Decisionmaking with Fuzzy-Belief-State-Based Reasoning

Lily R. Liang
Dept. of Computer Science
University of the District of Columbia
Washington, DC 20008 USA
LLiang@udc.edu

Carl G. Looney
Dept. of Computer Science & Engineering
University of Nevada, Reno
Reno, NV 89557 USA
looney@cs.unr.edu

Abstract

An outstanding problem is how to make decisions with uncertain and incomplete data from disparate sources without NP-hard algorithms. Here we introduce a new reasoning methodology, fuzzy-belief-state-based reasoning, to solve this problem. In this methodology, we first create a fuzzy-belief-state base for a system from its historical data. For any component n (n = 1,...,N) of the set of empirical state vectors, the values of that component are clustered into LOW, MEDIUM and HIGH fuzzy sets. Then each state vector is fuzzified into a fuzzy-belief-state vector of N triples, where the n-th triple contains the fuzzy truths of membership of the variable value in these respective three fuzzy sets. Each such vector of N triples is associated with a decision to form a fuzzy-belief-case and such cases comprise a fuzzy-belief-state base. Then, when given an observed state vector that is incomplete and uncertain, we mine fuzzy association rules from the fuzzy-belief-state base and apply them to infer the missing values and their fuzzy beliefs based on that incomplete observation. The completed observation is used to match fuzzy-belief-state vectors in the fuzzy-belief-state base. Decisions of the best matching cases are retrieved for use.

1. Introduction

Decisions underlie any action that a problem solver may take in structuring problems, reasoning about situations, allocating resources, retrieving and displaying information, or in controlling physical, organizational or political activities (see [27]). Decisions are needed, e.g., in pattern recognition and classification, diagnosis, prognosis, product design, marketing and military strategy, public policy, routing, scheduling and negotiations.

Given a situation where one of multiple competing choices, classes, actions, strategies, targets, etc. (called *decisions*), is to be selected, some considerations are: i) the completeness of the data upon which the decisionmaking is based; ii) the level of uncertainty of this data; iii) the consistency of the data that may be from disparate sources; iv) the degree of belief in the decision; and vi) the optimal cost of the decision. The expense of obtaining complete and essentially certain data makes it necessary and important to make decisions based on uncertain and incomplete data.

Each of the existing tools for decisionmaking under such conditions is either NP-hard, its parameters are very difficult and costly to obtain, or it lacks the flexibility of completing missing data. *Bayesian belief networks* (BBNs) [12, 24], *fuzzy belief networks* (FBNs) [17] and *fuzzy belief Petri nets* (FBPNs) [15] (also see [20]) can accommodate uncertainty, incompleteness and data from multiple and disparate sources. However, BBNs are NP-hard [6, 7] and the necessary conditional probabilities are difficult or impossible to obtain. But all types of these *influence networks*, which originated in 1921 [28] require knowledge of the structure of the influence network. Many BBN algorithms exist [22, 29, 30].

Other methodologies use beliefs of some kind but do not work without one or more of certainty, completeness and knowledge of the relationships between disparate sources. *Rule-based systems* [2, 4, 10, 11] attempt to deal with uncertainty by means of *certainty factors*, but these are neither axiomatic nor consistent. Completeness is required by the *theory of evidence* methodology [8, 26], *operations research (*see [21] for a connection to BBNs), *mathematical models* and *data mining* [5, 9, 14], which associates disparate data in tabular columns. *Fuzzy logic* methods require the completeness of the input data, as do *genetic algorithms*. *Rough sets* [23] require certainty as well. *Neural networks (see* [19]) generally require complete data as do pattern recognition and *classification* [18]. Case-based reasoning (CBR) is powerful [1, 13, 16, 25] but requires certainty and incompleteness degrades its performance.

In this paper we introduce a new reasoning methodology that we call *fuzzy-belief-state-based reasoning*, to solve this problem. Section 2 introduces how a fuzzy-belief-state base is created and Section 3 introduces our new fuzzy-belief-state-based reasoning approach. We applied this approach on real world data and show the results in Section 4. Section 5 discusses conclusions and future directions.

2. Fuzzy-belief-state base creation

Situations for a given system can be described by *state vectors* $\boldsymbol{s} = (s_1, s_2,...,s_N)$, where s_n (n = 1,...,N) is a state variable or parameter of the system. The state can vary over time as is indicated by $\boldsymbol{s}(t)$. A basic assumption here, as in other methodologies, is that the system is stationary rather than evolutionary, that is, the system at different times has the same relationships among the variables (the correlations are constant over time). For a set of Q state vectors $\mathbf{s}(t_1),...,\mathbf{s}(t_Q)$, we consider the set of all Q values $S_n = \{s_n(t_q): q \in Q\}$ for a fixed component n. Considering the Q vectors as Q rows, S_n is the set of all values in the nth column.

For each n = 1,...,N, we cluster the Q values in S_n into three clusters and find their Gaussian weighted centers [18]. This process weights the prototype (center) of each cluster according to the density of the surrounding points and locates the center among the most densely situated points. These three clusters of the S_n values are designated as the fuzzy sets LOW_n (L_n), $MEDIUM_n$ (M_n) and $HIGH_n$ (H_n) with fuzzy set membership functions centered on their centers. Figure 1 shows the components n across the horizontal axis, the range of values and cluster centers on the vertical axis and the fuzzy set membership function values coming out of the plane in the third dimension.

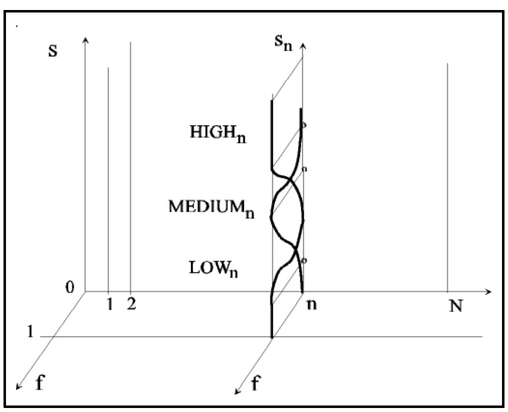

Figure 1. Fuzzy set membership functions of component n.

The fuzzy set membership functions (FSMFs) are Gaussians centered on the cluster centers except for the LOW and HIGH end functions where half Gaussians are used with the other half being a constant of unity. The FSMFs for L_n, M_n and H_n are

$$f_{Ln}(s_n) = \exp[-(s_n-\mu_{Ln})^2/(2\sigma_{Ln}^2)],\ s_n > \mu_{Ln},\ \text{else}\ f_{Ln}(s_n)=1 \quad (1)$$

$$f_{Mn}(s_n) = \exp[-(s_n-\mu_{Mn})^2/(2\sigma_{Mn}^2)] \quad (2)$$

$$f_{Hn}(s_n) = \exp[-(s_n-\mu_{Hn})^2/(2\sigma_{Hn}^2)],\ s_n < \mu_{Hn},\ \text{else}\ f_{Hn}(s_n)=1 \quad (3)$$

where, e.g., f_{Ln} is the fuzzy truth that component value s_n belongs to fuzzy set L_n, while μ_{Ln} and σ_{Ln} are the center and the standard deviation for the FSMF for L_n. Thus, via fuzzification of each component value $s_n(t)$, a state vector $\mathbf{s}(t) = (s_1(t), s_2(t),...,s_N(t))$ is mapped into a *fuzzy-belief-state* (FBS) vector of N triples as shown in Equation (4).

$$\mathbf{B}(t) = (f_{L1},f_{M1},f_{H1};\ f_{L2},f_{M2},f_{H2};...;\ f_{LN},\ f_{MN},\ f_{HN}) \quad (4)$$

For any s_n, there are two non-negligible fuzzy truths of membership (called fuzzy beliefs) respectively for two of the three fuzzy sets L_n, M_n and H_n. These two sufficiently describes the fuzzy set memberships of s_n. The lowest fuzzy truth is set to 0.

Table 1 displays a tabular base of Q fuzzy-belief-state vectors $\{\mathbf{B}(t)\}$ with each having a decision adjoined as an extra component. Such decisions are to have been previously selected so as to be optimal in some sense. The decision column can actually be taken to be any column in the state vector or a new component, but either way it represents the output of the decisionmaking process. Such a table is called a *fuzzy-belief-state base* (FBSB). A row (record) in the FBSB is a *fuzzy-belief-state case*.

Table 1. A fuzzy-belief-state base

	s_1			. . .	s_N			
t	f_{L1}	f_{M1}	f_{H1}	. . .	f_{LN}	f_{MN}	f_{HN}	**Decision**
t_1	0.0	0.5	0.8	. . .	0.7	0.5	0.0	Decision d_1
t_2	0.3	0.7	0.0	. . .	0.0	0.5	0.9	Decision d_2
..	. . .			. . .	. . .			. . .
t_Q	0.6	0.5	0.0	. . .	0.8	0.4	0.0	Decision d_Q

3. Fuzzy-belief-state-based reasoning

In the real world it is often true that not all of the variables in a state can be observed at a given time t. In this case we say the observation data is *incomplete* (due to missing component values). Also, the present measurements of the observed variables contain noise (uncertainty) and come from disparate sources such as sensors, reports, archived data, etc. From an uncertain and incomplete observed input state vector we propose to use the knowledge that exists within the data in the fuzzy-belief-state base to complete the observation, account for the uncertainty with beliefs, and perform a type of reasoning to select a decision as a response to that input.

The high level algorithm. The *fuzzy-belief-state-based reasoning* (FBSBR) approach assumes no functional relationship between the output decisions and the inputs, but uses only knowledge in the FBSB data. Given the incomplete and uncertain observation state vector $\boldsymbol{s}^{\sim}(t)$, we

1) *Fuzzification*: fuzzify the variables that are present to obtain an incomplete FBS vector $\mathbf{B}^{\sim}(t)$

2) *Data Association*: perform fuzzy association rule mining

3) *Belief Inferencing*: infer beliefs to get the complete the FBS vector $\mathbf{B}(t)$

4) *Decision Retrieval*: retrieve decisions of FBS cases from the FBSB that best match $\mathbf{B}(t)$

5) *Decision Adjustment*: adjust retrieved decisions and use

Figure 2 shows the above steps that make up fuzzy-belief-state-based reasoning. In the rest of this section we introduce these steps one by one.

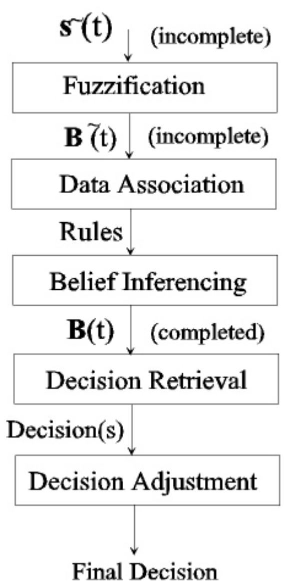

Figure 2. FBSBR flowchart.

Fuzzification. For a given incomplete observation $\boldsymbol{s}^{\sim}(t)$, we first fuzzify it to obtain an incomplete fuzzy-belief-state vector $\mathbf{B}^{\sim}(t)$. The fuzzification process is done for each present value $s_k(t)$ by obtaining its fuzzy truths f_{Lk}, f_{Mk} and f_{Hk} by putting it through the three FSMFs L_k, M_k and H_k. This step is similar to the fuzzification step in FBSB creation, which we introduced in Section 2. The lowest one of f_{Lk}, f_{Mk} and f_{Hk} is set to 0.

Data association. In this step, we implement our fuzzy association rule mining from the FBSB. Let X be a statement such as s_k *is* L_k and Y be a statement such as s_j *is* M_j. The *support of X* in the FBSB is the set of all cases in that FBSB where X is true. The *support of a rule* of the form X ⇒ Y is the set of all cases in the FBSB where both X and Y are true, i.e., the cases that support X∪Y. The *rule confidence* for X ⇒ Y is the ratio of the number of cases that support X∪Y to the number of cases that support X (sometimes expressed as a percentage). In our situation here, X will be a pair of statements where each has a nonzero fuzzy truth, and Y will be a single statement with its fuzzy belief. Thus X ⇒ Y will have a form such as

$$s_k \text{ is } L_k\ (f_{Lk})\ ;\ s_k \text{ is } M_k\ (f_{Mk}) \ \Rightarrow s_j \text{ is } M_j\ (f_{Mj}) \quad (5)$$

For each s_k that is present in the observation, we create rules of the above format with the pair of statements that are true for s_k as X, and a statement of each missing $s_j(t)$ being either L_j, M_j or H_j as the statement in Y. To determine the fuzzy truth of s_j in a rule, we examine all cases in the FBSB that support the statements of that rule and average their fuzzy beliefs for s_j.

The rule confidence is computed for each rule. The rules and their confidences are to be used in the fuzzy-belief inferencing, which is the next step.

Fuzzy-belief inferencing. We now do the fuzzy belief inferencing to estimate each missing value $s_j(t)$. All fuzzy rules obtained for $s_j(t)$ are applied and each rule votes for its consequent statement Y with its confidence. Letting F represent any of L_j, M_j and H_j, the fuzzy belief (f_F) that $s_j(t)$ belongs to F is

$$f_F = \frac{\sum_{(r=1,R(F))} C_r f_{Fr}}{\sum_{(r=1,R(F))} C_r} \quad (6)$$

where

R(F) = number of rules with statement Y = *($s_j(t)$ is F)*
f_{Fr} is the f_F in Rule r
C_r is the confidence of Rule r

Only the two highest f_F values for $s_j(t)$ are kept and the other one is set to 0. Thus we account for the missing datum s_j in $\mathbf{s}$ with a triple of fuzzy beliefs for $\mathbf{B}^{\sim}(t)$. Each missing component of $\mathbf{s}$ is accounted for in this way to complete the fuzzy-belief-state vector $\mathbf{B}(t)$.

Decision retrieval. Next, we measure the similarity of the completed $\mathbf{B}(t)$ to each FBS vector in the FBSB. The similarity measurement λ_q is determined by

$$\lambda_q = \sum_{(n=1,N)} \alpha_{qn} (1 - \delta_{qn}) \quad (7)$$

where

λ_q is the similarity of the q-th vector in the FBSB to **B**(t)

α_{qn} is the vote of template matching of the q-th vector to **B**(t) on the n-th dimension, when either both of non-zero L_n and M_n, or both of non-zero H_n and M_n, then $\alpha_{qn} = 1$, otherwise

$\alpha_{qn} = 0$

δ_{qn} is the square root of the mean square error between the beliefs of the q-th vector and **B**(t) on dimension n

We then retrieve the decisions of the cases with the maximum similarity λ_q and use either one decision or interpolated decisions.

4. Computer runs on real world data

We test the methodology on real world noisy and incomplete data. The Wisconsin breast-cancer database was originally provided by Dr. William H. Woldberg [3] and used by a number of researchers in pattern recognition. This database contains 699 cases, each of which is described by ten attributes in addition to the unique identification (code) number. Attribute 10 is the class decision here. The attributes are listed in table 2:

Table 2. Wisconsin breast-cancer data attributes

# Attribute	Domain
0. Sample code number	id number
1. Clump Thickness	1 - 10
2. Uniformity of Cell Size	1 - 10
3. Uniformity of Cell Shape	1 - 10
4. Marginal Adhesion	1 - 10
5. Single Epithelial Cell Size	1 - 10
6. Bare Nuclei	1 - 10
7. Bland Chromatin	1 - 10
8. Normal Nucleoli	1 - 10
9. Mitosis	1 - 10
10. Class:	(2 for benign, 4 for malignant)

Experiments with existing incomplete data. There are 16 records in the Wisconsin breast-cancer database with the value of Feature 6 missing. We use these 16 records as incomplete observations with their class label masked and use the rest of the data set for creating a FBSB. The FBSB is created and FBSBR is applied to make decisions for these observations as to their membership in the disjoint classes *benign* or *malignant*. Decisions were compared against the masked labels and accuracy was computed. The results are shown in Table 3. As can be seen, 14 out of the 16 cases were labeled correctly by the FBSBR. Cases 1096800 and 704168, which are highlighted in Table 3, were mislabeled to *malignant*. The accuracy for this test is 14/16, or 87.5%.

Table 3. Result with records of missing Feature 7

Sample code #	Label by FBSBR	Original Label
1057013	malignant	malignant
1096800	malignant	benign
1183246	benign	benign
1184840	benign	benign
1193683	benign	benign
1197510	benign	benign
1241232	benign	benign
169356	benign	benign
432809	benign	benign
563649	malignant	malignant
606140	benign	benign
61634	benign	benign
704168	malignant	benign
733639	benign	benign
1238464	benign	benign
1057067	benign	benign

Experiments with created incomplete data. We also did tests on the Wisconsin breast-cancer with the missing data taken out. We divided the 683 complete cases into 10 portions. Each time we drew one portion out of the ten to use as an observation set and used the rest for creating a small fuzzy-belief-state base.

Table 4. Accuracy with created incomplete data (missing one attribute each time)

	"Missing" Attribute								
Dataset	1	2	3	4	5	6	7	8	9
1st	81	81	81	85	85	87	84	78	82
2nd	97	97	94	96	99	91	99	97	97
3th	97	97	99	99	97	94	99	97	99
4th	88	96	91	94	90	90	91	91	91
5th	88	97	96	97	97	91	97	96	97
6th	99	99	97	99	99	97	99	99	97
6th	99	99	97	99	99	97	99	99	97
7th	96	96	94	94	96	96	96	96	96
8th	97	99	97	96	97	99	97	96	97
9th	100	100	100	100	100	100	97	100	100
10th	100	100	100	100	100	100	100	100	100

For the observation set, we deliberately masked their labels, and also cleared their values of one or more features to simulate real world incomplete observations. Then the FBSB was created and FBSBR was applied to infer about the missing data and to make decisions for these observations. The decisions were compared against the masked labels and the accuracy was computed. We repeated this for each different observation set. Tables 4 and 5 show the results.

Table 5. Accuracy with created incomplete data (missing three attributes each time)

	"Missing"Attributes		
Dataset	1, 4, 9	2, 5, 7	3, 6, 8
1st	82	84	91
2nd	96	97	90
3th	97	97	94
4th	93	91	90
5th	91	96	87
6th	99	97	96
7th	96	96	96
8th	99	97	100
9th	100	100	100
10 th	100	100	99

5. Conclusions and future work.

A new methodology, fuzzy-belief-state-based reasoning, is introduced for making decisions with uncertain and incomplete data from disparate sources. In this methodology, a FBSB is created from historical data of a system by clustering and fuzzification. When given observed incomplete and uncertain data, the FBSB is then mined for fuzzy association rules to infer about the missing data. After that, each case in it is matched against the inference-completed observation to retrieve decision(s) of the best match(es). The test results on real world data prove the effectiveness of this methodology.

FBSBR leverages the strength of fuzzy clustering, belief inferencing and case-based reasoning to provide an innovative and intuitive way of reasoning under uncertainty and incompleteness. Compared with the existing Bayesian network models of belief inferencing, FBSBR has the following advantages

1) ***it is not NP-hard.*** The fuzzy association rules are mined after an incomplete observation is made. Only the association rules for the observed ones and the unobserved ones are mined in a one-to-one manner. The time complexity for the rule mining is 3N.
2) ***it is flexible.*** This post-observation mining not only prunes the mining space, but also makes FBSBR much more flexible than any other model that is predefined before observation.
3) ***it avoids the graphical structure problem.*** FBSBR reasons at the data level, instead of using graphical representations for events or associations. This avoids the big graphical representation problem, which inhibits the utilization of existing network models.
4) ***it is self-expandable.*** As new cases are added and new rules are found and saved, the FBSBR system grows and learns.

Compared with the case-based reasoning (CBR) technique, FBSBR handles problems that CBR cannot handle without degrading the performance for retrieval of cases, i.e., FBSBR reasons under incompleteness.

Future work includes the investigation of evolutionary processes whose interrelationships change over time. It will also include searching for more complex rules. Future applications will include diagnosis, tracking, and threat analyses, for examples.

References

[1] A. Aamodt and E. Platz, Case-based reasoning: foundational issues, methodological variations and system approaches, IEEE AI Communications, 7(i), 1994, 39-59.

[2] A. Barr and E. A. Feigenbaum, *Handbook of Artificial Intelligence*, Morgan Kaufmann, Los Altos, CA, 1981.

[3] K. P. Bennett & O. L. Mangasarian: "Robust linear programming discrimination of two linearly inseparable sets", *Optimization Methods and Software* 1, 1992, 23-34

[4] B. G. Buchanan and E. H. Shortliffe, Rule-based Expert Systems: the MYCIN Experiment of the Stanford Heuristic Programming Project, Addison-Wesley, Reading, MA, 1984.

[5] K.C.C. Chan and W.-H. Au, "Mining Fuzzy Association Rules," in *Proc. of the 6th Int'l Conf. on Information and Knowledge Management*, Las Vegas, Nevada, 1997, pp. 209-215.

[6] G. F. Cooper, *Probabilistic Inference Using Belief Networks is NP-Hard*, Technical Report KSL-87-27, Medical Computer Science Group, Stanford University, 1987.

[7] P. Dagum and M. Luby, "Approximating probabilistic inference in Bayesian belief networks is NP-hard," *Artificial Intelligence* **60**(1), 1993, 141-153.

[8] A. P. Dempster, "A generalization of Bayesian inference," *J. Royal Statistical Soc*. **30**, 1968, 205-247.

[9] A. Fu, M.H. Wong, S.C. Sze, W.C. Wong, W.L. Wong, W.K. Yu, "Finding Fuzzy Sets for the Mining of Fuzzy Association Rules for Numerical Attributes", IDEAL 98, 1st International Symposium on Intelligent Data Engineering and Learning, Pages 263-268, Hong Kong, October, 1998.

[10] D.E. Heckerman, Probabilistic interpretations for MYCIN's certainty factors, in: L.N. Kanal and J.F. Lemmer (Eds.), *Uncertainty in Artificial Intelligence 3* , North-Holland, Amsterdam, 1986, 167-196.

[11] Eric J. Horvitz, John S. Breese, "Decision Theory in Expert Systems and Artificial Intelligence (Special Issue on Uncertainty in Artificial Intelligence)", *International Journal of Approximate Reasoning*, **2**, 1988, 247-302.

[12] Fin V. Jensen, *Bayesian Networks and Decision Graphs*,

Springer, New York, 2001.

[13] J. L. Kolodner, “Maintaining Organization in a Dynamic Long-Term Memory,” *Cognitive Science* **7**(4), 1983, 243-280.

[14] C.M. Kuok, A. Fu, M.H. Wong, "Mining Fuzzy Association Rules in Databases" *ACM SIGMOD Record,* vol. **27**(1), March 1998, 41-46.

[15] Lily Liang and Carl Looney, “Inference via fuzzy belief Petri nets,” *Proc. IEEE ACTAI 2003 Conf.*, Sacramento, 2003, 510-514.

[16] Carl G. Looney and Lily Rui Liang, “Cognitive Situation and Threat Assessment of Ground Battlespaces”, *Int. J. Information Fusion*, **4**(4), 2003, 297-308.

[17] Carl Looney and Lily Liang, “Inference via Fuzzy Belief Networks” *Proc. ISCA International Conference*, San Diego, Nov. 2002.

[18] Carl G. Looney, “Interactive clustering and merging with a new fuzzy expected value,” *Pattern Recognition* 35, 2002, 2413-2423.

[19] Carl G. Looney, *Pattern Recognition Using Neural Networks*, Oxford University Press, NY, 1997.

[20] Carl Looney, "Fuzzy Petri Nets for Rule Based Decisionmaking," *IEEE Trans. SMC*, **18**(1), Jan. 1988.

[21] Sucheta Nadkarni and Prakash Shenoy, “A Bayseian network approach to making inferences in causal maps,” *European J. Operations Research* **128**, 2001, 479-498.

[22] A. E. Nicholson and J. M. Brady, “Dynamic belief networks for discrete monitoring,” *IEEE Trans. SMC* **24**(11), 1994, 1593-1610.

[23] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic, Dordrecht, 1991.

[24] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, San Mateo, California, 1988.

[25] R. Schank, *Dynamic memory: a theory of reminding and learning in computers and people*, Cambridge University Press, Cambridge, UK, 1982.

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

[27] Glenn Shafter and Judea Pearl, Editors, *Readings in Uncertain Reasoning*, Morgan Kaufman, San Mateo,1990.

[28] S. Wright, “Correlation and causation,” *J. Agricultural Research* **20**, 1921, 557-585.

[29] Y. Xiang, “Belief updating in multiply sectioned Bayesian networks without repeated local propagations,” *Int. J. Approximate Reasoning* **23**, 2000, 1-21.

[30] Nevin L. Zhang, “Computational properties of two exact algorithms for Bayesian networks,” *Applied Intelligence* **9**, 1998, 173-183.