Appeared in: European Journal of Operational Research, Vol. 186, No. 1, April 1 2008, pp. 261-275.

Decision Making with Hybrid Influence Diagrams Using Mixtures of Truncated Exponentials

Barry R. Cobb [a,*] Prakash P. Shenoy [b]

[a] *Department of Economics and Business, Virginia Military Institute, Lexington, Virginia, 24450, U.S.A*

[b] *School of Business, University of Kansas, 1300 Sunnyside Ave., Summerfield Hall, Lawrence, Kansas, 66045-7585, U.S.A*

Abstract

Mixtures of truncated exponentials (MTE) potentials are an alternative to discretization for representing continuous chance variables in influence diagrams. Also, MTE potentials can be used to approximate utility functions. This paper introduces MTE influence diagrams, which can represent decision problems without restrictions on the relationships between continuous and discrete chance variables, without limitations on the distributions of continuous chance variables, and without limitations on the nature of the utility functions. In MTE influence diagrams, all probability distributions and the joint utility function (or its multiplicative factors) are represented by MTE potentials and decision nodes are assumed to have discrete state spaces. MTE influence diagrams are solved by variable elimination using a fusion algorithm.

Key words: Decision analysis, Influence diagrams, MTE potentials, Applied probability

1 Introduction

An influence diagram is a compact graphical representation for a decision problem under uncertainty. Initially, influence diagrams were proposed as a

* Corresponding author.
Email addresses: cobbbr@vmi.edu (Barry R. Cobb), pshenoy@ku.edu (Prakash P. Shenoy).

front-end for decision trees (Howard and Matheson 1984). Subsequently, Olmsted (1983) and Shachter (1986) developed methods for evaluating an influence diagram directly without converting it to a decision tree. These methods assume that all uncertain variables in the model are represented by discrete probability mass functions (PMF's). Several improvements to solution procedures for solving discrete influence diagrams have been proposed (see, e.g., Shenoy 1992, Shachter and Ndilikilikesha 1993, Jensen *et al.* 1994, Madsen and Jensen 1999, Lauritzen and Nilsson 2001, Madsen and Nilsson 2001).

Shachter and Kenley (1989) introduced Gaussian influence diagrams, which contain continuous variables with Gaussian distributions and a quadratic value function. In this framework, chance nodes have conditional linear Gaussian distributions, meaning each chance variable has a Gaussian distribution whose mean is a linear function of the variable's parents and whose variance is a constant. This framework does not allow discrete chance nodes; however, it does allow chance variables whose distributions are conditionally deterministic linear functions of their parents.

Poland and Shachter (1993) introduce mixture of Gaussians influence diagrams, which allow both discrete and continuous nodes where continuous variables are modeled as mixtures of Gaussians. In this framework, instantiating all discrete nodes reduces the model to a Gaussian influence diagram. The influence diagram must satisfy the condition that discrete chance nodes cannot have continuous parents. In this model, a quadratic value function is specified along with a utility function which represents risk-neutral behavior or a constant risk aversion. Poland (1994) proposes a procedure for solving such influence diagrams which uses discrete and Gaussian operations and reduces continuous chance variables before discrete chance variables. Madsen and Jensen (2005) describe a new procedure for exact evaluation of similar influence diagrams that contain an additively decomposing quadratic utility function.

Monte Carlo methods have also been proposed for solving decision problems with continuous and discrete variables. Bielza *et al.* (1999) uses Markov chain Monte Carlo methods to solve *single stage* problems with continuous decision and chance nodes. Charnes and Shenoy (2004) solve *multiple stage* decision problems using a multi-stage Monte Carlo sampling technique that takes advantage of local computation to limit the number of variables sampled at one time.

Mixtures of truncated exponentials (MTE) potentials are suggested by Moral *et al.* (2001) and Rumí (2003) as an alternative to discretization for solving Bayesian networks with a mixture of discrete and continuous chance variables. In this paper, we propose MTE influence diagrams, which are influence diagrams in which probability distributions and utility functions are represented by MTE potentials. We solve MTE influence diagrams using the fusion algorithm proposed by Shenoy (1993) for the case where the joint utility function decomposes multiplicatively.

The remainder of this paper is organized as follows. Section 2 introduces notation and definitions used throughout the paper. Section 3 defines MTE potentials. Section 4 reviews the operations required for solving an MTE influence diagram. Section 5 presents details of a method for approximating joint utility functions with MTE utility potentials. Section 6 contains two examples, including an adaptation of Raiffa's (1968) Oil Wildcatter problem, which are represented and solved using MTE influence diagrams. Finally, section 7 summarizes and states some directions for future research.

2 Notation and Definitions

This section contains notation and definitions used throughout the paper.

2.1 Notation

Variables will be denoted by capital letters, e.g., A, B, C. Sets of variables will be denoted by boldface capital letters, $\mathbf{Y}$ if all are discrete chance variables, $\mathbf{Z}$ if all are continuous chance variables, $\mathbf{D}$ if all are decision variables, or $\mathbf{X}$ if the components are a mixture of discrete chance, continuous chance, and decision variables. In this paper, all decision variables are assumed to be discrete. If $\mathbf{X}$ is a set of variables, $\mathbf{x}$ is a configuration of specific states of those variables. The discrete, continuous, or mixed state space of $\mathbf{X}$ is denoted by $\Omega_{\mathbf{X}}$.

MTE probability potentials and discrete probability potentials are denoted by lower-case greek letters, e.g., α, β, γ. Discrete probabilities for a specific element of the state space are denoted as an argument to a discrete potential, e.g. $\delta(0) = P(D = 0)$. MTE utility potentials are denoted by u_i, where the subscript i indexes both the initial MTE utility potential(s) specified in the influence diagram and subsequent MTE utility potentials created during the solution procedure.

In graphical representations, decision variables are represented by rectangular nodes, discrete chance variables are represented by single–border ovals, continuous chance variables are represented by double–border ovals, and utility functions are represented by diamonds.

3 Mixtures of Truncated Exponentials

3.1 MTE Potentials

A mixture of truncated exponentials (MTE) potential in an influence diagram has the following definition, which is a modification of the original definition proposed by Moral *et al.* (2001) and Rumí (2003).

MTE potential. Let $\mathbf{X}$ be a mixed n-dimensional variable. Let $\mathbf{Y} = (Y_1, \ldots, Y_f)$, $\mathbf{Z} = (Z_1, \ldots, Z_c)$, and $\mathbf{D} = (D_1, \ldots, D_g)$ be the discrete chance, continuous chance, and decision variable parts of $\mathbf{X}$, respectively, with $c + f + g = n$. A function $\phi : \Omega_{\mathbf{X}} \mapsto \mathcal{R}^+$ is an MTE potential if one of the next two conditions holds:

(1) The potential ϕ can be written as

$$\phi(\mathbf{x}) = \phi(\mathbf{y}, \mathbf{z}, \mathbf{d}) = \\ a_0 + \sum_{i=1}^{m} a_i \exp\left\{ \sum_{j=1}^{f} b_i^{(j)} y_j + \sum_{k=1}^{c} b_i^{(f+k)} z_k + \sum_{\ell=1}^{g} b_i^{(c+f+\ell)} d_\ell \right\} \tag{1}$$

for all $\mathbf{x} \in \Omega_{\mathbf{X}}$, where $a_i, i = 0, \ldots, m$ and $b_i^{(j)}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ are real numbers.

(2) There is a partition $\Omega_1, \ldots, \Omega_k$ of $\Omega_{\mathbf{X}}$ verifying that the domain of continuous chance variables, $\Omega_{\mathbf{Z}}$, is divided into hypercubes, the domain of the discrete chance and decision variables, $\Omega_{\mathbf{Y} \cup \mathbf{D}}$, is divided into arbitrary sets, and such that ϕ is defined as

$$\phi(\mathbf{x}) = \phi_i(\mathbf{x}) \qquad \text{if } \mathbf{x} \in \Omega_i, \tag{2}$$

where each $\phi_i, i = 1, ..., k$ can be written in the form of equation (1) (i.e. each ϕ_i is an MTE potential on Ω_i).

In the definition above, k is the number of *pieces*, and m is the number of exponential *terms* in each piece of the MTE potential. In this paper, all MTE probability and utility potentials are equal to zero in unspecified regions. Several methods for fitting MTE potentials have been proposed (see, e.g., Moral *et al.* 2002, Moral *et al.* 2003, Cobb *et al.* 2006).

A 2-piece, 3-term un-normalized MTE potential which approximates the nor-

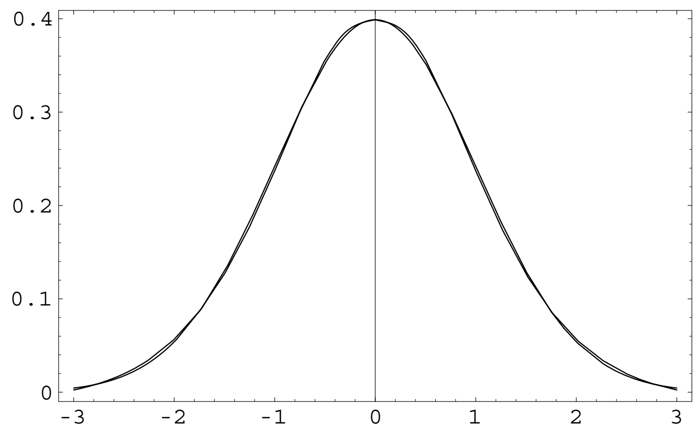

Fig. 1. 2-piece MTE approximation overlayed on the standard normal distribution.

mal PDF defined by Cobb and Shenoy (2006) is

$$\psi'(x) = \begin{cases} \sigma^{-1}(-0.0105643 + 197.0557202\exp\{2.2568434(\frac{x-\mu}{\sigma})\} \\ -461.4392506\exp\{2.3434117(\frac{x-\mu}{\sigma})\} \\ +264.7930371\exp\{2.4043270(\frac{x-\mu}{\sigma})\}) & \text{if } \mu - 3\sigma \le x < \mu \\ \sigma^{-1}(-0.0105643 + 197.0557202\exp\{-2.2568434(\frac{x-\mu}{\sigma})\} \\ -461.4392506\exp\{-2.3434117(\frac{x-\mu}{\sigma})\} \\ +264.7930371\exp\{-2.4043270(\frac{x-\mu}{\sigma})\}) & \text{if } \mu \le z \le \mu + 3\sigma. \end{cases} \tag{3}$$

Figure 1 shows a graph of the 2-piece, 3-term MTE approximation overlayed on the actual normal PDF for the case where $\mu = 0$ and $\sigma^2 = 1$ over the domain $[-3, 3]$. A normalized version of the 2-piece, 3-term MTE approximation to the normal PDF is $\psi(x) = (1/0.9973) \cdot \psi'(x)$.

3.2 MTE Probability Densities

MTE probability densities. Suppose ϕ' is an input MTE potential for $\mathbf{X} = \mathbf{Y} \cup \mathbf{Z} \cup \mathbf{D}$ representing a PDF for $Z \in \mathbf{Z}$ given its parents $\mathbf{X} \setminus \{Z\}$. If we can verify that

$$K(\mathbf{x}) = \int\limits_{\Omega_Z} \phi'(\mathbf{x}, z)\, dz = 1\ , \tag{4}$$

for all $\mathbf{x} \in \Omega_{\mathbf{X}\setminus Z}$, we state that ϕ' is an MTE density for Z. If $K(\mathbf{x}) \neq 1$ for any $\mathbf{x} \in \Omega_{\mathbf{X}\setminus Z}$, ϕ' can be normalized to form an MTE density by calculating $\phi = K(\mathbf{x})^{-1} \cdot \phi'$ for all $\mathbf{x} \in \Omega_{\mathbf{X}\setminus Z}$. We assume that all input MTE potentials which represent probability potentials in an MTE influence diagram are normalized to be MTE probability densities prior to the solution phase.

3.3 MTE Influence Diagrams

An MTE influence diagram is an influence diagram with discrete and/or continuous chance variables in which all probability distributions are MTE probability densities as in (1) which satisfy the normalization condition in (4), and the joint utility function or the multiplicative factors of the joint utility function are MTE potential(s) as in (1). We assume decision nodes have discrete state spaces so that we stay in the class of MTE potentials during the solution process and also avoid optimization problems associated with continuous state spaces.

4 Operations on MTE Influence Diagrams

This section will describe the operations required to solve MTE influence diagrams.

4.1 Combination

Combination of MTE potentials is pointwise multiplication. Let ϕ_1 and ϕ_2 be MTE potentials for $\mathbf{X}_1 = \mathbf{Y}_1 \cup \mathbf{Z}_1 \cup \mathbf{D}_1$ and $\mathbf{X}_2 = \mathbf{Y}_2 \cup \mathbf{Z}_2 \cup \mathbf{D}_2$. The combination of ϕ_1 and ϕ_2 is a new MTE potential for $\mathbf{X} = \mathbf{X}_1 \cup \mathbf{X}_2$ defined as follows

$$\phi(\mathbf{x}) = \phi_1(\mathbf{x}^{\downarrow \Omega_{\mathbf{X}_1}}) \cdot \phi_2(\mathbf{x}^{\downarrow \Omega_{\mathbf{X}_2}}) \tag{5}$$

for all $\mathbf{x} \in \Omega_{\mathbf{X}}$.

Combination of two MTE probability densities results in an MTE probability density. Combination of an MTE probability density and an MTE utility potential results in an MTE utility potential. Combination of two MTE utility potentials results in an MTE utility potential. Combination of an MTE potential consisting of k_1 pieces with an MTE potential consisting k_2 pieces results in an MTE potential consisting of at most $k_1 \cdot k_2$ pieces. If the domains of the potentials do not overlap in all pieces, the resulting MTE potential may have less than $k_1 \cdot k_2$ pieces.

4.2 Marginalization

4.2.1 Chance Variables

Marginalization of chance variables in an MTE influence diagram corresponds to summing over discrete chance variables and integrating over continuous chance variables. Let ϕ be an MTE potential for $\mathbf{X} = \mathbf{Y} \cup \mathbf{Z} \cup \mathbf{D}$. The marginal of ϕ for a set of variables $\mathbf{X}' = \mathbf{Y}' \cup \mathbf{Z}' \cup \mathbf{D} \subseteq \mathbf{X}$ is an MTE potential computed as

$$\phi^{\downarrow \mathbf{X}'}(\mathbf{y}', \mathbf{z}', \mathbf{d}) = \sum_{\mathbf{y} \in \Omega_{\mathbf{Y} \setminus \mathbf{Y}'}} \left(\int_{\Omega_{\mathbf{Z} \setminus \mathbf{Z}'}} \phi(\mathbf{y}, \mathbf{z}, \mathbf{d}) \, d\mathbf{z}'' \right) \tag{6}$$

where $\mathbf{z} = (\mathbf{z}', \mathbf{z}'')$, and $(\mathbf{y}', \mathbf{z}', \mathbf{d}) \in \Omega_{\mathbf{X}'}$.

Although we show the continuous variables being marginalized before the discrete variables in (6), the variables can be marginalized in any sequence, resulting in the same final MTE potential.

4.2.2 Decision Variables

Marginalization with respect to a decision variable is only defined for MTE utility potentials. Let u be an MTE utility potential for $\mathbf{X} = \mathbf{Y} \cup \mathbf{Z} \cup \mathbf{D}$, where $D \in \mathbf{D}$. The marginal of u for a set of variables $\mathbf{X} - \{D\}$ is an MTE utility potential computed as

$$u^{\downarrow (\mathbf{X} - \{D\})}(\mathbf{y}, \mathbf{z}, \mathbf{d}') = \max_{d \in \Omega_D} u(\mathbf{y}, \mathbf{z}, \mathbf{d}) \tag{7}$$

for all $(\mathbf{y}, \mathbf{z}, \mathbf{d}') \in \Omega_{\mathbf{X} - \{D\}}$ where $\mathbf{d} = (\mathbf{d}', d)$.

In order to use the fusion algorithm to solve MTE influence diagrams, marginalization of decision variables must result in an MTE potential. The following theorem ensures this result.

Theorem 1. Let u_1 be an MTE utility potential for $\mathbf{X} = \mathbf{Y} \cup \mathbf{Z} \cup \mathbf{D}$, where $D \in \mathbf{D}$. If $u_1(\mathbf{y}, \mathbf{z}, \mathbf{d}', D = 1), ..., u_1(\mathbf{y}, \mathbf{z}, \mathbf{d}', D = n)$ are MTE utility potential fragments defined over the same domain, $\Omega_{\mathbf{X}}$, then $u_1^{\downarrow (\mathbf{X} - \{D\})} = \mathrm{Max}\{u_1(\mathbf{y}, \mathbf{z}, \mathbf{d}', D = 1), \ldots, u_1(\mathbf{y}, \mathbf{z}, \mathbf{d}', D = n)\}$ can be represented as an MTE utility potential whose components are equal to one of the fragments $u_1(\mathbf{y}, \mathbf{z}, \mathbf{d}', D = 1), \ldots, u_1(\mathbf{y}, \mathbf{z}, \mathbf{d}', D = n)$ in each region of a hypercube of $\Omega_{\mathbf{Z}}$, where $\mathbf{Z}$ are the continuous chance variables in $\mathbf{X}$.

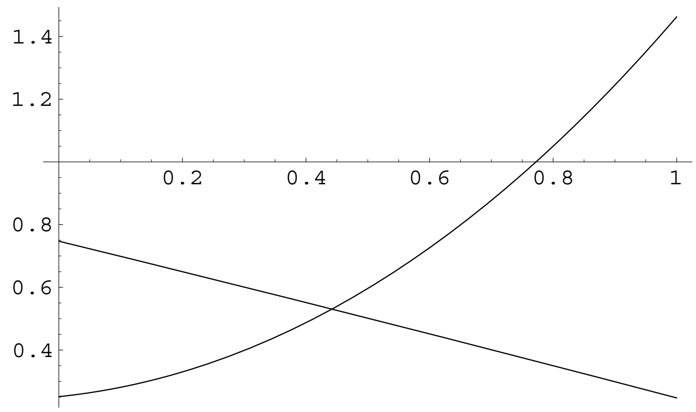

Fig. 2. The MTE utility potential fragments $u_1(x, D = 1)$ and $u_1(x, D = 2)$ in *Example 1.*

Proof. The proof of Theorem 1 is demonstrated using the following example. Consider the following MTE utility potential fragments defined over the interval $[0, 1]$:

$$\begin{aligned} u_1(x, D = 1) &= 7.439066 - 6.692019 \exp\{0.0720496x\} \\ u_1(x, D = 2) &= 176.548799 - 355.430039 \exp\{0.0720496x\} \\ &\quad +179.132456 \exp\{0.1440991x\} \end{aligned}$$

These MTE utility potential fragments are shown graphically in Figure 2.

The function $u_1^{\downarrow X} = \text{Max}\{u_1(x, D = 1), u_2(x, D = 2)\}$ can be represented as an MTE utility potential $u_{max}(x)$ by finding the point where the MTE utility potentials $u_1(x, D = 1)$ and $u_1(x, D = 2)$ are equal (if any), then determining which component maximizes the new potential on the resulting sub-intervals. To find the point where the two functions are equal, we use the procedure described in Section 4.3. The MTE utility potential fragments are equal at $x = 0.442$, so the new MTE utility potential is defined as:

$$u_{max}(x) = \begin{cases} 7.439066 - 6.692019 \exp\{0.0720496x\} & \text{if } 0 \leq x < 0.442 \\ 176.548799 - 355.430039 \exp\{0.0720496x\} & \\ +179.132456 \exp\{0.1440991x\} & \text{if } 0.442 \leq x \leq 1. \end{cases}$$

This MTE utility potential is shown graphically in Figure 3.

A similar maximization process can be used with multi-variate MTE utility potentials and/or when more than two MTE utility potentials are being evaluated to define a new component potential.

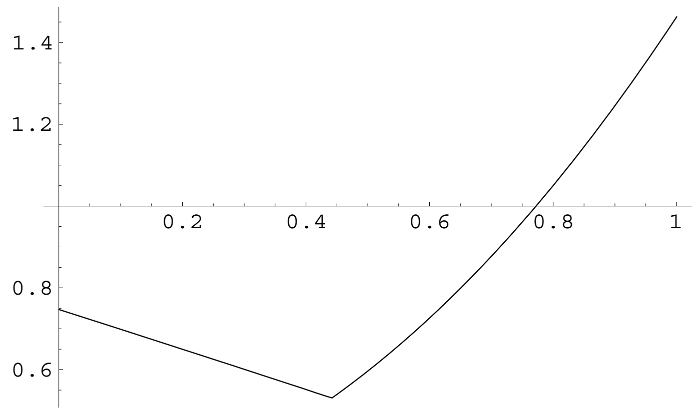

Fig. 3. The MTE utility potential u_{max} for X in the proof of Theorem 1.

4.3 Defining Component MTE Potentials

Theorem 1 states that the maximum of several MTE utility potentials can be defined as an MTE utility potential whose components are equal to one of the original MTE utility potentials. Defining the marginalized MTE utility potential involves finding the points where the component potentials are equal (if any), then evaluating each component potential on either side of these points to determine which components comprise the marginalized MTE utility potential.

Consider the problem of finding $f(x) = \text{Max}\{u_1(x), \ldots, u_n(x)\}$ where $\Psi = \{u_1(x), ..., u_n(x)\}$ are univariate MTE utility potentials with $\Omega_X = \{x : x \in [a, b]\}$. We use the bisection search method (see, e.g., Bazaraa *et al.* 1993) to find the points where $u_i(x) - u_j(x) = 0$, for all $i, j = 1, ..., n$, $i \neq j$. Let $E = \{e_1, ..., e_K\}$ be the set of points where two or more of the MTE utility potentials in Ψ are equal (with the exception that $e_1 = a$) and let ϵ be the allowable length of the final interval of uncertainty. Let $\Phi = \{\phi_1, ..., \phi_K\}$ be a vector of values defining the index of the MTE utility potential in Ψ that maximizes $f(x)$ on each of the sub-intervals created by the points in E, with $max_k = u_{\phi_k}(e_k + \epsilon)$. The pseudo-code of a procedure for finding the component MTE utility potentials in the function $f(x)$ described above is as follows:

INPUT: Ψ, a, b, ϵ
OUTPUT: E, Φ
INITIALIZATION
$E \leftarrow \{a\}$ /* E contains the equality points of potentials in Ψ */
$\phi_1 \leftarrow 1$
$\Phi \leftarrow \{\phi_1\}$ /* Φ contains the index for the maximum MTE utility potentials */
FOR $i = 1 : (n-1)$
Find points d_i where $u_i(x) - u_j(x) = 0$, for all $j = i+1, ...n$, $i \neq j$, within error bound ϵ using the bisection search method.

$E \leftarrow E \cup \{d_i\}$

$E \leftarrow$ Rank(E) /* Re-order E in ascending sequence */

END FOR

FOR $k = 1 : | E |$

$max_k = u_1(e_k + \epsilon)$

FOR $i = 2 : n$

IF $u_i(e_k + \epsilon) > max_k$

$\phi_k = n$

$max_k = u_i(e_k + \epsilon)$

END IF

END FOR

END FOR

The resulting component MTE utility potential defined in the procedure above is described as follows:

$$u_{max}(x) = \begin{cases} u_{\phi_1}(x) \text{ if } a < x < e_2 \\ u_{\phi_2}(x) \text{ if } e_2 \le x \le e_3 \\ \vdots \\ u_{\phi_K}(x) \text{ if } e_K < x < b \ . \end{cases}$$

4.4 Fusion Algorithm

A fusion algorithm for solving influence diagrams is described by Shenoy (1993). The fusion algorithm involves deleting variables from the network in a sequence which respects the precedence constraints (represented by arcs pointing to decision variables in influence diagrams) in the problem. In influence diagrams, for any random variable R and any decision variable D, if R is an information ancestor of D, then this means that the true value of R will be known at the point when a decision alternative has to be chosen at D. If R is not an information ancestor of D, then the true value of R is assumed to be not known. The set of such relations constitutes the precedence relations for the problem. A deletion sequence for the variables must obey the precedence relation in the sense that if R is known when a choice at D has to be made, then D must be deleted before R, and vice-versa. There may be several deletion sequences possible for an influence diagram, but all of these will lead to the same solution.

During the solution phase, the next variable, X, in the deletion sequence is selected and a "fusion" operation is performed on the utility and/or probability potentials containing X in their domain. The fusion operation combines all potentials with X in their domains, and the resulting potential is marginalized such that X is eliminated from the domain. The marginalization operation used depends on whether X is a continuous chance variable, discrete chance variable, or decision variable. The potentials without X in their domain remain unchanged and the solution proceeds similarly until all variables are eliminated from the network. The solution in Section 6.1 is an example utilizing the fusion algorithm.

The fusion method applies to problems where there is only one utility function (or a joint utility function which factors multiplicatively into several utility potentials) and uses only the operations of combination and marginalization as described above. Moral *et al.* (2001) shows that the class of MTE potentials is closed under marginalization (of chance variables) and combination. Theorem 1 states that the class of MTE potentials is closed under marginalization of discrete decision variables. Thus, MTE influence diagrams can be solved using the fusion algorithm, since only combinations and marginalizations are performed.

The general fusion algorithm described by Shenoy (1992) for additive decomposition of the joint utility function involves division of probability potentials. Since the class of MTE potentials is not closed under division, we restrict our discussion to the case of multiplicative decomposition as no divisions are required for this case.

5 Estimating MTE Utility Functions

In this paper, we consider problems with one joint utility function or a joint utility function which factors multiplicatively. The utility functions can be of any form as long as one can approximate them using MTE potentials. For instance, consider a class of utility functions as follows. The joint utility function (or one of its multiplicative factors) is a function of M variables in the influence diagram, $u = f(X_1, ..., X_M)$. This joint utility function may be composed of P additive factors, each a function of K_p multiplicative factors, where K_p is the number of multiplicative factors composing the p-th additive factor. Thus, we can restate the joint utility function as

$$u = f_{1,1}(\mathbf{x_i}) \cdots f_{1,K_1}(\mathbf{x_i}) + \ldots + f_{P,1}(\mathbf{x_i}) \cdots f_{P,K_P}(\mathbf{x_i}) \ , \tag{8}$$

where $\mathbf{x_i}$ is a vector of values for a subset of the variables in the influence diagram. We restrict the components $f_{p,k}(\mathbf{x_i})$ to be one of three forms:

(1) A real number.
(2) An MTE potential of the form in (1).

(3) An arbitrary polynomial function of one variable, x_i.

The class of utility functions described above contains the class of polynomial functions.

Example 1. Consider the joint utility function $u(x,y,z) = 3x^2y+4z^2+3xz^2+3y^2$.

This utility function contains four additive factors, $P = 4$. These additive factors can be decomposed into constants and the following linear functions: $g_1(x) = x, g_2(y) = y, g_3(z) = z$. Thus, $u(x,y,z)$ can be restated as

$$u(x,y,z) = 3[g_1(x)]^2 \cdot g_2(y) + 4[g_3(z)]^2 + 3g_1(x) \cdot [g_3(z)]^2 + 3[g_2(y)]^2 \ .$$

To create an MTE potential which approximates the joint utility function, we create an MTE approximation $\phi_{p,k}(x_i) = a_0^{(p,k)} + a_1^{(p,k)} \exp\{a_2^{(p,k)} x_i\}$ for each $f_{p,k}(x_i)$ (unless $f_{p,k}(\mathbf{x_i})$ is already of the form in (1) or is a constant), then re-combine the MTE approximations into a joint MTE utility potential. Functions that exhibit no changes in concavity/convexity over their domains can be well-approximated by an MTE potential with one independent term and one exponential term; if a term $f_{p,k}(x_i)$ has changes in concavity/convexity, we can simply fit additional exponential terms. The result of the combination is an MTE utility potential because the class of MTE potentials is closed under addition and multiplication (Moral *et al.* 2001).

To create the MTE approximations $\phi_{p,k}(x_i)$, we use unconstrained, non-linear optimization and solve

$$\underset{a_0^{(p,k)}, a_1^{(p,k)}, a_2^{(p,k)}}{\text{argmin}} \sum_{j=0}^{n} (f_{p,k}(x_{ij}) - \phi_{p,k}(x_{ij}))^2$$

where $x_{ij}, j = 0, ..., n$ is a set of points obtained by evenly dividing the domain of X_i. We then replace each $f_{p,k}(x_i)$ in (8) with its MTE approximation $\phi_{p,k}(x_i)$ and simplify the function accordingly.

6 Examples

6.1 Oil Wildcatter with Continuous Uncertainties

This example is an adaptation of the Oil Wildcatter problem from Raiffa (1968). We model some variables as continuous uncertainties.

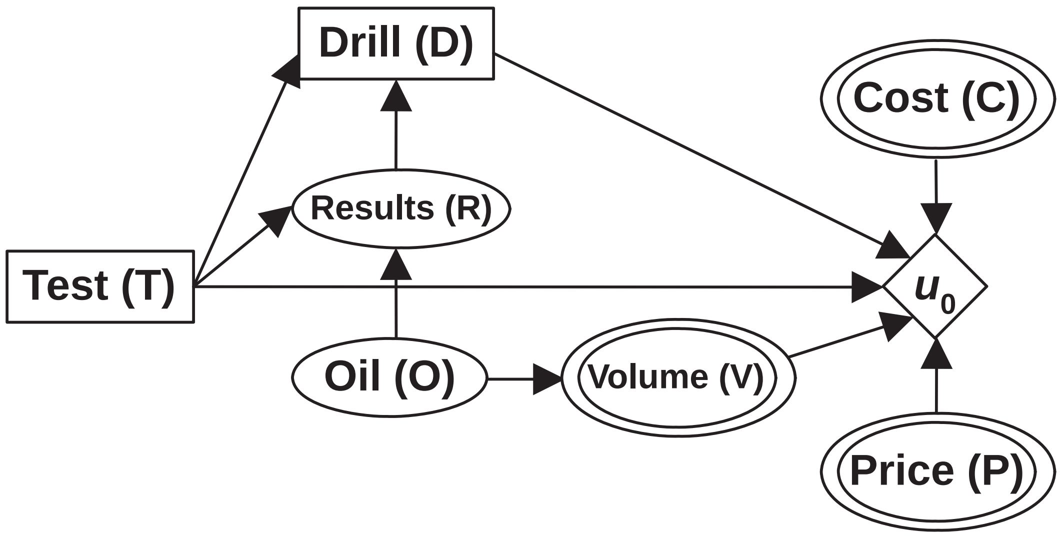

Fig. 4. A hybrid influence diagram representation of the Oil Wildcatter problem with continuous uncertainties.

An oil wildcatter must decide whether to drill ($D = 1$) or not drill ($D = 0$). He is uncertain whether the hole is dry ($O = 0$), wet ($O = 1$), or soaking ($O = 2$). The oil volume (V) extracted depends on the state of oil (O). If the hole is dry ($O = 0$), no oil is extracted. If $O = 1$, the amount of oil extracted follows a normal distribution with a mean of 6 thousand barrels and a standard deviation of 1 thousand barrels, i.e. $\pounds(V \mid O = 1) \sim N(6, 1^2)$. If $O = 2$, the amount of oil extracted follows a normal distribution with a mean of 13.5 thousand barrels and a standard deviation of 2 thousand barrels, i.e. $\pounds(V \mid O = 2) \sim N(13.5, 2^2)$. The cost of drilling ($C$) is normally distributed with a mean of 70 thousand dollars and a standard deviation of 10 thousand dollars, i.e. $\pounds(C) \sim N(70, 10^2)$. The log of oil prices (P) follows a normal distribution with a mean of \$2.75 and a standard deviation of \$0.7071, i.e. $\pounds(P) \sim LN(2.75, 0.7071^2)$. The wildcatter assumes potential θ for O with values $\theta(0) = P(O = 0) = 0.500$, $\theta(1) = P(O = 1) = 0.300$, and $\theta(2) = P(O = 2) = 0.200$.

At a cost of 10 thousand dollars, the wildcatter can conduct a seismic test which will help determine the geological structure at the site. The test results (R) will disclose whether the structure under the site has no structure ($R = 0$) (bad), open structure ($R = 1$) (so-so), or closed structure ($R = 2$) (very hopeful). Experts have provided Table 1 which shows the probabilities of test results (R) conditional on the state of oil (O) and test (T) (which we will refer to as potential δ for $\{R, O, T\}$). Figure 4 shows a hybrid influence diagram representation of the Oil Wildcatter problem with some discrete and some continuous chance variables.

Table 1
Probabilities of seismic test results conditional on the amount of oil and test.

Amount of Oil (O)	Seismic Test Results (R): No Structure ($R=0$)	Open Structure ($R=1$)	Closed Structure ($R=2$)	No Result (NR)
$P(R \mid O, T=1)$				
Dry ($O=0$)	0.60	0.30	0.10	0
Wet ($O=1$)	0.30	0.40	0.30	0
Soaking ($O=2$)	0.10	0.40	0.50	0
$P(R \mid O, T=0)$				
Dry ($O=0$)	0	0	0	1
Wet ($O=1$)	0	0	0	1
Soaking ($O=2$)	0	0	0	1

6.1.1 Representation

The single utility function in the problem has domain $\{C, P, V, D, T\}$ and can be stated (in \$000) as

$$
\begin{aligned}
u_0(v, p, c, D=1, T=1) &= v \cdot p - c - 10 \\
u_0(v, p, c, D=1, T=0) &= v \cdot p - c \\
u_0(v, p, c, D=0, T=1) &= -10 \\
u_0(v, p, c, D=0, T=0) &= 0\,.
\end{aligned}
$$

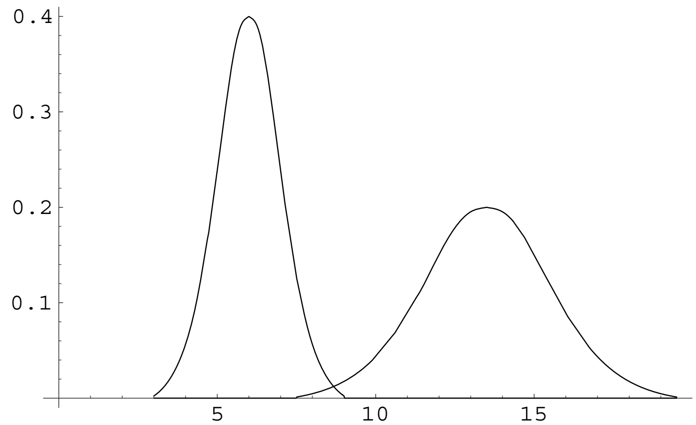

Fig. 5. MTE approximations of the PDF's for V given O.

The utility function u_0 can be approximated by an MTE potential u_1 by using the method described in Section 5. The resulting MTE potential is

$$
\begin{aligned}
&u_1(v,p,c,D=1,T=1) = \\
&\qquad 600,462,529.9767685 + 24,504.975886\ \exp\{-0.00004109695c\} \\
&\qquad -600,488,161.2450081\ \exp\{0.00004069868p\} \\
&\qquad +600,488,190.477144\ \exp\{0.00004069868p + 0.00004078953v\} \\
&\qquad -600,487,073.8393291\ \exp\{0.00004078953v\} \\
&u_1(v,p,c,D=1,T=0) = u_1(v,p,c,D=1,T=1) - 10 \\
&u_1(v,p,c,D=0,T=1) = -10 \\
&u_1(v,p,c,D=0,T=0) = 0\,.
\end{aligned}
$$

Normally distributed chance variables are modeled using the 2-piece MTE approximation to the normal PDF given in (3). The PDF's for V given $O=1$ and V given $O=2$ are approximated by the MTE potential fragments $\nu(v,O=1)$ and $\nu(v,O=2)$, respectively. These potential fragments constitute the potential ν for $\{V,O\}$ and are displayed graphically in Figure 5. The PDF for C is approximated by the MTE potential ϑ.

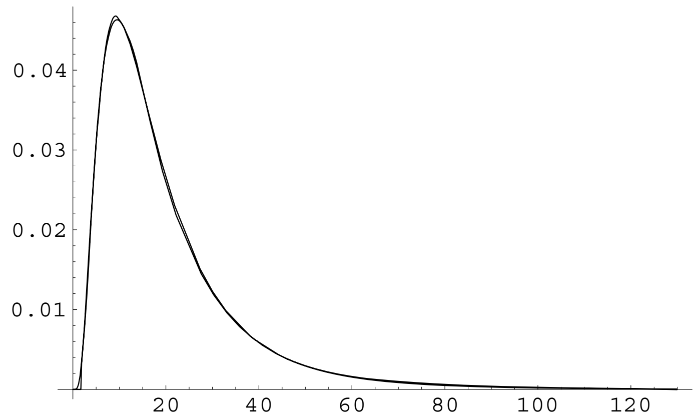

Fig. 6. The MTE approximation for the distribution of P overlayed on the actual $LN(2.75, 0.7071^2)$ distribution.

An MTE approximation of the lognormal PDF ρ for P is constructed using the procedure in Cobb *et al.* (2006). This MTE potential is as follows:

$$\rho(p) = f_P(p) =$$

$$\begin{cases} -0.024921 + 0.186834 \exp\{0.249714(p - 9.44687)\} & \\ +0.101347 \exp\{1.419659(p - 9.44687)\} & \text{if } 1.86706 \le p < 3.47531 \\ 0.174804 - 0.062119 \exp\{-0.116729(p - 9.44687)\} & \\ -0.066038 \exp\{0.116608(p - 9.44687)\} & \text{if } 3.47531 \le p < 9.44687 \\ 0.049064 + 0.000000154912 \exp\{1.480552(p - 9.44687)\} & \\ -0.002427 \exp\{0.287079(p - 9.44687)\} & \text{if } 9.44687 \le p < 15.57526 \\ -0.583002 + 0.057534 \exp\{-0.079477(p - 9.44687)\} & \\ +0.584025 \exp\{-0.000015(p - 9.44687)\} & \text{if } 15.57526 \le p \le 129.93107. \end{cases}$$

The potential ρ is displayed graphically in Figure 6 overlayed on the actual $LN(2.75, 0.7071^2)$ distribution.

6.1.2 Solution

To calculate the optimal strategy and expected profit associated with that strategy, we use the fusion algorithm and delete the variables in the sequence C, P, V, O, D, R, T.

To remove C, we calculate $u_2 = (u_1 \otimes \vartheta)^{\downarrow\{P,V,D,T\}}$. An example of this calcu-

Table 2
The utility function u_4 with domain $\{O, D, T\}$ resulting from the removal of variable V (\$000).

Amount of Oil (O)	Values of Drill (D) and Test (T) $D=1$, $T=1$	$D=1$, $T=0$	$D=0$, $T=1$	$D=0$, $T=0$
Dry ($O=0$)	–82.752	–72.752	–10.000	0
Wet ($O=1$)	40.955	50.955	–10.000	0
Soaking ($O=2$)	192.219	202.219	–10.000	0

lation and one of the resulting utility potential fragments is as follows:

$$u_2(v, p, D = 1, T = 1) = \int_{\Omega_C} (u_1(v, p, c, D = 1, T = 1) \cdot \vartheta(c))\, dc =$$
$$600,487,104.6908119$$
$$-600,488,301.3756697\ \exp\{0.00004069868p\}$$
$$+600,488,330.6078435\ \exp\{0.00004069868p + 0.00004078953v\}$$
$$-600,487,213.969722\ \exp\{0.00004078953v\}$$

To remove P, we calculate $u_3 = (u_2 \otimes \rho)^{\downarrow\{V,D,T\}}$. An example of this calculation and one of the resulting utility potential fragments is as follows:

$$u_3(v, D = 1, T = 1) = \int_{\Omega_P} (u_2(v, p, D = 1, T = 1) \cdot \rho(p))\, dp =$$
$$-494,334.1115046 + 494,254.0887283\ \exp\{0.00004078953v\}$$

To remove V, we calculate $u_4 = (u_3 \otimes \nu)^{\downarrow\{O,D,T\}}$, which is described in Table 2. The potentials remaining in the network after removal of V are u_4 with domain $\{O, D, T\}$, δ with domain $\{R, O, T\}$ and θ with domain $\{O\}$. Thus, to remove O, we calculate $u_5 = (u_4 \otimes \theta \otimes \delta)^{\downarrow\{D,R,T\}}$, which is described in Table 3.

Removing D involves simply maximizing the utility in Table 3 for each configuration of $\{R, T\}$. The resulting utility function u_6 is shown in Table 4. The optimal policy is drill ($D = 1$) if a test is performed and the results reveal open structure ($R = 1$) or closed structure ($R = 2$), not drill ($D = 0$) if a test is performed and the results reveal no structure ($R = 0$), and drill ($D = 1$) if no test is performed.

Table 3
The utility function u_5 with domain $\{D, R, T\}$ resulting from the removal of variable O ($000).

Results of Test (R)	$D = 1$, $T = 1$	$D = 1$, $T = 0$	$D = 0$, $T = 1$	$D = 0$, $T = 0$
	Values of Drill (D) and Test (T)			
No Result	0	19.354	0	0
No Struct. ($R = 0$)	–17.295	0	–4.100	0
Open Struct. ($R = 1$)	7.879	0	–3.500	0
Closed Struct. ($R = 2$)	18.770	0	–2.400	0

Table 4
The utility function u_6 with domain $\{R, T\}$ resulting from the removal of variable D ($000), with optimal policies.

Results of Test (R)	$T = 1$	$T = 0$
	Value of Test (T)	
No Result	0	19.354 ($D = 1$)
No Struct. ($R = 0$)	–4.100 ($D = 0$)	0
Open Struct. ($R = 1$)	7.879 ($D = 1$)	0
Closed Struct. ($R = 2$)	18.770 ($D = 1$)	0

Summing the values in Table 4 over the possible values of R gives $u_7(T = 1) = -4.100 + 7.879 + 18.770 = 22.550$ and $u_7(T = 0) = 19.354$. Thus, the optimal test decision is to test ($T = 1$), and the maximum expected profit is $22,550.

6.2 Oil Wildcatter with Continuous Test Results

Suppose that the seismic test in the Oil Wildcatter example yields a continuous reading (R) representing the location of the peak response, measured on the unit interval [0,1]. The PDF's for R given O and $T = 1$ are symmetric beta distributions as follows: $\pounds(R \mid O = 0, T = 1) \sim Beta(1, 1)$, $\pounds(R \mid O = 1, T = 1) \sim Beta(3.2, 3.2)$, and $\pounds(R \mid O = 2, T = 1) \sim Beta(4.2, 4.2)$. These distributions are approximated by MTE potential fragments using the

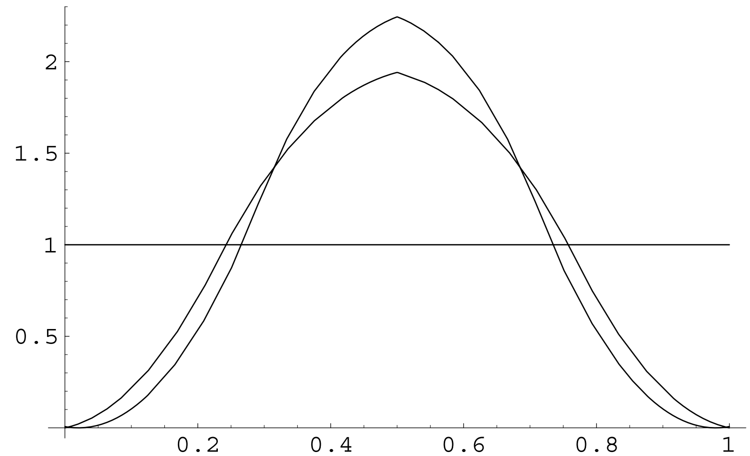

Fig. 7. The MTE potential fragments for the distribution of R given O and $T = 1$ representing the distributions of test results given the state of oil.

procedure in Cobb *et al.* (2006). The PDF for R given $\{O = 0, T = 1\}$ is the uniform distribution on the interval $[0, 1]$, which is already an MTE potential. The PDF for R given $\{O = 1, T = 1\}$ is approximated by an MTE potential fragment as follows:

$$\delta(r, O = 1, T = 1) = f_{\{O=1,T=1\}}(r) =$$

$$\begin{cases} 4,476.386259 - 8,644.466462 \exp\{0.0542577r\} & \\ +4,168.080202 \exp\{0.1127101r\} & \text{if } 0 < r < 0.228837 \\ -9.771378 - 15.207250 \exp\{-2.0318469r\} & \\ +23.169263 \exp\{-0.5936362r\} & \text{if } 0.228837 \leq r < 0.5 \\ -9.771378 - 15.207250 \exp\{-2.0318469(1-r)\} & \\ +23.169263 \exp\{-0.5936362(1-r)\} & \text{if } 0.5 \leq r < 0.771163 \\ 4,476.386259 - 8,644.466462 \exp\{0.0542577(1-r)\} & \\ +4,168.080202 \exp\{0.1127101(1-r)\} & \text{if } 0.771163 \leq r < 1. \end{cases}$$

The PDF for R given $\{O = 2, T = 1\}$ is approximated similarly by an MTE potential fragment and denoted as $\delta(r, O = 2, T = 1)$. The MTE potential fragments $\delta(r, O = 0, T = 1)$, $\delta(r, O = 1, T = 1)$ and $\delta(r, O = 2, T = 1)$ constitute the potential fragment δ for $\{R, O, T = 1\}$. These fragments are shown in Figure 7—$\delta(r, O = 0, T = 1)$ is the flat distribution, $\delta(r, O = 2, T = 1)$ is the most peaked distribution, and $\delta(r, O = 1, T = 1)$ is in-between. An observation in the middle of the unit interval will favor $O = 2$, an observation near the extremes $R = 0$ or 1 will favor $O = 0$, and an observation around $R = 0.275$ or 0.725 will favor $O = 1$.

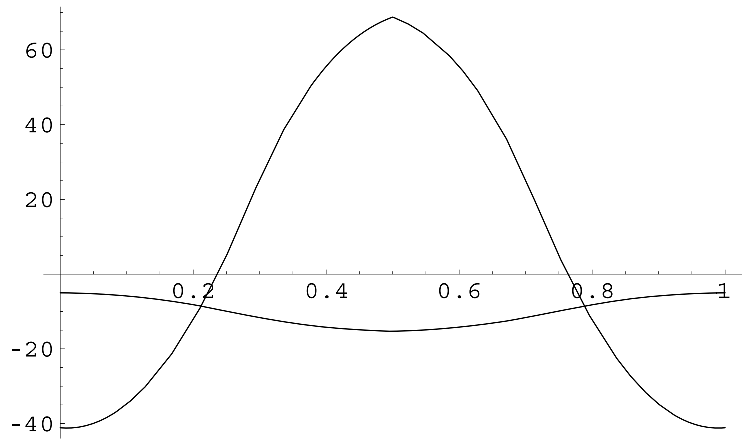

Fig. 8. Utility potential fragments $u_5(r, D = 1, T = 1)$ (increasing on (0,0.5]) and $u_5(r, D = 0, T = 1)$.

The solution remains the same as in Section 6.1 through the elimination of variable V. The potentials remaining in the network after removal of V are u_4 with domain $\{O, D, T\}$, δ with domain $\{R, O, T\}$ and θ with domain $\{O\}$. Thus, to remove O, we calculate $u_5 = (u_4 \otimes \theta \otimes \delta)^{\downarrow \{D,R,T\}}$. The utility fragments $u_5(r, D = 1, T = 1)$ and $u_5(r, D = 0, T = 1)$ are shown graphically in Figure 8. The other utility potential fragments constituting the utility potential u_5 are constants: $u_5(r, D = 1, T = 0) = 19.354$ and $u_5(r, D = 0, T = 0) = 0$.

For $T = 1$, removing D involves finding $\mathrm{Max}\{u_5(r, D = 1, T = 1), u_5(r, D = 0, T = 1)\}$ at each point in the domain of R. We can recover an MTE potential from this calculation by using the procedure described in Section 4.3. In this case, we find $u_5(r, D = 1, T = 1) \approx u_5(r, D = 0, T = 1)$ at 0.212 and 0.788. The resulting utility potential fragment $u_6(r, T = 1)$ (see Figure 9) is defined as follows:

$$u_6(r, T = 1) = \begin{cases} u_5(r, D = 0, T = 1) \text{ if } 0 < r < 0.212 \\ u_5(r, D = 1, T = 1) \text{ if } 0.212 \leq r \leq 0.788 \\ u_5(r, D = 0, T = 1) \text{ if } 0.788 < r < 1. \end{cases}$$

The optimal strategy is drill if the test result is in the interval $[0.212, 0.788]$ and not drill otherwise. For $T = 0$, removing D involves simply selecting the value of D which yields the highest utility; thus, we select $D = 1$ which gives $u_6(r, T = 0) = 19.354$.

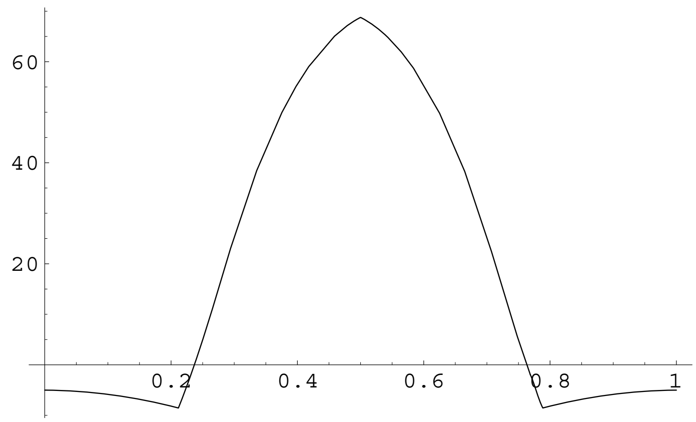

Fig. 9. The utility potential fragment $u_6(r, T = 1)$.

Removing R results in utility potential u_7, defined as follows:

$$u_7(T = 1) = \int_0^1 u_6(r, T = 1)\, dr = 19.802$$

$$u_7(T = 0) = 19.354\,.$$

Thus, the optimal decision is to test ($T = 1$) and the maximum expected profit is \$19,802.

7 Summary and Conclusions

We have described MTE influence diagrams and demonstrated a procedure for solving MTE influence diagrams with one joint utility function (or multiplicative factors of one joint utility function) when probability distributions are represented by MTE probability densities and utility functions are represented by MTE utility potentials. Any continuous PDF can be modeled by an MTE potential, so any continuous random variable can be represented in an MTE influence diagram. This includes, e.g., conditional linear Gaussian, gamma, beta, and lognormal distributions. The solution method presented places no restrictions on the arrangement of discrete and continuous chance variables in the influence diagram.

As described, MTE influence diagrams have some limitations. First, the numerical stability of the solution algorithm may be an issue in problems where the MTE approximations have very large and/or very small parameters. Rumí and Salmerón (2005) devised a method to improve computational efficiency and numerical stability of calculations with MTE potentials. In operations involving MTE probability potentials, their method can be used to "prune" terms that contribute an insignificant amount of density to the resulting potential, thus reducing the overall number of terms in the resulting potential, which eliminates some terms with very small parameters. We have also experimented with a method of operating with a linear transformation of an MTE utility potential during the solution phase in MTE influence diagrams. This method takes advantage of the fact that for a constant k, an MTE probability density ϕ for Z, and an MTE utility potential u for Z, the following holds:

$$k \cdot \int_{\Omega_Z} u(z) \cdot \phi(z)\, dz = \int_{\Omega_Z} (k \cdot u(z)) \cdot \phi(z)\, dz$$

Thus, we can transform the MTE utility potential such that the parameters are as close to one as possible, perform operations on the MTE influence diagram, then recover the appropriate expected utility at the end. Such a method can help avoid making successive calculations with increasingly smaller or larger numbers. We leave the detailed description of these two methods for improving the numerical stability of MTE influence diagrams for future research.

Second, MTE influence diagrams only allow for multiplicative factorization of the joint utility function. This is because solving an influence diagram with an additive factorization involves division of potentials, and the class of MTE potentials is not closed under division.

Third, MTE influence diagrams only allow discrete decision variables. This is because Theorem 1, which states that the class of MTE potentials is closed under marginalization of decision variables, holds only for discrete decision variables. Cobb (2006) describes an extension of MTE influence diagrams that allows continuous decision variables in certain configurations.

Acknowledgments

This research was partly funded by a graduate research assistantship provided to the first author from the Ronald G. Harper Professorship at the University of Kansas.

References

Bazaraa, M.S., Sherali, H.D. and C.M. Shetty, 1993. Nonlinear Programming: Theory and Algorithms, John Wiley and Sons, New York, NY.

Bielza, C., Müller, P. and D. Rios Insua, 1999. Decision analysis by augmented probability simulation, Management Science 45(7) 995–1007.

Charnes, J.M. and P.P. Shenoy, 2004. Multi-stage Monte Carlo method for solving influence diagrams using local computation, Management Science 50(3) 405–418.

Cobb, B.R., 2006. Continuous decision MTE influence diagrams. In: M. Studeny and J. Vomlel (Eds.), Proceedings of the Third European Conference on Probabilistic Graphical Models (PGM–06), pp. 67–74.

Cobb, B.R. and P.P. Shenoy, 2006. Inference in hybrid Bayesian networks with mixtures of truncated exponentials, International Journal of Approximate Reasoning 41(3) 257–286.

Cobb, B.R., Shenoy, P.P. and R. Rumí, 2006. Approximating probability density functions in hybrid Bayesian networks with mixtures of truncated exponentials, Statistics and Computing 16(3) 293–308.

Howard, R.A. and J.E. Matheson, 1984. Influence diagrams. In: R.A. Howard and J.E. Matheson (Eds.), Readings on the Principles and Applications of Decision Analysis, Volume 2, Strategic Decisions Group, Menlo Park, CA, pp. 719–762.

Jensen, F., Jensen, F.V. and S. Dittmer, 1994. From influence diagrams to junction trees. In: R. López de Mántaras and D. Poole (Eds.), Uncertainty in Artificial Intelligence, 10, Morgan–Kaufman, San Francisco, pp. 367–373.

Lauritzen, S.L. and D. Nilsson, 2001. Representing and solving decision problems with limited information, Management Science 47(9) 1238–1251.

Madsen, A.L. and F. Jensen, 2005. Solving linear-quadratic conditional Gaussian influence diagrams, International Journal of Approximate Reasoning 38(3) 263–282.

Madsen, A.L. and F.V. Jensen, 1999. Lazy evaluation of symmetric Bayesian decision problems. In: K.B. Laskey and H. Prade (Eds.), Uncertainty in Artificial Intelligence, 15, Morgan–Kaufman, San Francisco, pp. 382–390.

Madsen, A.L. and D. Nilsson, 2001. Solving influence diagrams using HUGIN, Shafer-Shenoy and Lazy propagation. In: J. Breese and D. Koller (Eds.), Uncertainty in Artificial Intelligence, 17, Morgan–Kaufman, San Francisco, pp. 337–345.

Moral, S., Rumí, R. and A. Salmerón, 2001. Mixtures of truncated exponentials in hybrid Bayesian networks. In: P. Besnard and S. Benferhart (Eds.), Symbolic and Quantitative Approaches to Reasoning under Uncertainty, Lecture Notes in Artificial Intelligence, 2143, Springer–Verlag, Heidelberg, pp. 156–167.

Moral, S., Rumí, R. and A. Salmerón, 2002. Estimating mixtures of truncated exponentials from data. In: J.A. Gamez and A. Salmerón (Eds.), Proceedings of the First European Workshop on Probabilistic Graphical Models, Cuenca, Spain, pp. 135–143.

Moral, S., Rumí, R. and A. Salmerón, 2003. Approximating conditional MTE distributions by means of mixed trees. In: T.D. Nielsen and N.L. Zhang (Eds.), Symbolic and Quantitative Approaches to Reasoning under Uncertainty, Lecture Notes in Artificial Intelligence, 2711, Springer–Verlag, Heidelberg, pp. 173–183.

Olmsted, S.M., 1983. On representing and solving decision problems, Ph.D. Thesis, Department of Engineering–Economic Systems, Stanford University, Stanford, CA.

Poland, W.B., 1994. Mixtures of Gaussians and minimum relative entropy techniques for modeling continuous distributions, Ph.D. Thesis, Department of Engineering–Economic Systems, Stanford University, Stanford, CA.

Poland, W.B. and R.D. Shachter, 1993. Mixtures of Gaussians and minimum relative entropy techniques for modeling continuous uncertainties. In: D. Heckerman and E.H. Mamdani (Eds.), Uncertainty in Artificial Intelligence, 9, Morgan–Kaufman, San Francisco, pp. 183–190.

Raiffa, H., 1968. Decision Analysis, Addison-Wesley, Reading, MA.

Rumí, R., 2003. Modelos De Redes Bayesianas Con Variables Discretas Y Continuas, Doctoral Thesis, Universidad de Almería, Departamento de Estadística y Matemática Aplicada, Almería, Spain.

Rumí, R. and A. Salmerón, 2005. Penniless propagation with mixtures of truncated exponentials. In: L. Godo (Ed.), Symbolic and Quantitative Approaches to Reasoning under Uncertainty, Lecture Notes in Artificial Intelligence, 3571, Springer–Verlag, Heidelberg, pp. 39–50.

Shachter, R.D., 1986. Evaluating influence diagrams, Operations Research 34(6) 871–882.

Shachter, R.D. and C.R. Kenley, 1989. Gaussian influence diagrams, Management Science 35(5) 527–550.

Shachter, R.D. and P. Ndilikilikesha, 1993. Using potential influence diagrams for probabilistic inference and decision making. In: D. Heckerman and E.H. Mamdani (Eds.), Uncertainty in Artificial Intelligence, 9, Morgan–Kaufman, San Francisco, pp. 383–390.

Shenoy, P.P., 1992. Valuation-Based Systems for Bayesian decision analysis, Operations Research 40(3) 463–484.

Shenoy, P.P., 1993. A new method for representing and solving Bayesian decision problems. In: D.J. Hand (Ed.), Artificial Intelligence Frontiers in Statistics: AI and Statistics III, Chapman and Hall, London, pp. 119–138.