Hybrid Propagation in Junction Trees

A. Philip Dawid[1], Uffe Kjærulff[2], Steffen L. Lauritzen[2]

[1] University College London, Gower Street, London WC1E 6BT, England
[2] Aalborg University, Fredrik Bajers Vej 7E, DK-9220 Aalborg, Denmark

Abstract. We introduce a methodology for performing approximate computations in complex probabilistic expert systems, when some components can be handled exactly and others require approximation or simulation. This is illustrated by means of a modified version of the familiar 'chest-clinic' problem.

1 Introduction

Markov Chain Monte-Carlo (MCMC) methods have over the last decade become increasingly popular as a computational tool in complex stochastic systems of various type (Gelfand and Smith 1990, Thomas et al. 1992, Gelman and Rubin 1992, Geyer 1992, Smith and Roberts 1993). The methods are flexible, easy to implement, and computing time tends to scale manageably with the size of the problem under consideration. Their main disadvantages are associated with their reliance on pseudorandom number generators (Ripley 1987), some difficulty in deciding whether the necessary equilibrium has been reached, as well as the fact that even moderately sized problems compute slowly.

Alternatively, fast and exact methods have been developed for computation in particular types of graphical models (Lauritzen and Spiegelhalter 1988, Shenoy and Shafer 1990, Jensen et al. 1990, Dawid 1992, Lauritzen 1992, Spiegelhalter et al. 1993) and implemented, for example, in the shell HUGIN (Andersen et al. 1989). Whenever exact methods apply, they are usually preferable, but the weakness of exact methods is precisely the limitations on the problems that they are currently able to deal with appropriately. Even though the development of exact methods has not stopped nor found its limit, there will always be interesting problems that cannot be appropriately dealt with by means of an exact analysis.

Currently the situation is that if just a single variable in a Bayesian network has a continuous distribution, then the exact analysis cannot be performed and the entire network must be handled with Monte-Carlo methods – except in special cases (Lauritzen 1992).

This paper is concerned with establishing the basic ideas behind a scheme that enables the combination of virtues of the two types of methods. The idea is to split up a graphical model into submodels (universes) that each use their favourite method of computation, organise the universes in a junction tree as also done in the case of exact methods, and letting them communicate with each other. The hope is that this will produce a scheme which has the flexibility of the

MCMC methods but exploits speed and correctness of exact methods whenever feasible. In the research horizon, more general schemes involving combination with belief functions and objects from other uncertainty formalisms appear as interesting possibilities.

2 A Basic Example

To illustrate our scheme we have chosen a modified version of the chest-clinic problem (Lauritzen and Spiegelhalter 1988). The directed acyclic graph in Fig. 1a shows the independence graph of the associated model with variables 'Visit to Asia?' (A), 'Smoker?' (S), 'Tuberculosis?' (T), 'Lung cancer?' (L), 'Bronchitis?' (B), 'Either tuberculosis or lung cancer?' (E), 'Positive X-ray?' (X), and 'Dyspnoea?' (D). All variables have states 'yes' and 'no', in the following denoted by 1 and 0, respectively. For convenience, we shall also use the notation $A = a$ for $A = 1$, $A = \overline{a}$ for $A = 0$, etc.

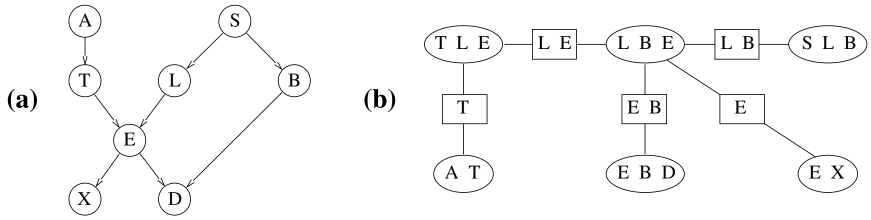

Fig. 1. Independence graph (a) and a junction tree (b) for the original chest-clinic model

The original conditional probabilities associated with the model are reproduced in Table 1.

Table 1. Conditional probabilities for the chest-clinic model

$P(A=1) \quad = 0.01$	$P(E=1 \mid L=1, T=1) = 1$
	$P(E=1 \mid L=1, T=0) = 1$
$P(T=1 \mid A=1) = 0.05$	$P(E=1 \mid L=0, T=1) = 1$
$P(T=1 \mid A=0) = 0.01$	$P(E=1 \mid L=0, T=0) = 0$
$P(S=1) \quad = 0.50$	$P(X=1 \mid E=1) = 0.98$
	$P(X=1 \mid E=0) = 0.05$
$P(L=1 \mid S=1) = 0.10$	
$P(L=1 \mid S=0) = 0.01$	$P(D=1 \mid B=1, E=1) = 0.90$
	$P(D=1 \mid B=1, E=0) = 0.80$
$P(B=1 \mid S=1) = 0.60$	$P(D=1 \mid B=0, E=1) = 0.70$
$P(B=1 \mid S=0) = 0.30$	$P(D=1 \mid B=0, E=0) = 0.10$

A junction tree corresponding to the graph of Fig. 1a is constructed via the operations of moralization (edges between E and B and between T and L) and triangulation (edge between L and B) (Lauritzen and Spiegelhalter 1988) and is shown in Fig. 1b, with the belief universes (cliques) marked by ovals, and the separators by rectangles. The initial (i.e., with no evidence) marginal probabilities for the individual variables can be calculated by appropriate marginalisations of the clique potentials after a propagation has been conducted. The result of these operations appear in Table 2.

Table 2. Marginal probabilities for the individual variables of the chest-clinic model

$P(A) = (0.0100, 0.9900)$	$P(B) = (0.4500, 0.5500)$
$P(T) = (0.0104, 0.9896)$	$P(E) = (0.0648, 0.9352)$
$P(S) = (0.5000, 0.5000)$	$P(X) = (0.1103, 0.8897)$
$P(L) = (0.0550, 0.9450)$	$P(D) = (0.4361, 0.5639)$

The modified chest-clinic model illustrating our hybrid propagation scheme is as follows. First, assume that we want to relax the constraint that bronchitis is either present or absent, and instead be able to express degrees of bronchitis on a continuous scale. We could accomplish this by defining the conditional distribution of acquiring bronchitis given smoking habits by some standard continuous distribution. If a beta distribution (Johnson and Kotz 1970) is found appropriate, we could define

$$B \,|\, S \sim \mathrm{Be}(3 + 3S, 4 + 3(1 - S)) \ , \tag{1}$$

which yields mean values $3/(3 + 7) = 0.3$ and $6/(6 + 4) = 0.6$ for $S = 0$ and $S = 1$, respectively, complying with the original specification of $P(B \,|\, S)$ (see Table 1).

Next, when $B \,|\, S$ has been specified as a beta distribution, we need to respecify the conditional probabilities of suffering from shortness of breath (dyspnoea) given bronchitis and lung cancer or tuberculosis. This could be accomplished through

$$\mathrm{logit}\ P(D = 1 \,|\, B, E) = \alpha B + \beta E + \gamma \ , \tag{2}$$

where $\mathrm{logit}\ x = \log x/(1 - x)$. For $\alpha = 2.5$, $\beta = 1.9$ and $\gamma = -1.65$ we get

$$\begin{aligned}
\mathrm{logit}\ P(D = 1 \,|\, B = 1, E = 1) &= 2.75\ (0.94)\\
\mathrm{logit}\ P(D = 1 \,|\, B = 1, E = 0) &= 0.85\ (0.70)\\
\mathrm{logit}\ P(D = 1 \,|\, B = 0, E = 1) &= 0.25\ (0.56)\\
\mathrm{logit}\ P(D = 1 \,|\, B = 0, E = 0) &= -1.65\ (0.16)
\end{aligned}$$

where the numbers in parentheses are the corresponding conditional probabilities obtained by $\mathrm{logit}^{-1}(2.75) = 0.94$, etc, where $\mathrm{logit}^{-1}(x) = e^x/(1+e^x)$. Note that these probabilities are somewhat different from those of Table 1, but the closest we can get when $P(D \,|\, B, E)$ is given by an expression as (2).

The modification of the model which we have introduced is sufficient to destroy the possibility of computing marginal probabilities at all nodes by exact methods. This example is small enough that direct Monte-Carlo or other numerical methods can be exploited to obtain the marginal probabilities. In this paper we exploit this fact to check our calculations. The example is supposed only to illustrate a scheme that has its practical interest in problems of larger size.

With the above modifications, the mean of B is $E(B) = E(B \mid S = 0)P(S = 0) + E(B \mid S = 1)P(S = 1) = 0.5 \cdot 0.3 + 0.5 \cdot 0.6 = 0.45$. The marginal distribution for D can be calculated to be $(0.4042, 0.5958)$. The last result has been found by numerical integration.

3 Propagation by Message Passing

Here we briefly recall the general idea of propagation by message passing in a junction tree as described by e.g. Dawid (1992). We use notation and terminology from this reference.

Each belief universe (clique) C and separator S in the junction tree is initially equipped with a *potential* such that the joint density satisfies

$$f(x) \propto \frac{\prod_{C \in \mathcal{C}} a_C(x_C)}{\prod_{S \in \mathcal{S}} b_S(x_S)} \ . \tag{3}$$

When evidence has been received, the potentials might have changed in some universes, but an expression of the form (3) continues to be valid.

Messages can now be sent from one universe C to another universe D via their separator S in the following way:

1. C calculates the *marginal* to S, $a_C^{\downarrow S}$, of its potential.
2. The potential on D *receives* the message by changing to

$$a_D^* = a_D \frac{a_C^{\downarrow S}}{b_S} = \frac{a_D}{b_S} a_C^{\downarrow S} \ . \tag{4}$$

3. The separator potential changes to $b_S^* = a_C^{\downarrow S}$.

Two important points should be noted here. Firstly, whenever a message has been sent, the correctness of (3) has not been affected. Secondly, if the messages are scheduled appropriately, then after a finite number of message passes, all messages are neutral, i.e., they do not change anything. We then say that the junction tree is in *equilibrium*. When this happens, marginals to all subsets of a universe calculated from its potential are identical to marginals calculated from the joint expression (3).

The basic idea behind this paper is to consider the potentials and (3) as purely formal objects, allowing the computer to represent these either directly as (discrete) probability tables (as in the current implementation of HUGIN), or as programs such as numerical integrators or Monte-Carlo machines that can draw at random from a distribution with density proportional to a given potential.

For this to be of any practical interest we must also think in terms of the universes being large networks and other more complex objects than just cliques in a graph. If we think of such objects as sets, they are organised in a junction tree of usual type.

The difficulties encountered in such a general approach is concerned with an elaboration of the abstract theory, as well as the specific description of types of universes and the scheduling of messages. The abstract theory is to be discussed in a subsequent paper and the other aspects in the following sections. One aspect of the abstract theory must be briefly mentioned to explain the following developments: The potentials should be replaced by measures rather than functions, and ratios interpreted as Radon-Nikodym derivatives. We avoid the technical details here.

However, we point out that in general it may happen in our scheme that a marginalised measure $\mu^{\downarrow S}$ is not absolutely continuous with respect to the separator measure μ_S. Then the second version of (4) must be used, rather than the first. In the first case, we say that D receives its message *as an update ratio*, and in the second case *as a measure*.

4 Universe Types

This section describes in general terms the main types of universes dealt with in our scheme and how they communicate with each other. We distinguish between the following three main types.

DE (discrete exact) The archetype of a belief universe. A DE-universe is simply given by a list of possible configurations and their associated weights (real numbers). The term 'exact' refers to the fact that it knows to do whatever computation it needs in an exact way.

E (other exact) There are other types of universes that can do exact computations. Examples of such E-universes are
- E-hugin: A universe of discrete variables represented in the HUGIN sense, i.e., by a junction tree of DE-universes.
- E-cghugin: A universe with possibly both discrete and continuous conditional Gaussian variables in the HUGIN sense. DE-universes and E-hugin universes become special cases.
- E-formula: A universe of discrete variables with a belief potential expressed through exact formulae. Computations can for example be handled through some form of computer algebra program.

N (numerical) This is a type of universe that in some way calculates only approximately. An N-universe is typically one of
- N-int: A universe where computations are made by numerical integration; such universes will not be discussed in the present paper, but we find it important to point at the possibility.
- N-mc: This is a universe that makes its calculations through some form of a Monte-Carlo method, for example Gibbs sampling in a moral graph, forward sampling in a directed, acyclic graph, or direct sampling without exploiting any graphical structure. The calculations can involve importance weights.

The different kinds of universes and the way they send and receive messages will be described in detail in Sects. 4.1–4.3.

Returning to our example, we choose to implement AT, TLE and EX as universes of type DE, $SLBE$ as a universe of type N-mc, and EBD as a universe of type E-formula, where we use AT as shorthand for $\{A, T\}$, etc. The universes are connected in the junction tree displayed in Fig. 2.

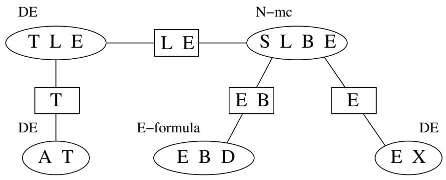

Fig. 2. Junction tree for the modified chest-clinic model with three DE-universes, one N-mc universe, and one E-formula universe

The belief potentials of the universes are instantiated as shown in Table 3, where $\phi(B) \sim U(0,1)$ in universe EBD, $\phi(E) = (1,1)$ in the universes $SLBE$ and EBD, and where $\phi(A) = P(A)$, $\phi(T \mid A) = P(T \mid A)$, etc. No evidence is assumed present. We shall later explain the precise meaning of these potentials in the specific cases.

Table 3. Initial instantiation of the belief potentials of the universes

Univ.	Kind	Initial potential(s)
AT	DE	$\phi(A), \phi(T \mid A)$
TLE	DE	$\phi(E \mid T, L)$
$SLBE$	N-mc	$\phi(S)$, $\phi(B \mid S)$, $\phi(L \mid S)$, $\phi(E)$
EBD	E-formula	$\phi(E)$, $\phi(B)$, $\phi(D \mid E, B)$
EX	DE	$\phi(X \mid E)$

4.1 Discrete Exact Universes

A DE-universe, U, is represented by a finite subset

$$\mathcal{X}'_U = \{x_U^1, \ldots, x_U^n\} \subseteq \mathcal{X}_U = \times_{v \in U} \mathcal{X}_v$$

of n possible states in the product space $\mathcal{X}_U$, and a set of associated weights $w_1, \ldots, w_n$, where w_i is the weight associated with x_U^i. As for other universes, its potential is given by a positive measure μ_U. For a function f, the integral of f with respect to μ_U is thus represented as

$$\int f(x)\mu_U(dx_U) = \sum_{i=1}^{n} w_i f(x_U^i) \ .$$

The universe marginalises to the space

$$\mathcal{X}'_A = \{x_A, x_U \in \mathcal{X}'_U\} \subseteq \mathcal{X}_A = \times_{v \in A} \mathcal{X}_v$$

by just adding the weights of points with the same A-marginal. That is,

$$w_i^{\downarrow A} = \sum_{j: x_A^j = x_A^i} w_j \ .$$

Messages are received as update ratios – messages sent from neighbouring universes as measures are transformed to update ratios upon division by separator measures. When receiving an update ratio ϕ_S, the revised weights are obtained as

$$w_i^* = w_i \cdot \phi_S(x_S^i) \ .$$

DE-universes can always receive and send messages when they are requested to do so.

In our basic example, the AT-universe has states $\{(a,t), (\overline{a},t), (a,\overline{t}), (\overline{a},\overline{t})\}$. The weights (potentials) are obtained by multiplying the appropriate numbers from Table 1. Similarly with the other DE-universes in our example.

4.2 Other Exact Universes

An E-universe, U, has some representation of a probability measure μ_U on $\mathcal{X}_U$ such that $\mu_U^{\downarrow S}$ can be computed in exact form for all separators S associated with U. We assume that if $\mu_U^{\downarrow S}$ is computable, so is $\mu_{U \setminus S \mid S}$. This seems no severe restriction.

As described above, U can be one of sub-types E-hugin, E-cghugin, or E-formula. Any kind of E-universe can send exact messages and receive any message. Upon receiving an update ratio, ϕ_S, the revised measure, μ_U^*, is obtained as

$$\mu_U^* = \mu_U \cdot \phi_S \ , \tag{5}$$

and upon receiving a measure μ_S, μ_U^* is obtained as

$$\mu_U^* = \mu_{U \setminus S \mid S} \cdot \mu_S \ , \tag{6}$$

but U may have to change to some other type such as N-type, since μ_U^* may not allow exact calculations.

In our basic example, the potential on the universe EBD is represented as a table (over DE) of algebraic expressions involving standard functions of B such as displayed in Table 4.

4.3 Numerical Universes

An N-universe, U, is equipped with some representation of a probability measure μ_U defined on $\mathcal{X}_U = \times_{v \in U} \mathcal{X}_v$, but it cannot calculate separator marginals exactly and must therefore make some approximate calculation.

If U is of type N-int, this approximation is obtained via numerical integration, possibly by approximating the entire universe and measure $(\mathcal{X}_U, \mu_U)$ by a discrete DE-universe $\mathcal{X}'_U$ with suitable weights. Thus, when integration has been performed, U may have changed type from N-int to DE.

Table 4. Initial exact potentials for the universe EBD

	$E=1$	$E=0$
$D=1$	$\dfrac{e^{2.5B+.25}}{1+e^{2.5B+.25}}$	$\dfrac{e^{2.5B-1.65}}{1+e^{2.5B-1.65}}$
$D=0$	$\dfrac{1}{1+e^{2.5B+.25}}$	$\dfrac{1}{1+e^{2.5B-1.65}}$

If U is of type N-mc, the approximation is obtained by using a program $\pi(\nu_U, \lambda_U, n)$ which performs forward sampling in a DAG on U drawing from ν_U, Gibbs sampling in an undirected graph on U drawing from ν_U, or Monte-Carlo sampling directly from ν_U without exploiting a graphical structure. Here we assume that $\mu_U = \lambda_U \nu_U$. The program returns a discrete DE-universe $\mathcal{X}'_U$ with points equal to those in the sample. If the function λ_U differs from unity, the samples are weighted according to λ_U. That is, each sample x^i_U is given importance weight $\lambda_U(x^i_U)$. The third argument determines the sample size. Thus, when the program terminates, U has changed type from N-mc to DE.

An N-universe can receive any kind of message (update ratio or measure), and the revised measure is obtained as described in (5) and (6). In any particular instance it must be specified whether the program representing the universe takes in the message by changing ν or by changing the importance weights λ. It may also have to change its Monte-Carlo sampling procedure from forward to Gibbs etc.

Since U becomes a DE-universe before sending, messages from U are sent as described in Sect. 4.1.

In our basic example we have implemented the universe $SLBE$ as an N-mc universe with forward sampling that receives its messages as importance factors. Hence, with the initial potential, it samples by first choosing S uniformly, then L from the distribution of L given S, B from a beta distribution with the correct parameters determined by the sampled value of S, and finally E uniformly. This gives a first point (s, l, b, e) in the approximating space, and the procedure is then repeated n times. If messages have been received from neighbouring universes, the points are weighted appropriately.

Note that this is not in any way an optimal implementation of our network, but it is chosen to illustrate the flexibility of the methods.

5 Network Architecture

In a certain sense the possibilities for scheduling of messages are like in the general case. The novel element is the fact that in the propagation process, universes may change type, in particular they may change into approximating universes, and these can be drastically different from the original ones. This happens for example when some universe with continuous state space changes to a discrete universe. This may destroy the property that the marginal of the universe is absolutely continuous with respect to a separator measure. When it has sent its message, this property will be restored in the separator, but then it may have been destroyed in the separators of the neighbouring universe etc., thus requiring further messages until the property is restored everywhere. We refer to this phenomenon as a *cascade* of messages.

Noting that messages never destroy correctness of the representation, we see that essentially the usual message schedule carries over, with the above modification. We shall illustrate a possible correct scheduling in our example. Recall that a message is termed *active* if it is sent from a universe that has received active messages from all its other neighbours. When active messages have been passed between any pair of universes, equilibrium has been reached. The scheduling in our example is as follows:

1. $AT \to TLE$;	5. $EBD \to SLBE$;
2. $EX \to SLBE$;	6. $SLBE \to TLE$;
3. $TLE \to SLBE$;	7. $TLE \to AT$;
4. $SLBE \to EBD$;	8. $SLBE \to EX$.

All messages are active and equilibrium is therefore reached. Below we indicate in more detail how the messages are sent.

AT → TLE This is a standard message. The potential is marginalised to T by summing the AT-table of numbers over the values of A. This is then placed as the new potential in the separator and is also multiplied onto the initial TLE-table. No division is needed because the separator initially contained a neutral potential.

EX → SLBE This message is almost a standard message. The universe EX sends the potential $a^*_E(e) = \sum_x a_{XE}(x, e)$ to the N-mc universe $SLBE$. The result is stored as an importance factor to be used when $SLBE$ is forward sampling. In our particular case the factor turns out to be unity, but this would not have been so if a chest X-ray had been taken.

TLE → SLBE This is a message of the same type as above. The universe TLE sends the separator potential $a^*_{LE}(l, e) = \sum_t a_{TLE}(t, l, e)$ to the N-mc universe $SLBE$ and this function is multiplied onto the existing importance factor.

SLBE → EBD For this message to be sent, we must first forward sample from $SLBE$ as described in Sect. 4.3, obtaining values $(s_i, l_i, b_i, e_i), i = 1, \ldots, n$ with weights w_i calculated from the importance factor potential.

Since B is a continuous variable, we cannot arrange our samples in a contingency table and we need to store all samples (see Table 5) and send the information pertaining to B and E to universe EBD.

When the sampling has ended, the universe has changed type to DE. It can now send to EBD in the usual way. Thus Table 5 is marginalised to BE by ignoring the S and L values and EBD receives the message by becoming of

Table 5. Sample values for universe $SLBE$

i	s_i	l_i	b_i	e_i	w_i
1	1	0	0.613862	0	0.989600
2	1	0	0.744140	0	0.989600
3	0	0	0.480592	1	0.010400
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n-1$	1	0	0.468109	0	0.989600
n	1	0	0.707468	1	0.010400

type DE with weights obtained from those in Table 5 by multiplication with appropriate values of the functions in Table 4.

When this message has been sent, all universes have become standard DE-universes and messages are sent in the usual way.

The hybrid propagation method given by the scheduling of messages described above has been implemented in the C programming language. Also, a forward sampling procedure has been implemented and run on our example. The marginal distributions on the individual variables obtained by running these two methods with a sample size of 1000 are displayed in Table 6 together with the exact values obtained by direct calculation (variables A, T, S, L, E, and X) and numerical integration (variables B and D). We observe no significant difference between the precision of the distributions provided by forward sampling and those provided by the hybrid propagation method. Since we have implemented the hybrid method in a very inefficient way in this example, this gives good promise for further developments.

Table 6. Exact marginal distributions and approximate ones obtained by forward sampling and hybrid propagation with a sample size of 1000. For variable B, the first number displayed is the mean and the second is the variance.

Variable	Forward	Hybrid	Exact
A	$(0.0110, 0.9890)$	$(0.0100, 0.9900)$	$(0.01, 0.99)$
T	$(0.0140, 0.9860)$	$(0.0107, 0.9893)$	$(0.0104, 0.9896)$
S	$(0.5190, 0.4810)$	$(0.4816, 0.5184)$	$(0.5, 0.5)$
L	$(0.0520, 0.9480)$	$(0.0611, 0.9389)$	$(0.055, 0.945)$
B	$(0.4579, 0.0416)$	$(0.4430, 0.0426)$	$(0.45, 0.04300)$
E	$(0.0660, 0.9340)$	$(0.0712, 0.9288)$	$(0.0648, 0.9352)$
X	$(0.1020, 0.8980)$	$(0.1162, 0.8838)$	$(0.110, 0.890)$
D	$(0.4080, 0.5920)$	$(0.3908, 0.6092)$	$(0.4042, 0.5958)$

Acknowledgements

This research was supported by ESPRIT Basic Research Action through project no. 6156 (DRUMS II) as well as the PIFT programme of the Danish Research

Councils. We are grateful to Ross Curds for providing the numerical integrations using *Mathematica*.

References

Andersen, S. K., Olesen, K. G., Jensen, F. V., Jensen, F.: HUGIN – A shell for building Bayesian belief universes for expert systems, *Proceedings of the 11th International Joint Conference on Artificial Intelligence* (1989) 1080–1085. Also reprinted in Shafer & Pearl (1990)

Dawid, A. P.: Applications of a general propagation algorithm for probabilistic expert systems, *Statistics and Computing* **2** (1992) 25–36

Gelfand, A. E., Smith, A. F. M.: Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association* **85** (1990) 398–409

Gelman, A., Rubin, D. B.: Inference from iterative simulation using single and multiple sequences (with discussion), *Statistical Science* **7** (1992) 457–511

Geyer, C. J.: Practical Markov Chain Monte Carlo (with discussion), *Statistical Science* **7** (1992) 473–511

Jensen, F. V., Lauritzen, S. L., Olesen, K. G.: Bayesian updating in causal probabilistic networks by local computation, *Computational Statistics Quarterly* **4** (1990) 269–282

Johnson, N. L., Kotz, S.: *Distributions in Statistics. Continuous Univariate Distributions*, Vol. 2, Wiley & Sons, New York (1970)

Lauritzen, S. L.: Propagation of probabilities, means and variances in mixed graphical association models, *Journal of the American Statistical Association* **86** (1992) 1098–1108

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

Ripley, B. D.: *Stochastic Simulation*, Wiley & Sons (1987)

Shafer, G. R., Pearl, J. (eds): *Readings in Uncertain Reasoning*, Morgan Kaufmann, San Mateo, California (1990)

Shenoy, P. P., Shafer, G. R.: Axioms for probability and belief-function propagation, *in* R. D. Shachter, T. S. Levitt, L. N. Kanal and J. F. Lemmer (eds), *Uncertainty in Artificial Intelligence IV*, North-Holland, Amsterdam, (1990) 169–198

Smith, A. F. M., Roberts, G. O.: Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods, *Journal of the Royal Statistical Society, Series B* **55**(1) (1993) 5–23

Spiegelhalter, D. J., Dawid, A. P., Lauritzen, S. L., Cowell, R. G.: Bayesian analysis in expert systems (with discussion), *Statistical Science* **8** (1993) 219–247 and 247–283

Thomas, A., Spiegelhalter, D. J., Gilks, W. R.: BUGS: A program to perform Bayesian inference using Gibbs sampling, *in* J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith (eds), *Bayesian Statistics 4*, Clarendon Press, Oxford, UK, (1992) 837–842

This article was processed using the LaTeX macro package with LLNCS style