Cosmetic company
PAH3	Porsche	Automotive manufacturer
VOW3	Volkswagen	Automotive manufacturer

Table 1: Companies considered from the German stock market in Figures 1-2

agents can contribute to or correct the machine agent parts of the system.

To exemplify our proposal, we discuss a running example about German market in Section 2, and in Section 3 we exploit one of the machines' strengths: Performing inductive reasoning with quantitive measures such as probabilities. We discuss a robust approach to handling uncertain information from a rather scarce dataset, namely Subjective Logic Bayesian Networks, an extension of Bayesian Networks using uncertain probabilities. This helps us towards overcoming one of the main issues related to Bayesian networks: the lack of information about the certainty of the trained model.

We then show, in Section 4, that Subjective Logic Bayesian Networks are well suited for situational understanding. Our tests show that they provide more accurate results compared to other approaches to Bayesian networks with uncertain probabilities, such as Credal networks [Zaffalon and Fagiuoli, 1998] and belief networks [Smets, 1993].

2 Human-Machine Coalitions for Situational Understanding

Let us suppose you are an advisor for investors who want to enter the German stock market. For brevity, let us suppose that a colleague has provided the two high-level dependency networks depicted in Figure 1, showing on one hand dependencies between Daimler, BMW, Continental, Porsche, and Volkswagen (automotive companies); and on the other hand dependencies between Bayer, Henkel, and Beiersdorf (cosmetic companies). Those dependencies suggest that the stock price of those companies are linked such that a significant variation of the stock price of Daimler will influence a variation in the stock price of BMW.

Let us suppose you have the privilege to use our conversational interface for interacting with such dependencies network, see Figure 2. Such a conversational interface would allow you to also to express additional information, in particular that there is a dependency between Bayer and Daimler. This enables the human user to therefore an as an 'oracle' as described previously, contributing relevant information to the machine agent based on their wider knowledge of the domain in question. Indeed, Daimler and Bayer are regularly traded by over-the-counter (OTC) list shares[1] like INTL FCStone Financial.[2]

3 Reasoning under Uncertainty with Limited Data

3.1 Dealing with Uncertainty: Subjective Logic

Subjective logic is a form of uncertain probabilistic reasoning [Jøsang, 2016]. It expands the notion of a probability of a variable value to a distribution of possible probabilities. In general, the variable can take on one of K mutually exclusive values. This paper considers binary variables such as X that can take on the value of true or false, i.e., $X = \mathfrak{x}$ or $X = \bar{\mathfrak{x}}$. The value of X does change over different instantiations, and there is an underlying ground truth value for the probability $\rho_X(x)$ of taking on the value in the domain $\mathbb{X} = \{\mathfrak{x}, \bar{\mathfrak{x}}\}$.

A subjective opinion can be formed by directly observing N_{ins} independent instantiations of X. If over these instantiations, n_x times $X = \mathfrak{x}$, $n_{\bar{x}} = N_{ins} - n_x$ times $X = \bar{\mathfrak{x}}$ and assuming an uninformative uniform prior, then the posterior knowledge of the ground truth outcome probability of X is known to follow the beta distribution

$$f_{\boldsymbol{\beta}}(p_x|\omega_X) = \frac{1}{\beta(\alpha_x, \alpha_{\bar{x}})} p_x^{\alpha_x - 1}(1-p_x)^{\alpha_{\bar{x}}-1} \tag{1}$$

for $0 \leq p_x \leq 1$, where $\beta(\cdot)$ is the beta function and the beta parameters $\boldsymbol{\alpha} = [\alpha_x, \alpha_{\bar{x}}] = [n_x+1, n_{\bar{x}}+1]$ are one particular representation of the opinion ω_X. The opinion ω_X in belief space is a tuple of belief $b_X = \frac{n_x}{s_X}$, disbelief $d_X = \frac{n_{\bar{x}}}{s_X}$ and uncertainty $u_X = \frac{2}{s_X}$, where $s_X = \alpha_x + \alpha_{\bar{x}}$ is the Dirichlet strength. Therefore, a tuple $\langle b_X, d_X, u_X \rangle$ identifies a point in a 3D space. However, since the belief masses are positive and sum to one, such a 3D space can be flattened into a 2D triangle as depicted in Figure 3. Following [Jøsang, 2016, p. 49] we can partition the 2D space of subjective logic opinions for (lossy) representation using fuzzy natural language terms such as "High Confidence" and "Very Likely". Such terms can be made even more consumable for human users when embedded within larger natural language sentences such as: "When BAYN stock price changes, there is *high confidence* that HEN3 stock price is *very likely to change*" that can summarise the subjective opinion $\langle 0.8, 0.1, 0.1 \rangle$.

[1] OTC trades refers to stock trades via a dealer network as opposed to on a centralised exchange.

[2] `https://goo.gl/lTruuv` (on 4th May 2017).

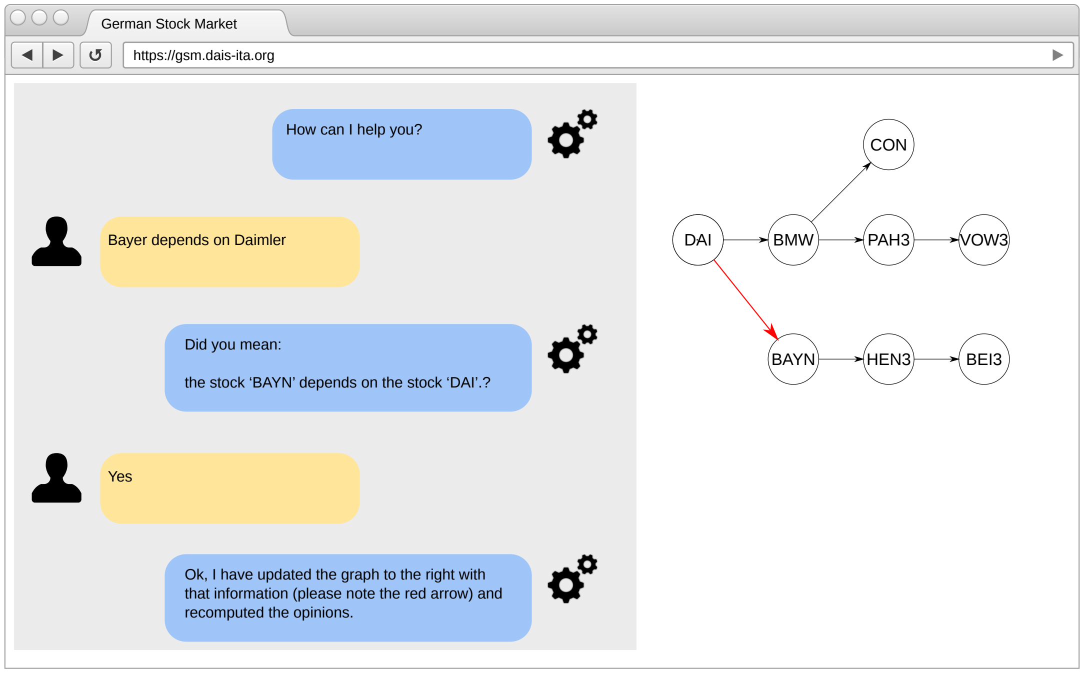


Figure 2: Mockup depicting the action of updating a dependency network through our proposed conversational interface

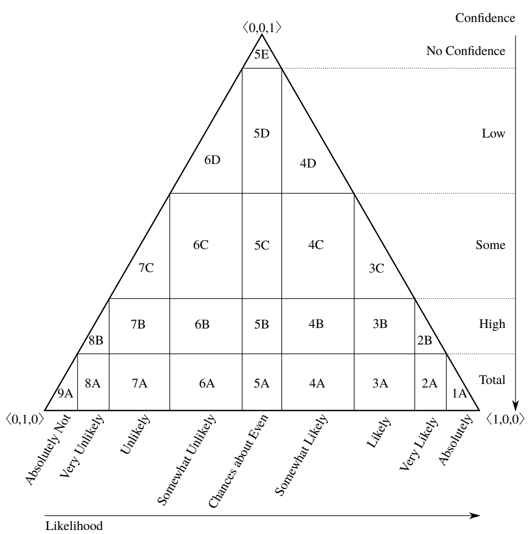


Figure 3: Subjective Logic 2D triangle and areas for fuzzy labels, adapted from [Jøsang, 2016, p. 49].

In this paper, it will be convenient to represent the subjective opinion ω_X by the mean and Dirichlet strength of the corresponding beta distribution. The mean represents the projected probability that converts the opinion in the pignistic probabilities, and is given by

$$P_X(\mathfrak{x}) = \frac{\alpha_x}{s_X} \quad \text{and} \quad P_X(\bar{\mathfrak{x}}) = \frac{\alpha_{\bar{x}}}{s_X}. \tag{2}$$

The mean of the corresponding beta distribution is identified by m_X while the variance of the corresponding beta distribution,

$$\sigma_X^2 = \frac{P_X(\mathfrak{x})P_X(\bar{\mathfrak{x}})}{s_X + 1}, \tag{3}$$

is a function of the projected probabilities and Dirichlet strength of the subjective opinion. This expression is used in the experiments to predict the root mean squared error between the projected probability $P_X(x)$ and the actual ground truth $\rho_X(x)$. Subjective opinions naturally extend to subjective conditional opinions, where for example, the opinion for X conditioned on Y and Z is interpreted as the set $\{\omega_{X|y,z} : y \in \mathbb{Y}, z \in \mathbb{Z}\}$, and $\omega_{X|y,z}$ represents the effective number of times that $X = \mathfrak{x}$ or $X = \bar{\mathfrak{x}}$ when $Y = y$ and $Z = z$ while jointly observing X, Y, and Z.

3.2 Dealing with Limited Data: Subjective Bayesian Network

The Subjective Bayesian network (SBN) was first proposed in [Ivanovska *et al.*, 2015], and it is an uncertain Bayesian network where the conditionals are subjective opinions instead of dogmatic probabilities. In other words, the conditional probabilities are known within a beta distribution. A SBN reflects the knowledge about a Bayesian network when limited historical data is used to learn the conditionals. The inference of a SBN leads to an opinion about the marginal probability of all the unobserved variables conditioned on the values of the observed variables. While different types of SBNs were discussed in [Ivanovska *et al.*, 2015], this paper focuses on the type that uses the beta distribution interpretation of the subjective opinion to compute uncertainty. This section reviews subjective belief propagation (SBP) which was introduced for trees in [Kaplan and Ivanovska, 2016] and extended for singly-connected networks in [Kaplan and Ivanovska, 2017] for this class of SBNs.

SBP extends the Belief Propagation (BP) inference method of Pearl [Pearl, 1986]. In BP, π- and λ-messages are passed from parents and children, respectively, to a node, i.e., variable. The node uses these messages to formulate the inferred marginal probability of the corresponding variable. The node also uses these messages to determine the π- and λ-messages to send to its children and parents, respectively. In SBP, the π- and λ-messages are subjective opinions characterized by a projected probability and Dirichlet strength.

The SBP formulation approximates output messages as beta distributed using the methods of moments and a first order Taylor series approximation to determine the mean and variance of the output messages in light of beta distributed input messages. The details of the derivations are provided in [Kaplan and Ivanovska, 2016; 2017]. Given a node X with m parents U_i for $i = 1, \ldots, m$, the subjective opinions of the π-messages sent to X are characterized by the projected probabilities $\pi_{U_i,X}(x)$ and Dirichlet strengths $s_{\pi_{U_i,X}}$. Likewise given that X has k children Y_j for $j = 1, \ldots, k$, the subjective opinions of the λ-messages sent to X are characterized by the projected probabilities $\lambda_{U_i,X}(x)$ and Dirichlet strengths $s_{\lambda_{U_i,X}}$. Node X processes these opinions to form the fused π opinion

$$\pi_X(x) = \sum_{u_1,\ldots,u_m} P(x|u_1, \ldots, u_m) \prod_{i=1}^{m} \pi_{U_i,X}(u_i), \quad (4)$$

$$s_{\pi_X} \quad = \quad \frac{\pi_X(x)(1 - \pi_X(x))}{\sigma^2_{\pi_X}} - 1, \quad (5)$$

where the variance $\sigma^2_{\pi_X} = V_{\pi_X} - m^2_{x|o}$,

$$V_{\pi_X} = \sum_{u_1,\ldots,u_m} \sum_{u'_1,\ldots,u'_m} g(x, x; u_1, \ldots, u_m; u'_1, \ldots, u'_m) \cdot \cdot \prod_{i=1}^{m} h(u_i, u'_i), \quad (6)$$

$$g(x, x'; u_1, \ldots, u_m; u'_1, \ldots, u'_m) = p_{x|u_1 \ldots u_m} p_{x|u'_1 \ldots u'_m} + (-1)^{x \neq x'} \delta_{\mathbf{u},\mathbf{u}'} \frac{p_{x|u_1 \ldots u_m}(1 - p_{x|u_1 \ldots u_m})}{s_{X|u_1 \ldots u_m} + 1} \quad (7)$$

where $\mathbf{u}$ is an arbitrary joint assignment of the variables $U_1, \ldots, U_m$,

$$\delta_{\mathbf{u},\mathbf{u}'} = \begin{cases} 1, & \text{if } u_j = u'_j, \text{for } j = 1, \ldots, m \\ 0, & \text{otherwise} \end{cases}$$

is the Kronecker delta function, and

$$h_\pi(u_i, u'_i) = \pi_{U_i,X}(u_i)\pi_{U_i,X}(u'_i) + (-1)^{u_i \neq u'_i} \frac{\pi_{U_i,X}(u_i)(1 - \pi_{U_i,X}(u_i))}{s_{\pi_{U_i,X}} + 1}. \quad (8)$$

The fused λ-message is

$$\lambda_X(x) = \alpha_\lambda \prod_{j=1}^{k} \lambda_{Y_j,X}(x) \ , \quad (9)$$

$$s_{\lambda_X} = \left(\sum_{j=1}^{k} \frac{\lambda_X(\mathfrak{x})\lambda_X(\bar{\mathfrak{x}})}{\lambda_{Y_j,X}(\mathfrak{x})\lambda_{Y_j,X}(\bar{\mathfrak{x}})} \frac{1}{s_{\lambda_{Y_j,X}} + 1} \right)^{-1} - 1,$$

where α_λ is a normalizing constant so that $\lambda_X(x)$ sums to one over its domain $\mathbb{X}$.

The π and λ-opinions are fused to determine the marginal opinion for node X:

$$P(x|o) = \alpha_f \pi_X(x) \lambda_X(x) \ , \quad (10)$$

$$s_X = \left(\frac{P_X(\mathfrak{x})P_X(\bar{\mathfrak{x}})}{\pi_X(\mathfrak{x})\pi_X(\bar{\mathfrak{x}})} \frac{1}{s_{\pi_X} + 1} + \frac{P_X(\mathfrak{x})P_X(\bar{\mathfrak{x}})}{\lambda_X(\mathfrak{x})\lambda_X(\bar{\mathfrak{x}})} \frac{1}{s_{\lambda_X} + 1} \right)^{-1} - 1,$$

where α_f is also a normalizing constant.

The opinion for the message that node X sends to parent U_i is

$$\lambda_{X,U_i}(u_i) = \alpha_b \sum_x \lambda_X(x) \sum_{\{u_1,\ldots,u_m\} \setminus \{u_i\}} P(x|u_1, \ldots, u_i, \ldots, u_m) \cdot \cdot \prod_{j \neq i} \pi_{U_j,X}(u_j), \quad (11)$$

$$s_{\lambda_{X,U_i}} = \frac{\lambda_{X,U_i}(u_i)(1 - \lambda_{X,U_i}(u_i))}{\sigma^2_{\lambda_{X,U_i}}} - 1, \quad (12)$$

where

$$\sigma^2_{\lambda_{X,U_i}} = \alpha_b^2 \left(\lambda^2_{X,U_i}(\bar{x})\sigma^2_{uu} + \lambda^2_{X,U_i}(x)\sigma^2_{\bar{u}\bar{u}} + -2\lambda_{X,U_i}(x)\lambda_{X,U_i}(\bar{x})\sigma^2_{u\bar{u}} \right), \quad (13)$$

$$\sigma^2_{zv} = \sum_x \sum_{x'} h_\lambda(x, x') \sum_{\{u_1,\ldots,u_m\} \setminus \{z\}} \sum_{\{u'_1,\ldots u'_m\} \setminus \{v\}} g(x, x'; u_1, \ldots, z, \ldots, u_m; u'_1, \ldots, v, \ldots, u'_m) \prod_{j \neq i} h_\pi(u_j, u'_j), \quad (14)$$

and

$$h_\lambda(x, x') = \lambda_X(x)\lambda_X(x') + (-1)^{x \neq x'} \frac{\lambda_X(x)(1 - \lambda_X(x))}{s_{\lambda_X} + 1}, \quad (15)$$

and α_b is a normalizing constant.

Finally, the opinion message sent to the children of X are

$$\pi_{X,Y_j}(x) = \alpha_\pi \prod_{i \neq j} \lambda_{Y_i,X}(x)\pi_X(x), \quad (16)$$

$$s_{\pi_{X,Y_j}} = \left(\frac{\pi_{X,Y_j}(\mathfrak{x})\pi_{X,Y_j}(\bar{\mathfrak{x}})}{\pi_X(\mathfrak{x})\pi_x(\bar{\mathfrak{x}})} \frac{1}{s_{\pi_X} + 1} + + \sum_{i \neq j} \frac{\pi_{X,Y_j}(\mathfrak{x})\pi_{X,Y_j}(\bar{\mathfrak{x}})}{\lambda_{Y_i,X}(\mathfrak{x})\lambda_{Y_i,X}(\bar{\mathfrak{x}})} \frac{1}{s_{\lambda_{Y_i,X}} + 1} \right)^{-1} - 1,$$

where α_π is a normalizing constant.

The equations for the projected probability updates in SBP mirror the update equations in standard belief propagation due to the first-order Taylor approximation. Actually, the normalizing constants α_λ and α_β are superfluous in standard belief propagation, but necessary in SBP so that the λ message are proper subjective opinions. In short, SBP provides the same answer as belief propagation in the mean value. The difference is that SBP also provides a quantification of the uncertainty through the Dirichlet strength. On a technical note, SBP will actually increase the Dirichlet strength as computed in the update equations to ensure that all belief values are non-negative. We refer the interested reader to [Kaplan and Ivanovska, 2016; 2017] for more details. Finally, the information flow in SBP is exactly the same as in belief propagation. Namely, a node can send a message to one particular neighbor once it receives messages from all of its other neighbors.

4 Experimentation

4.1 Methodology

Subjective Bayesian Networks can learn a model of the domain with very limited number of observations, however, the inferred opinions through such network will become more certain as the number of observations increases. To measure how well these models can be learned with limited data and measure the uncertainty associated with the inferences, we build gold standard models, which are Bayesian Networks that are generated from much larger amount of observations. Therefore, we expect the gold standard models to be more accurate and certain than other models that are built with much limited number of observations.

For structure learning of the gold standard models, we used well-known K2 algorithm [Lerner and Malka, 2011]. The K2 algorithm is used to learn the best structure of a singly-connected Bayesian network to represent the interactions between the random variables. The resulting network serves as a surrogate for a subject matter expert who would use their background knowledge to create the network structure, for example via the conversational interface (see Figure 2). Further discussion on this topic is provided in the conclusion of the paper. Then, the conditional and marginal probabilities at each node at the network are calculated in the traditional manner using the entire available data.

We use real data to evaluate the quality of the uncertainty (or Dirichlet strength) in the subjective opinion inferred by SBP to represent the actual spread between the corresponding 'projected' and 'ground truth' probabilities that are well captured by the gold standard models. The full data is then divided into non-overlapping segments of N_{ins} instantiations (i.e., observations). Each segment represents the sparse data that would actually be available to train a SBN. A SBN is trained for each segment, and the set of exterior nodes, i.e., nodes with one single neighbor (either a parent or child), are considered to be observed. For each combination of possible values for these exterior nodes the marginal opinions for the interior nodes are inferred by SBP. Likewise, to establish the ground truth, the marginal probabilities are inferred by standard belief propagation using the underlying gold standard Bayesian network for the same values of the observed exterior nodes. Then, the marginal opinions and ground truths for all interior nodes are determined over all combinations of observed values and non-overlapping segments. Finally, the uncertainty of the marginal opinions are evaluated.

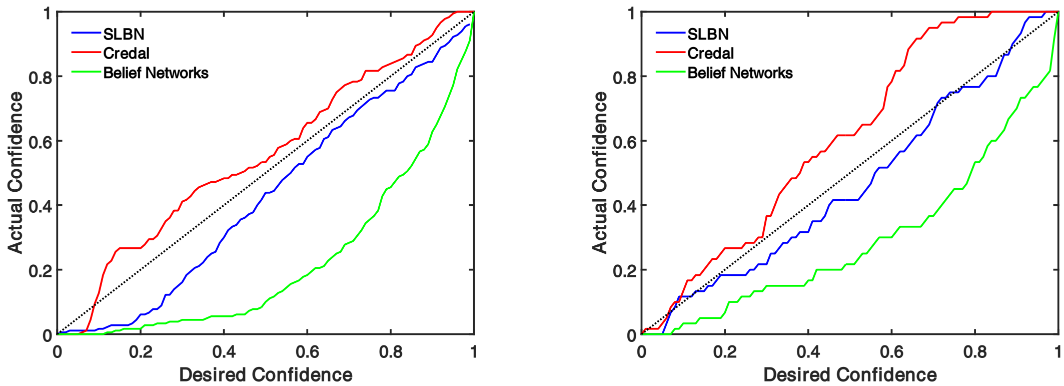
To evaluate the uncertainty quality, the actual root mean squared error (RMSE) between the projected and ground truth probabilities is calculated. Next, the predicted RMSE error is computed without knowledge of the ground truth as the square root of the average variance predicted from the opinions via (3). The similarity between the actual and predicted RMSE is one way to establish the quality of the uncertainty in the subjective opinions to characterize the spread between the projected and actual probabilities.

A even more precise method to determine the quality of the uncertainty characterization is to establish γ-confidence intervals from the opinions to capture the fraction of γ ground truths within these intervals. One then tabulates the fraction of times that the actual ground truth falls within the confidence interval. This is done for various values of $\gamma \in [0, 1]$, and the plot of the actual $\hat{\gamma}$ and the desired γ should follow a straight line as it should be the case that $\hat{\gamma} \approx \gamma$. A more detailed discussion can be found in [Kaplan *et al.*, 2015]. The quality of the subjective opinion ω_X should be judged on how well its expression of uncertainty captures the spread between its projected probability and the actual ground truth probability.

We compare the performance of SBP against previous methods to reason over uncertain probabilistic networks. Namely, we consider credal networks and belief networks, which are summarized below:

Credal Networks: A credal network over binary random variables extends a BN by replacing single probability values with closed intervals representing the possible range of probability values. The extension of Pearl's message-passing algorithm by the 2U algorithm for credal networks is described in [Zaffalon and Fagiuoli, 1998]. This algorithm works by determining the maximum and minimum value (an interval) for each of the target probabilities based on the given input intervals. It turns out that these extreme values lie at the vertices of the polytope dictated by the extreme values of the input intervals. As a result, the computational complexity grows exponentially with respect to the number of parents nodes. For the sake of comparison, we assume that our subjective network elicited from the given data corresponds to a credal network in the following way: If $\omega_x = [b_x, b_{\bar{x}}, u_X]$ is a subjective opinion on the probability p_x, then we have $[b_x, b_x + u_X]$ as an interval corresponding to this probability in the credal network. It should be noted that this mapping from the Beta distribution to an interval is consistent with past studies of credal networks [Karlsson *et al.*, 2008].

Belief Networks: In [Smets, 1993], Smets introduced a computationally efficient method to reason over networks via Dempster-Shafer theory. It is an approximation of a valuation-based system. Namely, a (conditional) subjective opinion $\omega_X = [b_x, b_{\bar{x}}, u_X]$ from our SBN obtained from data is converted to the following belief mass assignment: $m(x) = b_x$, $m(\bar{x}) = b_{\bar{x}}$ and $m(x \cup \bar{x}) = u_X$. (Note that in the binary case, the belief function overlaps with the belief mass assignment). The method exploits the disjunctive rule of combination (DRC) to compose beliefs conditioned on the Cartesian product space of the binary power sets. This enables both forward propagation and backward propagation after inverting the belief conditionals via the generalized Bayes' theorem (GBT). By operating in the Cartesian product space



(a) (b)

Figure 4: Comparing SBN against Belief Networks and Credal with $N_{train} = 10$ (sample size 2.74%) (a) and $N_{train} = 30$ (sample size 8.21%) (b) for the German stock exchange data. Best closest to the diagonal.

of the binary power sets, the computational complexity grows exponentially with respect to the number of parents, similar to the 2U algorithm for credal sets and our SBP method.

4.2 German Stock Exchange Predictions

Let us consider the case where a machine learning system is used to mine data from the German Stock Market, Börse Frankfurt. To simplify the scenario, let us consider a binary variable per company listed in Börse, where such a variable is $true$ if there is a significant increase (i.e. $+0.5\%$) in the company's stock value on a day, and $false$ otherwise. Let us then suppose that a well-known off-the-shelf algorithm for structure learning of dependencies among selected variables, such as K2 [Lerner and Malka, 2011] has been used. Using such an algorithm, the dependency networks highlighted in Figures 1(a) and 1(b) are derived. Table 1 explains the variables used in the dependency networks.

Figure 1(a) shows how there is a dependency between Daimler stock variations and BMW; between BMW and Porsche; between Porche and Volkswagen (all automotive manufacturers); and between BMW and Continental, a tyre manufacturer. Similarly, Figure 1(b) depicts the dependencies between Bayer—a pharmaceutical company—and Henkel—a company producing a variety of chemical products including ingredients of cosmetics; and between Henkel and Beiersdorf, cosmetic companies. Those dependencies are far from being a surprise, given that they are companies working in similar, or related, segments of the market. These two networks have then been merged to produce the single network described in Figure 2, thus introducing the dependency between Daimler and Bayer.

The gold standard Bayesian network is obtained by using all available data for (365 days) to determine the conditional probabilities. Then N_{train} days were used to generate $floor(365/N_{train})$ SBNs. Binary values were generated for the three nodes that have one edge, and the marginal probabilities (ground truth) and marginal opinions were generated via belief propagation and subjective belief propagation over the Bayesian and Subjective Bayesian networks, respectively. Table 2 lists actual and predicted RMSE for different approaches for different amount of observations. It indicates that SBN achieves pretty good error rate even with 10 days of observations (sample size 2.74%) and the error decreases to 0.05 when 30 days of data is used (sample size 8.21%). Figure 4 shows the ratio of the times the ground truth falls within the bounds—set at various significance levels—when building subjective logic Bayesian networks over 10 and 30 days. Our results indicate that SBN can estimate the ground truth probabilities much more accurate then Credal networks and Belief Networks. Especially, when $N_{train} = 30$, confidence level of the SBN is around the desired one, i.e., diagonal on the figures. Moreover, Table 2 lists actual and predicted RMSE for our approach and benchmark approaches when different amount of observations are used. SBN constantly provides the best error rate.

$N_{train} = 10$ (sample size 2.74%)

	SBN	Credal	Belief Networks
Actual RMSE	**0.124**	0.198	0.176
Predicted RMSE	**0.101**	0.187	0.132

$N_{train} = 30$ (sample size 8.21%)

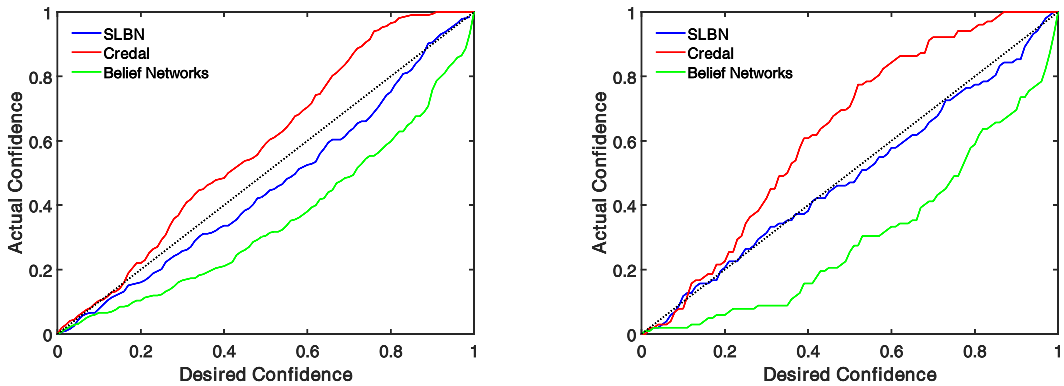
	SBN	Credal	Belief Networks
Actual RMSE	**0.047**	0.062	0.075
Predicted RMSE	**0.049**	0.089	0.061

Table 2: Error for the German stock exchange dataset. Gold standard trained with $N_{train} = 365$. Best results in bold.

4.3 Istanbul Stock Market Predictions

We also considered the dataset first derived in [Akbilgic *et al.*, 2014],[3] which considers stock exchange returns for several indexes, including those listed in Table 3. It is quite straight-

[3] `https://goo.gl/XzAZUX` (on 4th May 2017)



(a)

(b)

Figure 5: Comparing SBN against Belief Networks and Credal with $N_{train} = 10$ (sample size 1.86%) (a) and $N_{train} = 30$ (sample size 5.60%) (b) for the Istanbul stock market data. Best closest to the diagonal.

	Comment
SP	Standard & Poor's 500 Index Return.
DAX	Germany Stock Market Return
FTSE	UK Stock Market Return
NIK	Japan Stock Market Return
BVSP	Brazil Stock Market Return
EU	MSCI European Index Return
EM	MSCI Emerging Markets Index Return

Table 3: Indexes considered from the Istanbul Stock Exchange Data Set [Akbilgic *et al.*, 2014] in Figure 6.

forward to derive a dependency network such as Figure 6 between those indexes.

Standard & Poor's 500 index includes leading US companies and captures approximately 80% of available US market capitalisation. Those companies are trading heavily with the rest of the world, including Asia, and notably Japan; and with South America, notably Brazil. Moreover, Brazil's economy heavily affects the MSCI Emerging Markets Index. According to the Foreign Trade figures from the United States Census Bureau, within Europe US has a strong commercial partnership with Germany,[4] much stronger than with the second commercial ally, namely UK.[5] Therefore, it is straightforward to see how the return for Standard & Poor's has a significant statistical dependence with the German Stock Market. Moreover, with 15% of the imports coming from Germany, the UK economy is also significantly dependent on the German market[6] (instead Germany imports mostly from the Netherlands and exports mostly to the US).[7] Finally, the MSCI European Index return is heavily affected by the first economy in the European Union, namely Germany.

We also used this dataset of 536 entries to evaluate our approach using different amount of observed data. Table 4 lists actual and predicted RMSE for our approach and benchmark approaches when different amount of observations are used. It shows that SBN can achieve the best error rate even with trained with only 10 days of observations and the error drops to around 0.09 when 30 days of data is used.

[4] `https://goo.gl/8PdBl1` (on 4th May 2017)
[5] `https://goo.gl/n2V89z` (on 4th May 2017)
[6] `https://goo.gl/v1tXD4` (on 4th May 2017)
[7] `https://goo.gl/ZPJLdR` (on 4th May 2017)

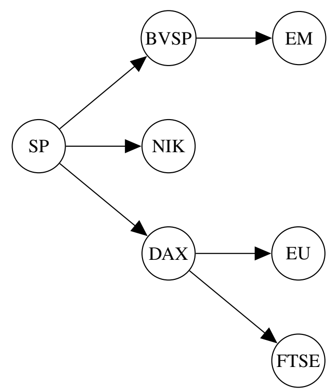


Figure 6: Istanbul Stock Exchange Data Set [Akbilgic *et al.*, 2014] dependency network.

$N_{train} = 10$ (sample size 1.86%)

	SBN	**Credal**	**Belief Networks**
Actual RMSE	**0.131**	0.170	0.172
Predicted RMSE	**0.146**	0.223	0.124

$N_{train} = 30$ (sample size 5.60%)

	SBN	**Credal**	**Belief Networks**
Actual RMSE	**0.088**	0.089	0.104
Predicted RMSE	**0.093**	0.140	0.068

Table 4: Error for the Istanbul stock exchange dataset. Gold standard trained with $N_{train} = 536$. Best results in bold.

Figure 5 demonstrates our results in terms of γ-confidence intervals. Even for data of 10 days, the confidence for inferences with SBN only slightly diverges from the desired confidence levels. When training data is increased to 30 days, the confidence interval for SBN approximate the desired one very closely. Again, in this dataset, the best performance belongs to SBN in terms of γ-confidence intervals.

5 Conclusion

In this paper we presented a methodology to exploit human-machine coalitions for situational understanding, i.e. the ability to relate relevant information and form logical conclusions as well as identifying gaps in information. This process requires the ability to reason inductively, for which we will exploit the machines' ability to learn from data, although important phenomena are often rare in occurrence, severely limiting the availability of instance data and hence the applicability of many machine learning approaches.

To this end, we discussed at length the benefits of Subjective Logic Bayesian Networks, especially their need for less data, and in Section 4 we proved they are superior to previous methods to reason over uncertain probabilistic networks, Credal networks and Belief Networks. We considered two different data-sets, both related to the financial domain, and the experimental results are remarkably similar. In future work we will address graph structures other than trees.

Moreover, we also discussed the role that would be played by humans in situational understanding. Differently from other approaches aimed at explaining high-dimensional, multivariate feature spaces and dependencies to humans, e.g. [Letham *et al.*, 2015; Timmer *et al.*, 2017], we believe a conversational interface like the one depicted in Figure 2 can provide the right level of interactivity in the coalition of humans and machines for situational understanding, and we are currently working towards an implementation of it to support future experiments in this space.

This opens a large spectrum of future work, including the ability to evaluate the human expertise and the quality of data. If a human user adds a dependency that is not supported by data, it might suggest that the user is not correct in their assertion, or that the data is either biased or corrupted.

References

[Akbilgic *et al.*, 2014] Oguz Akbilgic, Hamparsum Bozdogan, and M. Erdal Balaban. A novel Hybrid RBF Neural Networks model as a forecaster. *Statistics and Computing*, 24(3):365–375, 2014.

[Brannon *et al.*, 2009] N.G. Brannon, J.E. Seiffertt, T.J. Draelos, and D.C. Wunsch II. Coordinated machine learning and decision support for situation awareness. *Neural Networks*, 22(3):316 – 325, 2009.

[Dhami *et al.*, 2015] Mandeep K. Dhami, David R. Mandel, Barbara A. Mellers, and Philip E. Tetlock. Improving intelligence analysis with decision science. *Perspectives on Psychological Science*, 10(6):753–757, 2015.

[Dostal, 2007] Brad C. Dostal. Enhancing situational understanding through employment of unmanned aerial vehicle. *Army Transformation Taking Shape: Interim Brigade Combat Team Newsletter*, 01-18, 2007.

[Ivanovska *et al.*, 2015] Magdalena Ivanovska, Audun Jøsang, Lance Kaplan, and Francesco Sambo. Subjective networks: Perspectives and challenges. In *Proc. of the 4th International Workshop on Graph Structures for Knowledge Representation and Reasoning*, pages 107–124, Buenos Aires, Argentina, 2015.

[Jøsang, 2016] Audun Jøsang. *Subjective Logic: A Formalism for Reasoning Under Uncertainty*. Springer, 2016.

[Kaplan and Ivanovska, 2016] Lance Kaplan and Magdalena Ivanovska. Efficient subjective Bayesian network belief propagation for trees. In *ISIF/IEEE Intl. Conf. on Information Fusion (FUSION)*, pages 1300–1307, 2016.

[Kaplan and Ivanovska, 2017] Lance Kaplan and Magdalena Ivanovska. Efficient subjective bayesian network belief propagation for singly-connected graphs. Submitted to *International Journal of Approximate Reasoning*, 2017.

[Kaplan *et al.*, 2015] Lance Kaplan, Murat Şensoy, Supriyo Chakraborty, and Geeth de Mel. Partial observable update for subjective logic and its application for trust estimation. *Information Fusion*, 26:66–83, 2015.

[Karlsson *et al.*, 2008] Alexander Karlsson, Ronnie Johansson, and Sten F Andler. An empirical comparison of Bayesian and credal networks for dependable high-level information fusion. In *Intl. Conf. on Information Fusion (FUSION)*, pages 1–8, 2008.

[LeCun *et al.*, 2015] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[Lerner and Malka, 2011] Boaz Lerner and Roy Malka. Investigation of the k2 algorithm in learning bayesian network classifiers. *Applied Artificial Intelligence*, 25(1):74–96, 2011.

[Letham *et al.*, 2015] Benjamin Letham, Cynthia Rudin, Tyler H. McCormick, and David Madigan. Interpretable classifiers using rules and Bayesian analysis: Building a better stroke prediction model. *The Annals of Applied Statistics*, 9(3):1350–1371, sep 2015.

[Pearl, 1986] Judea Pearl. Fusion, propagation, and structuring in belief networks. *Artificial Intelligence*, 29(3):241–288, 1986.

[Smets, 1993] P. Smets. Belief functions: The disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning*, 9:1– 35, 1993.

[Timmer *et al.*, 2017] Sjoerd T. Timmer, John-Jules Ch. Meyer, Henry Prakken, Silja Renooij, and Bart Verheij. A two-phase method for extracting explanatory arguments from Bayesian networks. *International Journal of Approximate Reasoning*, 80(C):475–494, 2017.

[Zaffalon and Fagiuoli, 1998] M. Zaffalon and E. Fagiuoli. 2U: An exact interval propagation algorithm for polytrees with binary variables. *Artificial Intelligence*, 106(1):77–107, 1998.