Decision Analysis

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Directed Expected Utility Networks

http://orcid.org/0000-0002-2562-5192Manuele Leonelli, Jim Q. Smith

informs
http://pubsonline.informs.org/journal/deca/

DECISION ANALYSIS
Vol. 14, No. 2, June 2017, pp. 108–125
ISSN 1545-8490 (print), ISSN 1545-8504 (online)

Directed Expected Utility Networks

Manuele Leonelli,[a] **Jim Q. Smith**[b]

[a] School of Mathematics and Statistics, University of Glasgow, Glasgow G12 8QW, United Kingdom; [b] Department of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom

Contact: manuele.leonelli@glasgow.ac.uk, http://orcid.org/0000-0002-2562-5192 (ML); j.q.smith@warwick.ac.uk (JQS)

Received: August 1, 2016
Revised: October 26, 2016; February 7, 2017
Accepted: February 16, 2017
Published Online in Articles in Advance: May 9, 2017

https://doi.org/10.1287/deca.2017.0347

Abstract. A variety of statistical graphical models have been defined to represent the conditional independences underlying a random vector of interest. Similarly, many different graphs embedding various types of preferential independences, such as, for example, conditional utility independence and generalized additive independence, have more recently started to appear. In this paper, we define a new graphical model, called a directed expected utility network, whose edges depict both probabilistic and utility conditional independences. These embed a very flexible class of utility models, much larger than those usually conceived in standard influence diagrams. Our graphical representation and various transformations of the original graph into a tree structure are then used to guide fast routines for the computation of a decision problem's expected utilities. We show that our routines generalize those usually utilized in standard influence diagrams' evaluations under much more restrictive conditions. We then proceed with the construction of a directed expected utility network to support decision makers in the domain of household food security.

Funding: The work of M. Leonelli was partly supported by Capes, while the work of J. Q. Smith was partly supported by the Engineering and Physical Sciences Research Council (EPSRC) [Grant EP/K039628/1] and the Alan Turing Institute under EPSRC [Grant EP/N510129/1].

Keywords: Bayesian networks • expected utility • graphical models • utility diagrams

1. Introduction

The Bayesian paradigm provides a coherent platform to frame the beliefs and the preferences of decision makers (DMs). Once a DM has specified these in the form of a probability distribution and a utility function, then under the subjective expected utility paradigm she would act rationally by choosing a decision that maximizes her expected utility, i.e., the expectation of the utility function with respect to the probability distribution elicited from her. Although other paradigms expressing different canons of rationality exist (e.g., Giang and Shenoy 2005, Hong and Choi 2000, Smets 2002), applied decision making problems have been most commonly addressed within this Bayesian framework (Gómez 2004, Heckerman et al. 1995).

One of the reasons behind the widespread use of Bayesian methods is the existence of formally justifiable methods that can be used to decompose utility functions and probability distributions into several others, each of which has a smaller dimension than those of a naive representation of the problem. This decomposition offers both computational advantages and more focused decision making, since the DM needs only to elicit beliefs on small dimensional subsets of variables. This in turn has led to larger and larger problems being successfully and accurately modelled within this Bayesian framework.

The decomposition of the probabilistic part of the world is usually achieved via the notion of conditional independence (Dawid 1979). It was long ago recognized that graphical representations of the relationships between random variables directly express a collection of conditional independences. These independences enabled large dimensional joint probabilities to be formally written as products of local distributions of smaller dimension, needing many fewer probability specifications than a direct, full specification. Many formal statistical graphical models were subsequently defined, most notably Bayesian networks (BNs) (Pearl 1988, Smith 2010), that exploited these conditional independences to represent the qualitative structure of a multivariate random vector through a directed graph.

There are also many independence concepts related to utility that can be used to factorize a utility function into terms with a smaller number of arguments. Standard independence concepts are based on the notion of (generalized) additive independence and (conditional) utility independence (Keeney and Raiffa 1993). These both entail some additive or multiplicative decomposition of the utility function. Fairly recently it was recognized that sets of such statements could also be represented by a graph, which in turn could be used to develop fast elicitation routines (see, e.g., Abbas and Howard 2005; Abbas 2009, 2010, 2011; Braziunas and Boutilier 2005; Engel and Wellman 2008; Gonzales and Perny 2004).

The influence diagram (ID; Howard and Matheson 2005, Nielsen and Jensen 2009, Smith and Thwaites 2008) was one of the first graphical methods to contemporaneously depict probabilistic dependence, the form of the utility function, and the structure of the underlying decision space. Fast routines to compute expected utilities and identify optimal decisions that exploit the underlying graph have been defined for a long while (e.g., Jensen et al. 1994, Shachter 1986). However, these are almost exclusively designed to work when the utility can be assumed to factorize additively, i.e., assuming that the utility can be written as a linear combination of smaller dimensional functions over disjoint subsets of the decision problem's attributes. An exception is the multiplicative influence diagram (Leonelli et al. 2017), whose evaluation algorithm works not only for additive factorizations but also for more general multiplicative ones (Keeney 1974).

In this paper, we develop a class of graphical models that can depict both probabilistic independence and sets of (conditional) utility independence statements expressible by a utility diagram (Abbas 2010). We call these *directed expected utility networks* (*DEUNs*). We here develop two fast algorithms for the computation of expected utilities using these diagrams. The first one applies to any DEUN and consists of a sequential application of a conditional expectation operator, analogous to the chance node removal of Shachter (1986). The second algorithm is valid only for a subset of DEUNs, ones that we call here *decomposable*. After a transformation into a new junction tree representation of the problem, this routine computes the overall expected utility via variable elimination just as in Jensen et al. (1994), but now applied to our much more general family of utilities. We are able to demonstrate that the elimination step in DEUNs almost exactly coincides with that of standard ID's evaluation algorithms. Therefore, both additional theoretical results, as, for example, approximated propagation, and code already available for IDs, designed originally for use with additive utilities, can be fairly straightforwardly generalized to be used in conjunction with a much more general utility structure.

The motivation for this work stems from a decision support system we are currently building to help local authorities evaluate the impacts of different policies in the light of endemic food poverty (Smith et al. 2015a, b). In the initial study by Barons et al. (2016)—to keep the analysis as simple as possible—the underlying preferential structure was assumed to factorize additively, as commonly assumed in ID modelling and many applied decision analyses. Discussions during the elicitation process, however, showed that this assumption was far from ideal in this application. Currently available technology would not enable us to formally perform a decision analysis under the required much milder preferential conditions. We have thus taken on this challenge and developed new algorithms for the computation of expected utilities that enable DMs to perform much more general decision analyses.

The only other attempt in the literature we are aware of to represent utility and probabilistic dependence in a unique graph is the expected utility network of La Mura and Shoham (1999). This is an undirected graphical model with two types of edges to represent probabilistic and preferential dependence. However, this method is built on a nonstandard notion of a conditional utility function. Furthermore, no fast routines for the computation of the associated expected utility have yet been developed using this framework. In contrast, DEUNs are based on commonly used concepts of utility independences characterized by various preference relationships and so directly apply to standard formulations of decision problems.

This paper is structured as follows. In Section 2 we review the Bayesian paradigm for decision making. In Sections 3 and 4 we review independence concepts and their graphical representations for probabilities and utilities, respectively. In Section 5 we define our DEUN graphical model, and in Section 6 we develop algorithms for the computation of the DEUN's expected utilities. Section 7 presents an application of DEUNs to household food security. In Section 8 we demonstrate that our algorithms can be straightforwardly adapted to decision problems described by what we will call a *canonical bidirected expected utility network* (*BDEUN*), which is capable of representing even more flexible utility structures. We conclude in Section 9 with a discussion.

2. Bayesian Decision Making

Let d be a decision within some set $\mathbb{D}$ of available decisions, $n \in \mathbb{N}$ and $[n] = \{1, \ldots, n\}$. Let $\mathbf{Y} = (Y_i)_{i\in[n]}$ be an absolutely continuous random vector including the attributes of the problem, i.e., the arguments over which a utility function u is defined. For a subset $A \subseteq [n]$, we let $\mathbf{Y}_A = (Y_i)_{i\in A}$, $\mathbb{Y}_A = \times_{i\in A} \mathbb{Y}_i$, where $\mathbb{Y}_i$ is the sample space of Y_i, and we denote with y_i and $\mathbf{y}_A$ instantiations of Y_i and $\mathbf{Y}_A$, respectively, $i \in [n]$. Last, let $\mathbf{y}_{[n]} = \mathbf{y}$ and $\mathbb{Y}_{[n]} = \mathbb{Y}$.

In this paper, we assume the utility function u to be continuous and normalized between zero and one so that $u\colon \mathbb{Y}\times\mathbb{D} \to [0,1]$. In addition, we assume that for each attribute Y_i there are two reference values $y_i^0, y_i^* \in \mathbb{Y}_i$ such that $u(y_i^*, \mathbf{y}_{-i}, d) > u(y_i^0, \mathbf{y}_{-i}, d)$ for every $d \in \mathbb{D}$, where, for a set $A \subset [n]$, $\mathbf{y}_{-A} = (y_j)_{j\in[n]\setminus A}$.

The expected utility $\bar{u}(d)$ of a decision $d \in \mathbb{D}$—the expectation of $u(\mathbf{y}, d)$ with respect to a probability density $p(\mathbf{y} \mid d)$—is then

$$\bar{u}(d) = \mathbb{E}(u(\mathbf{y}, d)) = \int_{\mathbb{Y}} u(\mathbf{y}, d) p(\mathbf{y} \mid d)\, \mathrm{d}\mathbf{y}. \tag{1}$$

A rational decision maker would then choose to enact an *optimal* decision $d^* = \arg\max_{d\in\mathbb{D}}\{\bar{u}(d)\}$.

This framework, though conceptually straightforward, can become very challenging to apply in practice. As soon as the number of attributes grows even moderately, a faithful elicitation of the probability and utility functions becomes prohibitive. In addition to the knowledge issues in eliciting multivariate functions, the computation of the expected utility in Equation (1) requires an integration over an arbitrary large space $\mathbb{Y}$, which again may become infeasible in high-dimensional settings. For these two reasons, various additional models and independence conditions have been imposed. We review these types of conditions in the next two sections.

For ease of notation in the following, we leave implicit the dependence of all arguments of functions of interest on the decision $d \in \mathbb{D}$. On one hand we can assume that both the probabilistic and the utility independence structure are invariant to the choice of $d \in \mathbb{D}$. This is an assumption commonly made in standard influence diagram modelling. Now $p(\mathbf{y} \mid d)$ and $u(\mathbf{y}, d)$ may well be functions of $d \in \mathbb{D}$—we simply assume that the underlying conditional independence structure and preferential independences are shared by all $d \in \mathbb{D}$. But for any finite discrete space $\mathbb{D}$, we could alternatively apply our methods under the more general assumption that, for each $d \in \mathbb{D}$, the DM's problem could be depicted by a possibly different network. We could then apply the theory we develop below to each of these networks in turn and finally optimize over these separate evaluations.

3. Probability Factorizations

The concept used in probabilistic modelling to simplify density functions is *conditional independence* (Dawid 1979). For three random variables Y_i, Y_j, and Y_k with strictly positive joint density, we say that Y_i is conditional independent of Y_j given Y_k, and write $Y_i \perp\!\!\!\perp Y_j \mid Y_k$ if the conditional density of Y_i can be written as a function of Y_i and Y_k only, i.e., $p(y_i \mid y_j, y_k) = p(y_i \mid y_k)$. This means that the only information to infer Y_i from Y_j and Y_k is from Y_k.

Sets of conditional independence statements can then be depicted by a graph whose vertices are associated to the random variables of interest. We next briefly introduce some terminology from graph theory and then define one of the most common statistical graphical models, namely, the *Bayesian network*.

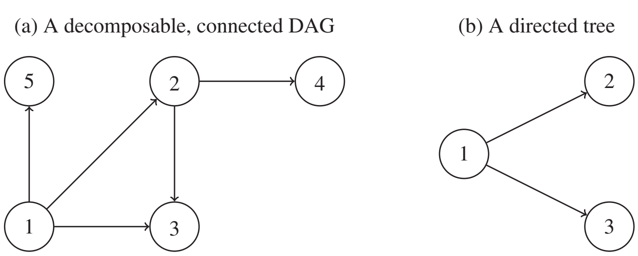
3.1. Graph Theory

A directed *graph* $\mathcal{G}$ is a pair $\mathcal{G} = (V(\mathcal{G}), E(\mathcal{G}))$, where $V(\mathcal{G})$ is a finite set of *vertices*, and $E(\mathcal{G})$ is a set of ordered pairs of vertices, called *edges*. A directed *path* of length m from i_1 to i_m in a graph $\mathcal{G}$ is a sequence of m vertices such that for any two consecutive vertices i_j and i_{j+1} in the sequence, $(i_j, i_{j+1}) \in E(\mathcal{G})$. If there is a directed path from i to j in $\mathcal{G}$, we write $i \to j$. We use the symbol $i \nrightarrow j$ if there is no such directed path in $\mathcal{G}$. Conversely, an undirected path is a sequence of vertices such that either $(i_j, i_{j+1}) \in E(\mathcal{G})$ or $(i_{j+1}, i_j) \in E(\mathcal{G})$. A *cycle* is a directed path with the additional condition that $i_1 = i_m$. For $i, j \in V(\mathcal{G})$, we say that i and j are connected if there is an undirected path between i and j. A graph $\mathcal{G}$ is *connected* if every pair of vertices $i, j \in V(\mathcal{G})$ are connected. A *directed acyclic graph* (*DAG*) is a directed graph with no cycles. For these graphs the labelling of the vertices can be constructed, not uniquely, so that $i < j$ if $(i, j) \in E(\mathcal{G})$.

Now let $\mathcal{G}$ be a DAG. If $(i, j) \in E(\mathcal{G})$ we say that i is a *parent* of j and that j is a *child* of i. The set of parents of i is denoted by Π_i. A vertex of a DAG with no children is called *leaf*, while a *root* is a vertex with no parents. A DAG is said to be *decomposable* if all pairs of parents of the same child are joined by an edge. A subset C of $V(\mathcal{G})$ is a *clique* of $\mathcal{G}$ if any pair $i, j \in C$ is connected by an edge and there is no other $C' \subseteq V(\mathcal{G})$ with the same property such that $C \subset C'$. Let $\mathcal{G}$ have m cliques $\{C_1, \dots, C_m\} = \mathcal{C}$, and suppose the elements of $\mathcal{C}$ are ordered according to their indexing. A *separator* S_i of $\mathcal{G}$, $i \in [m]\backslash\{1\}$, is defined as $S_i = C_i \cap \bigcup_{j=1}^{i-1} C_j$. The cliques of $\mathcal{G}$ are said to respect the running intersection property if $S_i \subseteq C_j$ for at least one $j < i$, $i \in [m]\backslash\{1\}$.

Example 1. The directed graph in Figure 1(a) can be clearly seen to be a DAG with a vertex set equal to [5]. This is decomposable since the two parents of vertex 3, i.e., 1 and 2, are connected by an edge. This DAG is also connected since every two vertices are connected by an undirected path. The cliques of the DAG in Figure 1(a) are $C_1 = \{1, 2, 3\}$, $C_2 = \{2, 4\}$, and $C_3 = \{1, 5\}$ and its separators $S_2 = \{2\}$ and $S_3 = \{1\}$. So with this indexing, the cliques of this DAG respect the running intersection property.

Figure 1. Example of Two DAGs



A graph of interest in this paper is the *directed tree* $\mathcal{T}$. This is a DAG with the following two properties: it has a unique vertex with no parents called *root*, and all other vertices have exactly one parent. The DAG in Figure 1(b) can be clearly seen to be a directed tree with root 1 and leaves 2 and 3.

3.2. Bayesian Networks

We are now ready to define the statistical graphical model that underpins the probabilistic part of the DEUN model we define below.

Definition 1. A *BN model for a random vector* $\mathbf{Y} = (Y_i)_{i\in[n]}$ is defined by

- $n-1$ conditional independence statements of the form $Y_i \perp\!\!\!\perp \mathbf{Y}_{[i-1]\backslash\Pi_i} \mid \mathbf{Y}_{\Pi_i}$, where $\Pi_i \subseteq [i-1]$;
- a DAG $\mathcal{G}$ with vertex set $V(\mathcal{G}) = [n]$ and edge set $E(\mathcal{G}) = \{(i, j): j \in [n], i \in \Pi_j\}$;
- conditional distributions $p(y_i \mid \mathbf{y}_{\Pi_i})$ for $i \in [n]$.

It can be shown (e.g., Lauritzen 1996) that the density of a BN can then be written as

$$p(\mathbf{y}) = \prod_{i\in[n]} p(y_i \mid \mathbf{y}_{\Pi_i}).$$

Example 2. Consider the DAG in Figure 1(a). A BN with this associated graph implies the conditional independences $Y_4 \perp\!\!\!\perp (Y_1, Y_3) \mid Y_2$ and $Y_5 \perp\!\!\!\perp (Y_2, Y_3, Y_4) \mid Y_1$. The probability distribution then factorizes as $p(\mathbf{y}) = p(y_5 \mid y_1)p(y_4 \mid y_2)p(y_3 \mid y_1, y_2)p(y_2 \mid y_1)p(y_1)$.

4. Utility Factorizations

4.1. Independence and Factorizations

While conditional independence is universally acknowledged as the gold standard to simplify probabilistic joint densities, for utility functions, a variety of independence concepts have been used. One very common assumption is that a utility has additively independent attributes implying the additive utility factorization

$$u(\mathbf{y}) = \sum_{i\in[n]} k_i u(y_i), \tag{2}$$

where $k_i = u(y_i^*, \mathbf{y}_{-i}^0)$ is a *criterion weight* and $u(y_i) = u(y_i, \mathbf{y}_{-i}^*) = u(y_i, \mathbf{y}_{-i}^0)$, $i \in [n]$. A generalization of this independence concept applies to subsets of $[n]$ that are possibly nondisjoint (Braziunas and Boutilier 2005, Fishburn 1967).

A second approach for defining multivariate utility factorizations is to first identify *utility independences*. For this purpose we introduce the *conditional utility function* of $\mathbf{y}_A$ given $\mathbf{y}_{-A}$, $A \subset [n]$,

$$u(\mathbf{y}_A \mid \mathbf{y}_{-A}) = \frac{u(\mathbf{y}) - u(\mathbf{y}_A^0, \mathbf{y}_{-A})}{u(\mathbf{y}_A^*, \mathbf{y}_{-A}) - u(\mathbf{y}_A^0, \mathbf{y}_{-A})},$$

where $\mathbf{y}_A^0 = (y_i^0)_{i\in A}$ and $\mathbf{y}_A^* = (y_i^*)_{i\in A}$.

Definition 2. We say that $\mathbf{Y}_A$ *is utility independent of* $\mathbf{Y}_B$ *given* $\mathbf{Y}_C$, denoted $\mathbf{Y}_A$ UI $\mathbf{Y}_B \mid \mathbf{Y}_C$, for $A \cup B \cup C = [n]$, if and only if

$$u(\mathbf{y}_A \mid \mathbf{y}_B, \mathbf{y}_C) = u(\mathbf{y}_A \mid \mathbf{y}_C).$$

Utility independences then imply joint utility functions that have a simpler form. Let $A \subseteq [n]$ be a totally ordered set, and let, for each $i \in A$, iP and iF be the set of indices that precede and follow i in A, respectively. Let $\mathbb{Y}_A^{0*}$ be the set comprising all possible instantiations of $\mathbf{Y}_A$, where each element is either y_i^0 or y_i^*, $i \in A$, and let $\mathbf{y}_A^{0*}$ be an element of $\mathbb{Y}_A^{0*}$. Abbas (2010) showed that by sequentially applying conditional utility independence statements according to the order of the elements in A, any utility function can then be written as

$$u(\mathbf{y}) = \sum_{\mathbf{y}_A^{0*}\in\mathbb{Y}_A^{0*}} u(\mathbf{y}_A^{0*}, \mathbf{y}_{-A}) \prod_{i\in A} g(y_i \mid \mathbf{y}_{iP}^{0*}, \mathbf{y}_{iF}), \qquad (3)$$

where

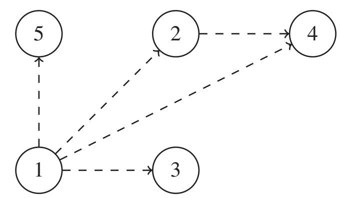
$$g(y_i \mid \mathbf{y}_{iP}^{0*}, \mathbf{y}_{iF}) = \begin{cases} u(y_i \mid \mathbf{y}_{iP}^{0*}, \mathbf{y}_{iF}), & \text{if } y_i = y_i^* \text{ in } u(\mathbf{y}_A^{0*}, \mathbf{y}_{-A}), \\ \hat{u}(y_i \mid \mathbf{y}_{iP}^{0*}, \mathbf{y}_{iF}), & \text{otherwise,} \end{cases}$$

and $\hat{u}(y_i \mid \mathbf{y}_{iP}^{0*}, \mathbf{y}_{iF}) = 1 - u(y_i \mid \mathbf{y}_{iP}^{0*}, \mathbf{y}_{iF})$ is the *disutility function*. So, for example, if each Y_i is utility independent of $\mathbf{Y}_{-i}$, then Equation (3) can be reexpressed as

$$u(\mathbf{y}) = \sum_{\mathbf{y}^{0*}\in\mathbb{Y}^{0*}} u(\mathbf{y}^{0*}) \prod_{i\in[n]} g(y_i \mid \mathbf{y}_{-i}^0). \qquad (4)$$

This special case can be identified as the well-known multilinear utility factorization Keeney and Raiffa (1993).

Figure 2. Example of a Directional Utility Diagram



4.2. Utility Diagrams

Graphical models depicting various types of preferential independences have now begun to appear. In this paper we consider a specific class of models called *utility diagrams* (Abbas 2010).

Definition 3. A *utility diagram* is a directed graph with vertex set $[n]$, and its edge set is such that the absence of an edge (i, j), $i, j \in [n]$, implies Y_j UI $Y_i \mid \mathbf{Y}_{-ij}$.

Note that Abbas (2010) defined utility diagrams as bidirectional graphs. However, given that our definition of a directed graph allows vertices to be connected by more than one edge, the model in Definition 3 is equivalent to the one of Abbas (2010) where a bidirected edge between two vertices is replaced by two edges, one pointing in each direction.

A utility diagram with an empty edge set corresponds to a multilinear factorization of the utility function as in Equation (4). Here we introduce a subclass of utility diagrams that has some important properties.

Definition 4. A utility diagram is said to be *directional* if its graph is a DAG.

Example 3. The utility diagram in Figure 2 is directional and implies the following conditional utility independences:

$$\begin{array}{ll} Y_1 \text{ UI } Y_2 \mid Y_3, Y_4, Y_5, & Y_1 \text{ UI } Y_3 \mid Y_2, Y_4, Y_5, \\ Y_1 \text{ UI } Y_4 \mid Y_2, Y_3, Y_5, & Y_1 \text{ UI } Y_5 \mid Y_2, Y_3, Y_4, \\ Y_2 \text{ UI } Y_3 \mid Y_1, Y_4, Y_5, & Y_2 \text{ UI } Y_4 \mid Y_1, Y_3, Y_5, \\ Y_2 \text{ UI } Y_5 \mid Y_1, Y_3, Y_4, & Y_3 \text{ UI } Y_2 \mid Y_1, Y_4, Y_5, \\ Y_3 \text{ UI } Y_4 \mid Y_1, Y_2, Y_5, & Y_3 \text{ UI } Y_5 \mid Y_1, Y_2, Y_4, \\ Y_4 \text{ UI } Y_3 \mid Y_1, Y_2, Y_5, & Y_4 \text{ UI } Y_5 \mid Y_1, Y_2, Y_3, \\ Y_5 \text{ UI } Y_2 \mid Y_1, Y_3, Y_4, & Y_5 \text{ UI } Y_3 \mid Y_1, Y_2, Y_4, \\ Y_5 \text{ UI } Y_4 \mid Y_1, Y_2, Y_3. & \end{array}$$

Directional utility diagrams have the unique property that their utility function can be written in terms of criterion weights and univariate utility functions only.

Although not explicitly depicted by a utility graph, such a property underlies the algorithms developed in Leonelli and Smith (2015) that apply to some specific generalized additively independent models only.

Lemma 1. *For a directional utility diagram there exists an expansion order over* $[n]$ *such that Equation* (3) *is a linear combination of terms involving only criterion weights and conditional utility functions having as argument a single attribute.*

This result follows by observing that the terms $u(\mathbf{y}_I^{0*}, \mathbf{y}_{-I})$ in Equation (3) coincide with $u(\mathbf{y}^{0*})$ since the expansion can be performed over all the attributes. These terms are functions of criterion weights. Furthermore, the conditional independence structure underlying a directed utility diagram is such that there is an expansion order where $\mathbf{Y}_i$ UI $Y_{iF} \mid \mathbf{Y}_{iP}$. Thus, $g(y_i \mid \mathbf{y}_{iP}^{0*}, \mathbf{y}_{iF})$ in Equation (3) is equal to $g(y_i \mid \mathbf{y}_{iP}^{0*})$ for every $i \in [n]$.

Example 4. The utility factorization associated to the diagram in Figure 2 is equal to the sum of the entries of Table 1, where r_I, $I \subseteq [n]$, is equal to $u(\mathbf{y}_I^*, \mathbf{y}_{-I}^0)$. This in general consists of 2^n terms each made of $n+1$ indeterminates, where n is number of vertices in the diagram.

The subclass of directed utility diagrams has the great computational advantage of enabling the computation of expected utilities through an associated backward inductive routine that at each step computes a finite number of integrals of univariate functions. More general utility dependence structures are examined in Section 8, where we adapt one of the algorithms for DEUNs to graphs embedding nondirectional utility diagrams.

5. Directed Expected Utility Networks

We are now ready to define our graphical model, which embeds both probabilistic and utility independence statements.

Definition 5. A *directed expected utility network* $\mathcal{G}$ consists of a set of vertices $V(\mathcal{G}) = [n]$, a probabilistic edge set $E_p(\mathcal{G})$, denoted by solid arrows, and a utility edge set $E_u(\mathcal{G})$, denoted by dashed arrows, such that

- $(V(\mathcal{G}), E_p(\mathcal{G}))$ is a BN model;
- $(V(\mathcal{G}), E_u(\mathcal{G}))$ is a directional utility diagram such that if $(i, j) \in E_u(\mathcal{G})$ then $j \nrightarrow i$ in $(V(\mathcal{G}), E_p(\mathcal{G}))$.

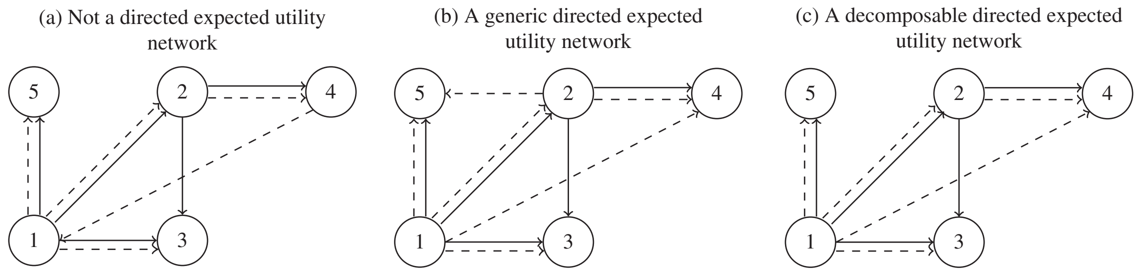
Example 5. Consider the diagrams in Figure 3. Figure 3(a) includes a graph that is not a DEUN since there is a utility edge from 4 to 1. This edge would make the computation of expected utilities via backward induction impossible. Figures 3(b) and (c) are DEUNs since for these $(V(\mathcal{G}), E_p(\mathcal{G}))$ is a BN and $(V(\mathcal{G}), E_u(\mathcal{G}))$ is a directed utility diagram, both including only edges (i, j) such that $j \nrightarrow i$ in $(V(\mathcal{G}), E_p(\mathcal{G}))$. Note that all three diagrams embed the BN in Figure 1(a), while only the diagram in Figure 3(c) embeds the utility diagram in Figure 2.

We next introduce a subclass of DEUNs that entails fast computation routines.

Table 1. Terms in the Utility Expansion Associated to the Utility Diagram in Figure 2

$r_{\varnothing}\hat{u}(y_1)\hat{u}(y_2 \mid y_1^0)\hat{u}(y_3 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^0)\hat{u}(y_5 \mid y_1^0)$	$r_1 u(y_1)\hat{u}(y_2 \mid y_1^*)\hat{u}(y_3 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^0)\hat{u}(y_5 \mid y_1^*)$
$r_2\hat{u}(y_1)u(y_2 \mid y_1^0)\hat{u}(y_3 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^*)\hat{u}(y_5 \mid y_1^0)$	$r_3\hat{u}(y_1)\hat{u}(y_2 \mid y_1^0)u(y_3 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^0)\hat{u}(y_5 \mid y_1^0)$
$r_4\hat{u}(y_1)\hat{u}(y_2 \mid y_1^0)\hat{u}(y_3 \mid y_1^0)u(y_4 \mid y_1^0, y_2^0)\hat{u}(y_5 \mid y_1^0)$	$r_5\hat{u}(y_1)\hat{u}(y_2 \mid y_1^0)\hat{u}(y_3 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^0)u(y_5 \mid y_1^0)$
$r_{12}u(y_1)u(y_2 \mid y_1^*)\hat{u}(y_3 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^*)\hat{u}(y_5 \mid y_1^*)$	$r_{13}u(y_1)\hat{u}(y_2 \mid y_1^*)u(y_3 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^0)\hat{u}(y_5 \mid y_1^*)$
$r_{14}u(y_1)\hat{u}(y_2 \mid y_1^*)\hat{u}(y_3 \mid y_1^*)u(y_4 \mid y_1^*, y_2^0)\hat{u}(y_5 \mid y_1^*)$	$r_{15}u(y_1)\hat{u}(y_2 \mid y_1^*)\hat{u}(y_3 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^0)u(y_5 \mid y_1^*)$
$r_{23}\hat{u}(y_1)u(y_2 \mid y_1^0)u(y_3 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^*)\hat{u}(y_5 \mid y_1^0)$	$r_{24}\hat{u}(y_1)u(y_2 \mid y_1^0)\hat{u}(y_3 \mid y_1^0)u(y_4 \mid y_1^0, y_2^*)\hat{u}(y_5 \mid y_1^0)$
$r_{25}\hat{u}(y_1)u(y_2 \mid y_1^0)\hat{u}(y_3 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^*)u(y_5 \mid y_1^0)$	$r_{34}\hat{u}(y_1)\hat{u}(y_2 \mid y_1^0)u(y_3 \mid y_1^0)u(y_4 \mid y_1^0, y_2^0)\hat{u}(y_5 \mid y_1^0)$
$r_{35}\hat{u}(y_1)\hat{u}(y_2 \mid y_1^0)u(y_3 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^0)u(y_5 \mid y_1^0)$	$r_{45}\hat{u}(y_1)\hat{u}(y_2 \mid y_1^0)\hat{u}(y_3 \mid y_1^0)u(y_4 \mid y_1^0, y_2^0)u(y_5 \mid y_1^0)$
$r_{123}u(y_1)u(y_2 \mid y_1^*)u(y_3 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^*)\hat{u}(y_5 \mid y_1^*)$	$r_{124}u(y_1)u(y_2 \mid y_1^*)\hat{u}(y_3 \mid y_1^*)u(y_4 \mid y_1^*, y_2^*)\hat{u}(y_5 \mid y_1^*)$
$r_{125}u(y_1)u(y_2 \mid y_1^*)\hat{u}(y_3 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^*)u(y_5 \mid y_1^*)$	$r_{134}u(y_1)\hat{u}(y_2 \mid y_1^*)u(y_3 \mid y_1^*)u(y_4 \mid y_1^*, y_2^0)\hat{u}(y_5 \mid y_1^*)$
$r_{135}u(y_1)\hat{u}(y_2 \mid y_1^*)u(y_3 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^0)u(y_5 \mid y_1^*)$	$r_{145}u(y_1)\hat{u}(y_2 \mid y_1^*)\hat{u}(y_3 \mid y_1^*)u(y_4 \mid y_1^*, y_2^0)u(y_5 \mid y_1^*)$
$r_{234}\hat{u}(y_1)u(y_2 \mid y_1^0)u(y_3 \mid y_1^0)u(y_4 \mid y_1^0, y_2^*)\hat{u}(y_5 \mid y_1^0)$	$r_{235}\hat{u}(y_1)u(y_2 \mid y_1^0)u(y_3 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^*)u(y_5 \mid y_1^0)$
$r_{245}\hat{u}(y_1)u(y_2 \mid y_1^0)\hat{u}(y_3 \mid y_1^0)u(y_4 \mid y_1^0, y_2^*)u(y_5 \mid y_1^0)$	$r_{345}\hat{u}(y_1)\hat{u}(y_2 \mid y_1^0)u(y_3 \mid y_1^0)u(y_4 \mid y_1^0, y_2^0)u(y_5 \mid y_1^0)$
$r_{1234}u(y_1)u(y_2 \mid y_1^*)u(y_3 \mid y_1^*)u(y_4 \mid y_1^*, y_2^*)\hat{u}(y_5 \mid y_1^*)$	$r_{1235}u(y_1)u(y_2 \mid y_1^*)u(y_3 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^*)u(y_5 \mid y_1^*)$
$r_{1245}u(y_1)u(y_2 \mid y_1^*)\hat{u}(y_3 \mid y_1^*)u(y_4 \mid y_1^*, y_2^*)u(y_5 \mid y_1^*)$	$r_{1345}u(y_1)\hat{u}(y_2 \mid y_1^*)u(y_3 \mid y_1^*)u(y_4 \mid y_1^*, y_2^0)u(y_5 \mid y_1^*)$
$r_{2345}\hat{u}(y_1)u(y_2 \mid y_1^0)u(y_3 \mid y_1^0)u(y_4 \mid y_1^0, y_2^*)u(y_5 \mid y_1^0)$	$r_{12345}u(y_1)u(y_2 \mid y_1^*)u(y_3 \mid y_1^*)u(y_4 \mid y_1^*, y_2^*)u(y_5 \mid y_1^*)$

Figure 3. Graphical Representations of Probabilistic and Utility Independences



Definition 6. A DEUN is said to be *decomposable* if

- $(V(\mathcal{G}), E_p(\mathcal{G}))$ is decomposable;
- $(i,j) \in E_u(\mathcal{G})$ only if $i \longrightarrow j$ in $(V(\mathcal{G}), E_p(\mathcal{G}))$.

Example 6. The DEUN in Figure 3(b) is not decomposable since $(2,5) \in E_u(\mathcal{G})$ but these two vertices are not connected by a directed path in the underlying BN. Conversely, the network in Figure 3(c) is decomposable. Note that the semantics of our model permit two vertices to be connected by both probabilistic and utility edges, by just one of the two, or potentially none. So, for example, $(1,2) \in E_p(\mathcal{G})$ and $(1,2) \in E_u(\mathcal{G})$, while $(1,4) \notin E_p(\mathcal{G})$ and $(1,4) \in E_u(\mathcal{G})$.

Just as in the triangulation step for probabilistic propagation (e.g., Lauritzen 1996), it can be easily shown that any nondecomposable DEUN can be transformed into a decomposable one.

Proposition 1. *Let $\mathcal{G}$ be a nondecomposable DEUN with vertex set $V(\mathcal{G})$ and edges $E_u(\mathcal{G})$ and $E_p(\mathcal{G})$. Let $\mathcal{G}'$ be a DEUN with vertex set $V(\mathcal{G}') = V(\mathcal{G})$ and edges $E_u(\mathcal{G}') = E_u(\mathcal{G})$ and $E_p(\mathcal{G}') = E_p(\mathcal{G}) \cup B_1 \cup B_2$, where*

$$B_1 = \{(i,j) \in E_u(\mathcal{G}) \colon i \nrightarrow j \text{ in } (V(\mathcal{G}), E_p(\mathcal{G}))\},$$
$$B_2 = \{(i,j) \colon (i,k) \text{ and } (j,k) \in E_p(\mathcal{G}) \cup B_1, k \in V(\mathcal{G})\}.$$

Then $\mathcal{G}'$ is decomposable.

This holds by noting that the set B_1 simply adds a probabilistic edge connecting two vertices linked by a utility edge which breaks the decomposability condition. The set B_2 then simply transforms the graph $(V(\mathcal{G}), E_p(\mathcal{G}) \cup B_1)$ into a decomposable DAG.

Example 7. For the nondecomposable network in Figure 3(b), the decomposability condition is achieved by simply adding $(2,5)$ to $E_p(\mathcal{G})$.

6. Computation of Expected Utilities

We next consider the computation of expected utilities for both nondecomposable and decomposable DEUNs and define algorithms based on backward inductive routines. All these routines have in common an operation applied to vectors of (expected) utility functions that we define next. Let Π_i^u and Π_i^p be the parent sets of i with respect to $E_u(\mathcal{G})$ and $E_p(\mathcal{G})$, respectively. We let $\mathbf{u}_i(y_i \mid \mathbf{y}^{0*}_{\Pi_i^u}) = (u(y_i \mid y^{0*}_{\Pi_i^u}), \hat{u}(y_i \mid y^{0*}_{\Pi_i^u}))_{y^{0*}_{\Pi_i^u} \in \mathbb{Y}^{0*}_{\Pi_i^u}}$ be the vector comprising the conditional utilities and disutilities given all possible combinations of the parents at the reference values and $\mathbf{u}_0(\mathbf{y}^{0*}) = (u(\mathbf{y}^{0*}))_{\mathbf{y}^{0*} \in \mathbb{Y}^{0*}}$.

Example 8. The vector $\mathbf{u}_5(y_5 \mid \mathbf{y}^{0*}_{\Pi_5^u})$ for the DEUN in Figure 3(c) has as its components

$$u(y_5 \mid y_1^*), \quad u(y_5 \mid y_1^0), \quad \hat{u}(y_5 \mid y_1^*), \quad \hat{u}(y_5 \mid y_1^0), \tag{5}$$

while the vector $\mathbf{u}_4(y_4 \mid \mathbf{y}^{0*}_{\Pi_4^u})$ has the utility components

$$\begin{array}{lll} u(y_4 \mid y_1^*, y_2^*), & u(y_4 \mid y_1^0, y_2^*), & u(y_4 \mid y_1^*, y_2^0), \\ u(y_4 \mid y_1^0, y_2^0), & \hat{u}(y_4 \mid y_1^*, y_2^*), & \hat{u}(y_4 \mid y_1^0, y_2^*), \\ \hat{u}(y_4 \mid y_1^*, y_2^0), & \hat{u}(y_4 \mid y_1^0, y_2^0). & \end{array} \tag{6}$$

We next introduce an elementwise operation, denoted by $\circ$, which multiplies an element of one vector, $\mathbf{u}_i(\cdot)$, with any element of another vector, $\mathbf{u}_j(\cdot)$, if these have compatible instantiations, i.e., if the common conditioning variables are instantiated to the same value.

Example 9. Consider the vectors $\mathbf{u}_5(y_5 \mid \mathbf{y}^{0*}_{\Pi_5^u})$ and $\mathbf{u}_4(y_4 \mid \mathbf{y}^{0*}_{\Pi_4^u})$ of Example 8 and suppose we need to compute $\mathbf{u}_5(y_5 \mid \mathbf{y}^{0*}_{\Pi_5^u}) \circ \mathbf{u}_4(y_4 \mid \mathbf{y}^{0*}_{\Pi_4^u})$. The terms $u(y_5 \mid y_1^*)$ and $\hat{u}(y_5 \mid y_1^*)$ in Equation (5) are multiplied to any term in Equation (6) conditioning on y_1^*: namely, $u(y_4 \mid y_1^*, y_2^*)$, $u(y_4 \mid y_1^*, y_2^0)$, $\hat{u}(y_4 \mid y_1^*, y_2^*)$, and $\hat{u}(y_4 \mid y_1^*, y_2^0)$. Conversely, $u(y_5 \mid y_1^0)$ and $\hat{u}(y_5 \mid y_1^0)$ multiply the remaining four entries of $\mathbf{u}_4(y_4 \mid \mathbf{y}^{0*}_{\Pi_4^u})$, i.e., those conditioning on y_1^0. So the product $\mathbf{u}_5(y_5 \mid \mathbf{y}^{0*}_{\Pi_5^u}) \circ \mathbf{u}_4(y_4 \mid \mathbf{y}^{0*}_{\Pi_4^u})$ returns a vector with entries

$$\begin{array}{ll} u(y_5 \mid y_1^*)u(y_4 \mid y_1^*, y_2^*), & u(y_5 \mid y_1^*)u(y_4 \mid y_1^*, y_2^0), \\ u(y_5 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^*), & u(y_5 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^0), \\ \hat{u}(y_5 \mid y_1^*)u(y_4 \mid y_1^*, y_2^*), & \hat{u}(y_5 \mid y_1^*)u(y_4 \mid y_1^*, y_2^0), \\ \hat{u}(y_5 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^*), & \hat{u}(y_5 \mid y_1^*)\hat{u}(y_4 \mid y_1^*, y_2^0), \\ u(y_5 \mid y_1^0)u(y_4 \mid y_1^0, y_2^*), & u(y_5 \mid y_1^0)u(y_4 \mid y_1^0, y_2^0), \\ u(y_5 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^*), & u(y_5 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^0), \\ \hat{u}(y_5 \mid y_1^0)u(y_4 \mid y_1^0, y_2^*), & \hat{u}(y_5 \mid y_1^0)u(y_4 \mid y_1^0, y_2^0), \\ \hat{u}(y_5 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^*), & \hat{u}(y_5 \mid y_1^0)\hat{u}(y_4 \mid y_1^0, y_2^0). \end{array}$$

If the vertices i and j are such that $\Pi_i^u \cap \Pi_j^u = \varnothing$ and $\mathbf{u}_i(\cdot)$ and $\mathbf{u}_j(\cdot)$ include, respectively, n_i and n_j elements, then $\mathbf{u}_i(\cdot) \circ \mathbf{u}_j(\cdot)$ returns a vector of $n_i \times n_j$ entries consisting of all possible multiplications between elements of the vectors. This operation can be encoded by defining the vectors to have elements appropriately ordered so that the standard elementwise multiplication returns only terms having compatible instantiations, just as in Leonelli et al. (2017).

6.1. Computations in Generic Directed Expected Utility Networks

The expected utility associated to any DEUN can now be computed via a backward induction, which at each step computes a conditional expectation, just as in the chance node removal step of Shachter (1986). This is formalized in the following theorem. Let, for a vector $\mathbf{a}$, $|\mathbf{a}|$ denote the sum of its elements.

Theorem 1. *The expected utility score $\bar{u}$ associated to a DEUN $\mathcal{G}$ can be computed according to the following algorithm*:

1. *compute*:
$$\bar{\mathbf{u}}_n = \int_{\mathbb{Y}_n} \mathbf{u}_n(y_n \mid \mathbf{y}^{0*}_{\Pi_n^u}) p(y_n \mid \mathbf{y}_{\Pi_n^p})\, \mathrm{d}y_n, \qquad (7)$$
2. *for i from $n-1$ to 1, compute*:
$$\bar{\mathbf{u}}_i = \int_{\mathbb{Y}_i} (\bar{\mathbf{u}}_{i+1} \circ \mathbf{u}_i(y_i \mid \mathbf{y}^{0*}_{\Pi_i^u})) p(y_i \mid \mathbf{y}_{\Pi_i^p})\, \mathrm{d}y_i, \qquad (8)$$
3. *return*:
$$\bar{u} = |\mathbf{u}_0(\mathbf{y}^{0*}) \circ \bar{\mathbf{u}}_1|. \qquad (9)$$

The above algorithm can be applied directly to any DEUN and computes expected utilities relatively fast and in a distributed fashion by marginalization of individual random variables. However, we also notice that the speed of such a routine can be improved since the computation and transmission of terms that it uses are not strictly necessary. To see this, consider the network in Figure 3(c). The algorithm starts from vertex 5 and computes a marginalization of $\mathbf{u}(y_5 \mid y_1)$ with respect to the density $p(y_5 \mid y_1)$. The result of this operation, $\bar{\mathbf{u}}_5$ is then a function of y_1 only. In the algorithm in Theorem 1, $\bar{\mathbf{u}}_5$ is then passed to 4, and a marginalization with respect to density $p(y_4 \mid y_2)$ is computed over $\bar{\mathbf{u}}_5 \circ \mathbf{u}_4(y_4 \mid y_2, y_1)$. But $\bar{\mathbf{u}}_5$ is not a function of y_4. It therefore does not carry any information about this variable that would need to be formally accounted for during its marginalization. Furthermore, since $\mathbf{u}_4(y_4 \mid y_2, y_1)$ is a function of not only y_1 but also y_2, the $\circ$ product computes a potentially very large number of terms that are not relevant at this stage of the evaluation. This inefficiency becomes even larger for nonconnected networks, since the contribution of each of the components can be collated together at the very end of the evaluation. This is because the only joint information these provide lies in the terms $u(\mathbf{y}^{0*})$.

6.2. Computations in Decomposable Directed Expected Utility Networks

To address these inefficiencies, we introduce next a much faster algorithm that works over a transformation of the original graph into a tree structure, just as in standard junction tree algorithms (see, e.g., Jensen et al. 1994, for BNs and IDs). Let $\mathscr{C} = \{C_i : i \in [m]\}$ be the cliques of the DAG $(V(\mathcal{G}), E_p(\mathcal{G}))$, let $S_2, \ldots, S_m$ be its separators, and assume the cliques are ordered to respect the running intersection property.

Definition 7. We call the *junction tree* of a decomposable DEUN $\mathcal{G}$ the directed tree $\mathcal{T}$ with vertex set $V(\mathcal{T}) = \mathscr{C}$ and edges (C_i, C_j) for one $i \in [j-1]$ such that $S_j \subseteq C_i$, $j \in [m]$.

Note that to construct such a tree we can apply in a straightforward way any of the algorithms already devised for both BNs and IDs (see, e.g., Cowell et al. 2007). Furthermore, as for BNs and IDs, a DEUN can have more than one junction tree representation.

Example 10. The junction tree associated to the DEUN in Figure 3(c) is shown in Figure 4.

In contrast to an algorithm based directly on Theorem 1, we instead propagate using "potentials," just as in many propagation algorithms of BNs and IDs. This enables us to demonstrate that our evaluation algorithm mirrors those commonly used to compute expected utilities in IDs, but now for utility functions that are not necessarily additive.

Recall that a potential Φ_A, $A \subset [n]$, is a function $\Phi_A\colon \mathbb{Y}_A \to \mathbb{R}$. Just as for IDs, we have two types of potentials: utility and probability potentials. For a clique $C \in \mathscr{C}\setminus\{C_1\}$ with an associated separator S, its probability potential Φ_C and its utility potential Ψ_C are defined as

$$\Phi_C = \prod_{i \in C\setminus S} p(y_i \mid \mathbf{y}_{\Pi_i^p}), \quad \Psi_C = \circ_{i \in C\setminus S}\mathbf{u}(y_i \mid \mathbf{y}^{0*}_{\Pi_i^u}),$$

and $\Phi_{C_1} = \prod_{i \in C_1} = p(y_i \mid \mathbf{y}_{\Pi_i^p})$ and $\Psi_{C_1} = \mathbf{u}(\mathbf{y}^{0*}) \circ \circ_{i \in C_1}\mathbf{u}(y_i \mid \mathbf{y}^{0*}_{\Pi_i^u})$. Call $\Phi_{\mathscr{T}} = \prod_{C \in \mathscr{C}} \Phi_C$ and $\Psi_{\mathscr{T}} = |\circ_{C \in \mathscr{C}} \Psi_C|$ and note that $p(\mathbf{y}) = \Phi_{\mathscr{T}}$ and $u(\mathbf{y}) = \Psi_{\mathscr{T}}$.

Now let C_i be the parent of C_j in $\mathscr{T}$. We say that C_i absorbs C_j if the utility potential of C_i, Ψ_{C_i}, maps to $\Psi_{C_i}^{C_j}$ where

$$\Psi_{C_i}^{C_j} = \Psi_{C_i} \circ \int_{\mathbb{Y}_{C_j\setminus S_j}} \Psi_{C_j}\Phi_{C_j}\,\mathrm{d}\mathbf{y}_{C_j\setminus S_j}. \tag{10}$$

For a leaf L of $\mathscr{T}$, call $\Phi_{\mathscr{T}\setminus L}$ and $\Psi_{\mathscr{T}\setminus L}$ the probability and utility potentials, respectively, of the junction tree obtained by absorbing L into its parent and removing L from $\mathscr{T}$.

Theorem 2. *After absorption of a leaf L with separator S into its parent, we have*

$$\int_{\mathbb{Y}_{L\setminus S}} \Phi_{\mathscr{T}}\Psi_{\mathscr{T}}\,\mathrm{d}\mathbf{y}_{L\setminus S} = \Phi_{\mathscr{T}\setminus L}\Psi_{\mathscr{T}\setminus L}.$$

Theorem 2 provides the basic step for computing the expected utility of a decomposable DEUN. Suppose the junction tree is connected. Then, by sequentially absorbing leaves into parents (for example, by following in reverse order the indices of the cliques), we obtain a tree consisting of a vertex/clique only, coinciding with the initial root of the junction tree. Let $\Psi_{C_1}^{C_2}$ be its utility potential resulting from the absorption of all the other cliques, assuming C_2 was the last clique to be absorbed. It then follows that the overall expected utility $\bar{u}$ is given by

$$\bar{u} = \int_{\mathbb{Y}_{C_1}} \Phi_{C_1}|\Psi_{C_1}^{C_2}|\,\mathrm{d}\mathbf{y}_{C_1} = \left|\int_{\mathbb{Y}_{C_1}} \Phi_{C_1}\Psi_{C_1}^{C_2}\,\mathrm{d}\mathbf{y}_{C_1}\right|. \tag{11}$$

If, on the other hand, the junction tree is not connected, and this is the case whenever the DEUN is not connected, $\bar{u}$ simply equals the $\circ$ product of the contributions of the roots of each nonconnected component after all other vertices have been absorbed. More formally, let $R_1, \ldots, R_k$ be the roots of the nonconnected components of the junction tree, and let $\Psi_{R_i}^{C_i}$ be their utility potentials resulting from the absorption of all other cliques, where C_i was the last child of R_i to be absorbed, for $i \in [k]$. We then have that

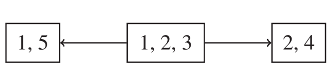
$$\bar{u} = \left|\int_{\mathbb{Y}_{R_1}} \Phi_{R_1}\Psi_{R_1}^{C_1}\,\mathrm{d}\mathbf{y}_{R_1} \circ \cdots \circ \int_{\mathbb{Y}_{R_k}} \Phi_{R_k}\Psi_{R_k}^{C_k}\,\mathrm{d}\mathbf{y}_{R_k}\right|.$$

It is interesting to highlight that the junction tree evaluations of DEUNs and IDs follow the same backward inductive routine, formalized in Theorem 2, which sequentially absorbs a leaf of the tree. Given our definition of the clique potentials, this absorption for DEUNs entails an updating of the utility potential only, which consists of an $\circ$ product. In contrast, for standard IDs, this operation corresponds to a simple sum. To see this, suppose that for an ID with m cliques, the utility potential for C_i, $i = 2, \ldots, m$, is $\Psi_{C_i} = \sum_{j \in C_i\setminus S_i} k_j u(y_j)$, and that for C_1 is $\Psi_{C_1} = \sum_{j \in C_1} k_j u(y_j)$. Let $\Psi_{\mathscr{T}} = \sum_{C \in \mathscr{C}} \Psi_C$. The absorption of a clique C_j, supposing C_j includes only chance nodes, into its parent C_i in an ID with these potentials then transforms Ψ_{C_i} into

$$\Psi_{C_i} + \int_{\mathbb{Y}_{C_j\setminus S_j}} \Psi_{C_j}\Phi_{C_j}\,\mathrm{d}\mathbf{y}_{C_j}. \tag{12}$$

Equation (12) can be seen to be almost identical to Equation (10), which specifies the absorption step in DEUNs. The only difference lies in the different operation: a sum for IDs and a $\circ$ product for DEUNs.

Figure 4. Junction Tree Representation of the DEUN in Figure 3(c)



7. An Application in Food Security

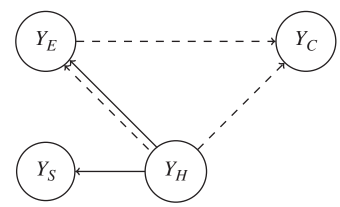
To illustrate the construction process of a DEUN, we next discuss a decision analysis in a food security application. The resulting DEUN is then used to illustrate the workings of our algorithms in practice. Food insecurity, defined as the "limited or uncertain availability of nutritionally adequate and safe foods or limited or uncertain ability to acquire acceptable foods in socially acceptable ways" (Bickel et al. 2000, p. 6), is not only an endemic issue in third world countries, but also a growing threat to wealthy nations. To support UK local governments tackling the complexity of the evaluation of various policies to ensure household food security, we have started building a probabilistic decision support tool modelling the food system.

7.1. Network Structure

After a series of decision conferences with local authorities, stakeholders, and potential decision makers, Barons et al. (2016) identified three areas that are impacted by increasing household food insecurity: educational attainment (Y_E), health (Y_H), and social cohesion (Y_S). Of course the cost (Y_C) associated with the enactment of any policy is deemed relevant in this domain. Measurable indices were then developed for each of these areas; for instance, educational attainment was assessed by the percentage of pupils not failing a combination of UK school examinations. Suppose these indices take values in $[0, 100]$. Details about the form of the various attributes are beyond the scope of this paper, and we refer the reader to Barons et al. (2016) for a discussion of these.

Of course, such a decision support system needs to model the probabilistic dependence over a much larger vector of variables to be accounted for in a reliable description of the food system. But for the illustrative purposes of this example, we assume the dependence structure between the four indexes above is summarized by the DEUN in Figure 5. This states that the variable cost is independent of all others and that, given a specific value of the health index, educational attainment and social cohesion are independent. For the preferential part, although a plausible assumption might be that the utilities of both health and social cohesion do not change when all the other attributes are varied, the utility of various levels of educational attainment did appear to sometimes be a function of health. Similarly, the utility of the costs associated with policies' implementations appeared to be a function of both educational attainment and health. These assumptions are represented in the DEUN in Figure 5

Figure 5. DEUN Representing the Food Security Example of Section 7



by the dashed arcs, depicting an underlying directional utility diagram.

For this illustrative example, we consider a decision space $\mathbb{D}$ including three policies: an increase (d_0), a decrease (d_1), and no change (d_2) of the number of pupils eligible for free school meals nationally. The UK government has already implemented this type of policy to give pupils a healthy start in life, since evidence seems to point toward an improvement in the development and social skills of young children who eat a healthy meal together at lunchtime (Kitchen et al. 2013). In this setting, we define the variables Y_E, Y_H, and Y_S as the variation in two years time of the corresponding current index value, while Y_C measures the change in the percentage of the government budget for the free school meal program. We assume that each policy directly influences Y_H, Y_E, and Y_C, while Y_S is only affected indirectly by a decision taken.

Initial discussions during the elicitation process suggested that a simple normal regression model could be sufficient to depict the probabilistic part of the system. This is defined by the distributions

$$Y_H \sim \mathcal{N}(\theta^d_{0H}, \sigma^d_H), \qquad Y_E \mid Y_H \sim \mathcal{N}(\theta^d_{0E} + \theta^d_{HE} Y_H, \sigma^d_E),$$
$$Y_C \sim \mathcal{N}(\theta^d_{0C}, \sigma^d_C), \qquad Y_S \mid Y_H \sim \mathcal{N}(\theta_{0S} + \theta_{HS} Y_H, \sigma_S),$$

where the parameters θ and σ take values in $\mathbb{R}$ and $\mathbb{R}_+$, respectively, and a superscript d denotes a different parameter value for each available policy. Notice that the above definitions are compatible with the underlying BN of Figure 5.

We assume the utilities to be exponentials and of the form specified in Table 2, where the parameters δ take values $\mathbb{R}_+$. These then need to be normalized. For an attribute Y, this can be done using the formula $u(y) = (\dot{u}(y) - m)/(M - m)$, where $\dot{u}$ is the unnormalized util-

Table 2. Unnormalized Utility Functions of the Free School Meals Example

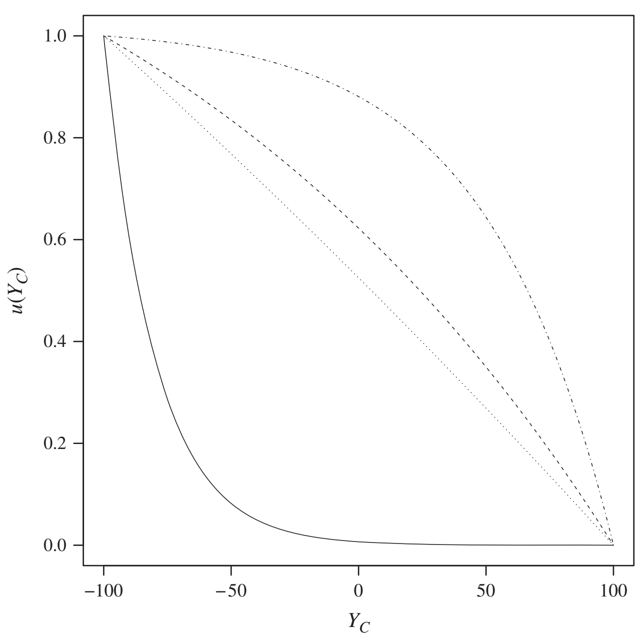
$\dot{u}(y_C \mid y_E^0, y_H^0) = \exp(-\delta_C^{00} y_C)$	$\dot{u}(y_E \mid y_H^0) = \exp(\delta_E^0 y_E)$
$\dot{u}(y_C \mid y_E^0, y_H^*) = 1 - \exp(\delta_C^{0*} y_C)$	$\dot{u}(y_E \mid y_H^*) = \exp(\delta_E^* y_E)$
$\dot{u}(y_C \mid y_E^*, y_H^0) = 1 - \exp(\delta_C^{*0} y_C)$	$\dot{u}(y_H) = \exp(\delta_H y_H)$
$\dot{u}(y_C \mid y_E^*, y_H^*) = 1 - \exp(\delta_C^{**} y_C)$	$\dot{u}(y_S) = \exp(\delta_S y_S)$

ity function, $m = \min(\dot{u}(y))$ and $M = \max(\dot{u}(y))$. So, for example, Figure 6 shows the normalized version of the utility functions of costs conditional on the boundary values of educational attainment and health, for a specific choice of the parameters δ. Again these utility definitions are compatible with the DEUN structure of Figure 5.

7.2. The Algorithm

Given the definitions of the DEUN structure and of the specific form of the probability and utility functions, we can now proceed with an illustration of our evaluation algorithm. Notice that the DEUN in Figure 5 is nondecomposable. For this reason, we first illustrate the evaluation algorithm in Theorem 1 that works over any DEUN. In Section 7.3, we consider the evaluation

Figure 6. Utility Functions of $Y_C \mid \{Y_E, Y_H\}^{0*}$



Notes. Full line, y_E^0, y_H^0; dotted line, y_E^*, y_H^0; dashed line, y_E^0, y_H^*; dotted dashed line, y_E^*, y_H^*.

algorithm based on Theorem 2 for a decomposable version of our DEUN in Figure 5. There are many variable orderings that the algorithm could follow, but we here choose the sequence (Y_E, Y_S, Y_H, Y_C).

First notice that the vector $\mathbf{u}_E$ consists of the four entries $u(y_E \mid y_H^0)$, $\hat{u}(y_E \mid y_H^0)$, $u(y_E \mid y_H^*)$, and $\hat{u}(y_E \mid y_H^*)$. The first step of the algorithm, as formalized in Equation (7), computes the expectation of these utilities with respect to the conditional probability function of Y_E given Y_H. This consists of the computation of the moment generating function of a normal random variable. Recall that for a normal random variable Y with mean μ and variance σ and a $t \in \mathbb{R}$, we have that $\mathbb{E}(\exp(tY)) = \exp(t\mu + 0.5t^2\sigma^2)$. Thus,

$$\bar{\mathbf{u}}_E = \left(\frac{E_d^0 - m_E^0}{M_E^0 - m_E^0}, \frac{M_E^0 - E_d^0}{M_E^0 - m_E^0}, \frac{E_d^* - m_E^*}{M_E^* - m_E^*}, \frac{M_E^* - E_d^*}{M_E^* - m_E^*}\right)$$

where M_E and m_E, with the appropriate superscript, denote the maximum and the minimum of the utility function, respectively, and

$$E_d^0 = \exp\left(\delta_E^0 \theta_{0E}^d + \delta_E^0 \theta_{HE}^d Y_H + 0.5(\delta_E^0 \sigma_E^d)^2\right),$$
$$E_d^* = \exp\left(\delta_E^* \theta_{0E}^d + \delta_E^* \theta_{HE}^d Y_H + 0.5(\delta_E^* \sigma_E^d)^2\right).$$

Next the algorithm considers the node Y_S. As specified by Equation (8), it first computes $\bar{\mathbf{u}}_E \circ \mathbf{u}_S$, where $\mathbf{u}_S = (u(y_S), 1 - u(y_S))$. This $\circ$ product is given by $(\bar{\mathbf{u}}_E u(y_S), \bar{\mathbf{u}}_E(1 - u(y_S)))$ since $\bar{\mathbf{u}}_E$ is not a function of Y_S. Then Equation (8) computes $\bar{\mathbf{u}}_S$ as the expectation of each entry of $\bar{\mathbf{u}}_E \circ \mathbf{u}_S$ with respect to $p(y_S \mid y_H)$. This gives the vector

$$\bar{\mathbf{u}}_S = \left(\bar{\mathbf{u}}_E(S - m_S)/(M_S - m_S), \bar{\mathbf{u}}_E(M_S - S)/(M_S - m_S)\right),$$

where $S = \exp(\delta_S(\theta_{0S} + \theta_{HS} y_H) + 0.5\delta_S^2\sigma_S^2)$.

At this point, the algorithm moves to Y_H and computes $\bar{\mathbf{u}}_S \circ \mathbf{u}_H$. Notice that $\bar{\mathbf{u}}_S$ is already a function of Y_H. Specifically the first, second, fifth, and sixth entries of $\bar{\mathbf{u}}_S$ refer to y_H^0 and therefore need to be multiplied by $1 - u(y_H)$, while the others need to be multiplied by $u(y_H)$. Then Equation (8) computes the expectation of this product with respect to $p(y_H)$, giving an eight-dimensional vector $\bar{\mathbf{u}}_H$ whose entries $\bar{\mathbf{u}}_H(i)$, $i \in [8]$, are given in Appendix C.1 with indeterminates defined in Appendix C.2.

The algorithm then moves to node Y_C. Since $\bar{\mathbf{u}}_H$ is not a function of Y_C, $\bar{\mathbf{u}}_H \circ \mathbf{u}(y_C \mid y_E^{0*}, y_H^{0*})$ returns the elements

$$\begin{aligned}
&\bar{\mathbf{u}}_H(1)u(y_C \mid y_E^*, y_H^0), && \bar{\mathbf{u}}_H(2)u(y_C \mid y_E^0, y_H^0),\\
&\bar{\mathbf{u}}_H(3)u(y_C \mid y_E^*, y_H^*), && \bar{\mathbf{u}}_H(4)u(y_C \mid y_E^0, y_H^*),\\
&\bar{\mathbf{u}}_H(5)u(y_C \mid y_E^*, y_H^0), && \bar{\mathbf{u}}_H(6)u(y_C \mid y_E^0, y_H^0),\\
&\bar{\mathbf{u}}_H(7)u(y_C \mid y_E^*, y_H^*), && \bar{\mathbf{u}}_H(8)u(y_C \mid y_E^0, y_H^*),\\
&\bar{\mathbf{u}}_H(1)\hat{u}(y_C \mid y_E^*, y_H^0), && \bar{\mathbf{u}}_H(2)\hat{u}(y_C \mid y_E^0, y_H^0),\\
&\bar{\mathbf{u}}_H(3)\hat{u}(y_C \mid y_E^*, y_H^*), && \bar{\mathbf{u}}_H(4)\hat{u}(y_C \mid y_E^0, y_H^*),\\
&\bar{\mathbf{u}}_H(5)\hat{u}(y_C \mid y_E^*, y_H^0), && \bar{\mathbf{u}}_H(6)\hat{u}(y_C \mid y_E^0, y_H^0),\\
&\bar{\mathbf{u}}_H(7)\hat{u}(y_C \mid y_E^*, y_H^*), && \bar{\mathbf{u}}_H(8)\hat{u}(y_C \mid y_E^0, y_H^*).
\end{aligned}$$

The expectation of the above terms with respect to $p(y_C)$ then follows by simply applying the moment generating function relationships for normal random variables, since $\bar{\mathbf{u}}_H$ is not a function of Y_C. We denote the resulting vector as $\bar{\mathbf{u}}_C = (\bar{\mathbf{u}}_C(i))_{i\in[16]}$.

As formalized in Equation (9), the algorithm then terminates by taking the sum of the element of $\bar{\mathbf{u}}_C$ multiplied by the appropriate weighting term $u(\mathbf{y}^{0*})$. Specifically, the overall expected utility for a decision $d \in \mathbb{D}$ equals the sum of the terms

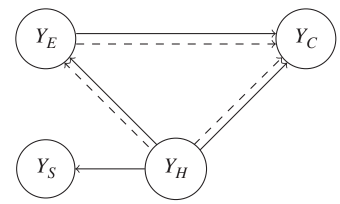
$$\begin{aligned}
&\bar{\mathbf{u}}_C(1)u(y_E^*, y_S^*, y_H^0, y_C^*), && \bar{\mathbf{u}}_C(2)u(y_E^0, y_S^*, y_H^0, y_C^*),\\
&\bar{\mathbf{u}}_C(3)u(y_E^*, y_S^*, y_H^*, y_C^*), && \bar{\mathbf{u}}_C(4)u(y_E^0, y_S^*, y_H^*, y_C^*),\\
&\bar{\mathbf{u}}_C(5)u(y_E^*, y_S^0, y_H^0, y_C^*), && \bar{\mathbf{u}}_C(6)u(y_E^0, y_S^0, y_H^0, y_C^*),\\
&\bar{\mathbf{u}}_C(7)u(y_E^*, y_S^0, y_H^*, y_C^*), && \bar{\mathbf{u}}_C(8)u(y_E^0, y_S^0, y_H^*, y_C^*),\\
&\bar{\mathbf{u}}_C(9)u(y_E^*, y_S^*, y_H^0, y_C^0), && \bar{\mathbf{u}}_C(10)u(y_E^0, y_S^*, y_H^0, y_C^0),\\
&\bar{\mathbf{u}}_C(11)u(y_E^*, y_S^*, y_H^*, y_C^0), && \bar{\mathbf{u}}_C(12)u(y_E^0, y_S^*, y_H^*, y_C^0),\\
&\bar{\mathbf{u}}_C(13)u(y_E^*, y_S^0, y_H^0, y_C^0), && \bar{\mathbf{u}}_C(14)u(y_E^0, y_S^0, y_H^0, y_C^0),\\
&\bar{\mathbf{u}}_C(15)u(y_E^*, y_S^0, y_H^*, y_C^0), && \bar{\mathbf{u}}_C(16)u(y_E^0, y_S^0, y_H^*, y_C^0).
\end{aligned} \tag{13}$$

The exact form of the overall expected utility can be shown to be a highly nonlinear function of the problem's parameters. But it has a closed-form expression, and this form is the same for all available decisions. Thus, the identification of an optimal strategy can then be carried out by simply plugging in the different numerical specifications associated to different policies. In Appendix D, we give plausible values to the parameters of the free school meal example. For such values, the decision d_0 of increasing the number of eligible pupils would be optimal, having expected utility score 0.29, compared to 0.19 and 0.21 for policies d_1 and d_2 respectively.

7.3. The Algorithm for Decomposable Networks

The same ranking of the available policies $d \in \mathbb{D}$ would have been achieved by evaluating the DEUN in Figure 5, after an appropriate transformation, with our algorithm for decomposable DEUNs. A decomposable version of the original DEUN can be deduced by simply adding the probabilistic edges (Y_H, Y_C) and (Y_E, Y_C). This is shown in Figure 7. This new DEUN consists of two cliques, $C_1 = \{Y_C, Y_E, Y_H\}$ and $C_2 = \{Y_S, Y_H\}$, with separator $S_2 = \{Y_H\}$. The associated junction tree is then one simply connecting the two cliques by the edge (C_1, C_2).

Figure 7. Decomposable Version of the DEUN in Figure 5



To apply Theorem 2, we first need to define the probability and utility potentials that for this application equal

$$\Phi_{C_1} = p(y_C \mid y_E, y_H)p(y_E \mid y_H)p(y_H), \quad \Phi_{C_2} = p(y_S \mid y_H),$$
$$\Psi_{C_1} = \mathbf{u}(y_C \mid y_E^{0*}, y_H^{0*}) \circ \mathbf{u}(y_E \mid y_H^{0*}) \circ \mathbf{u}(y_H) \circ \mathbf{u}(\mathbf{y}^{0*}),$$
$$\Psi_{C_2} = \mathbf{u}(y_S).$$

Then the evaluation algorithm starts with the absorption of the leaf C_2 into the root C_1 by computing the integral in Equation (10):

$$\int_{\mathbb{Y}_{C_2\setminus S_2}} \Phi_{C_2}\Psi_{C_2}\,\mathrm{d}\mathbf{y}_{C_2\setminus S_2} = \int_{\mathbb{Y}_S} \mathbf{u}(y_S)p(y_S \mid y_H)\,\mathrm{d}y_S = \left(\frac{S - m_S}{M_S - m_S}, \frac{M_S - S}{M_S - m_S}\right), \tag{14}$$

where $S = \exp(\delta_S(\theta_{0S} + \theta_{HS}y_H) + 0.5\delta_S^2\sigma_S^2)$ as in Section 7.2. At this stage, to compute $\Psi_{C_1}^{C_2}$ as in Equation (10), we need to $\circ$ multiply the right-hand side of Equation (14) with Ψ_{C_1}. This operation returns the vector with entries listed in Appendix E.1.

Since the junction tree of this example has one root only, the evaluation algorithm ends by computing Equation (11), which first entails the marginalization of $\Phi_{C_1}\Psi_{C_1}^{C_2}$ with respect to $\mathbf{y}_{C_1}$ and then the sum of the entries of the resulting vector. To perform the marginalization step, notice (from Sullivant et al. 2010) that the random vector (Y_H, Y_E, Y_C) follows a normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix Σ, where

$$\boldsymbol{\mu} = (\beta^d_{0H}, \beta^d_{0E} + \beta^d_{HE}\beta^d_{0H}, \beta^d_{0C} + \beta^d_{0H}\beta^d_{HC} + \beta^d_{EC}(\beta^d_{0E} + \beta^d_{HE}\beta^d_{0H})),$$

$$\Sigma = \begin{pmatrix} \sigma^d_H & \beta^d_{HE}\sigma^d_H & \beta^d_{HC}\sigma^d_H + \beta^d_{EC}\beta^d_{HE}\sigma^d_H \\ \beta^d_{HE}\sigma^d_H & \sigma^d_E & \beta^d_{HC}\sigma^d_H + \beta^d_{EC}\beta^d_{HE}\sigma^d_H \\ \Sigma_{31} & \Sigma_{32} & \sigma^d_C \end{pmatrix}, \quad \text{where}$$

$$\Sigma_{31} = \beta^d_{HC}\sigma^d_H + \beta^d_{EC}\beta^d_{HE}\sigma^d_H, \quad \Sigma_{32} = \beta^d_{EC}\sigma^d_E + \beta^d_{HC}\beta^d_{HE}\sigma^d_H$$

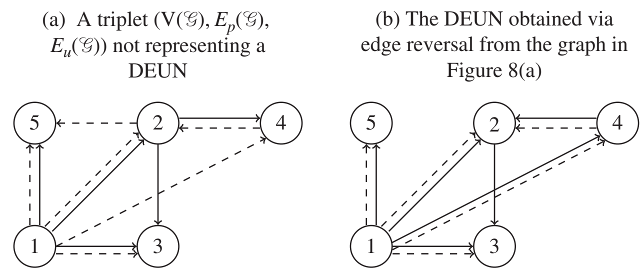
and the notation is straightforwardly adapted from Section 7.2 to describe the DEUN in Figure 7. It can be noticed that the entries of $\Psi^{C_2}_{C_1}$ in Section E.1 are a linear combination of terms $a\exp(\mathbf{t}^{\mathrm{T}}\mathbf{y}_{C_1})$, for $a \in \mathbb{R}$ and $\mathbf{t} \in \mathbb{R}^3$, and thus, by the properties of normal vectors, the marginalization maps each term $a\exp(\mathbf{t}^{\mathrm{T}}\mathbf{y}_{C_1})$ into $a\exp(\mathbf{t}^{\mathrm{T}}\mathbf{u} + 0.5\mathbf{t}^{\mathrm{T}}\Sigma\mathbf{t})$. Using this result, it can be shown that the entries of $\int_{\mathbb{Y}_{C_1}} \Phi_{C_1}\Psi^{C_2}_{C_1}\,d\mathbf{y}_{C_1}$ are equal to those reported in Equation (13), and the overall expected utility $\bar{u}$ is then the sum of these terms.

The evaluation order of the variables in Section 7.2 was chosen to have the simplest notation, but for other variables' orders, the algorithm in Theorem 1 would have required the transmission and computation of a larger number of terms not relevant during the evaluation. Conversely, the algorithm for decomposable DEUNs enables fast propagation routines to be applied using the properties of random vectors and matrix calculus, as briefly illustrated by this example.

8. Extensions

Although the algorithms presented in Section 6 allow for the computation of expected utilities in settings much more general than standard ID modelling, the requirement of a directed utility diagram of the same directionality of the underlying BN model might not be entertained in some applications. For this reason, here we first present a method to transform a triplet $(V(\mathcal{G}), E_p(\mathcal{G}), E_u(\mathcal{G}))$, where $E_u(\mathcal{G})$ is a directional utility diagram, into a DEUN. Second, we discuss an extension of our DEUN framework that can accommodate more flexible utility structures but that can still use a distributed routine in a fashion similar to that in Theorem 1.

Figure 8. Illustration of the Edge Reversal Process



8.1. Edge Reversal in DEUNs

Changing the directionality of some of the probabilistic edges of an ID model is an operation commonly done to remove some vertices that are not strictly relevant for the ranking of the available policies. However, when one edge is reserved, some additional edges might have to be included to prevent the new graph to represent conditional independences that were not implied by the original DAG. Such edge reversal operations can be equally applied to the probabilistic side of DEUNs.

First we formalize the edge reversal operation. Let $\mathcal{G} = (V(\mathcal{G}), E_p(\mathcal{G}))$ be a DAG. For $i, j \in V(\mathcal{G})$, let $i \in \Pi^p_j$ and (i, j) be the only directed path between i and j. Then $\mathcal{G}$ and $\mathcal{G}'$ imply the same probability distribution where $V(\mathcal{G}) = V(\mathcal{G}')$ and

$$E(\mathcal{G}') = E(\mathcal{G})\setminus\{(i,j)\} \cup \{(k,i) \colon k \in \{\Pi^p_j \cup j\}\setminus i\} \cup \{(k,j) \colon k \in \Pi^p_i\}.$$

In practice, an edge reversal of (i, j) entails the addition of an edge from any parent of i to j and an edge from any parent of j to i.

We next describe how this operation can be used to transform a triplet $(V(\mathcal{G}), E_p(\mathcal{G}), E_u(\mathcal{G}))$ into a DEUN.

Example 11. Consider the graph in Figure 8(a). It can be noticed that this is not a DEUN since $(2,4) \in E_p(\mathcal{G})$ and $(4,2) \in E_u(\mathcal{G})$. However, if we reverse the edge $(2,4) \in E_p(\mathcal{G})$ and consequently add the edge $(1,4)$ to $E_p(\mathcal{G})$, since $1 \in \Pi^p_2$, then the result graph is a DEUN. This is shown in Figure 8(b).

The following result generalizes the transformation illustrated in the above example.

Proposition 2. *Any triplet $(V(\mathcal{G}), E_p(\mathcal{G}), E_u(\mathcal{G}))$ can be transformed into a DEUN if $E_u(\mathcal{G})$ is a directional utility diagram.*

This holds by noting that in the worst-case scenario a sequential use of edge reversals can return a DAG $(V(\mathcal{G}), E_p(\mathcal{G}))$ that is complete but with edges in the same direction as in the utility diagram. Thus our algorithms can be applied after a series of edge reversal operations and, if required, a "triangulation step" to make the resulting DEUN decomposable, whenever $E_u(\mathcal{G})$ is directional.

8.2. A Class of Nondirectional Utility Diagrams

Of course, in some applications the requirement of a directional utility diagram can be too restrictive. It is thus important to develop a methodology for the quick and distributed computation of expected utilities embedding more flexible utility structures. Such a full theoretical development is beyond the scope of this paper, although we refer to the discussion for some details on how we believe this problem could be approached. However, we note here that there happens to be a class of bidirectional utility diagrams where our evaluation algorithms can be used after some small adaptations. Such diagrams are called *canonical* in Abbas (2010). For a set K, let $\#K$ denote the number of elements of K.

Definition 8. Let $K \subseteq [n]$ be the set comprising vertices i for which $\Pi_i^u \neq [n]\backslash i$. A utility diagram is *canonical* if

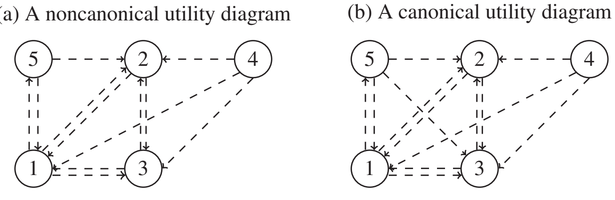
- $\#K < 2$;
- if $\#K \geq 2$, there are no edges connecting any two vertices in K.

The set K comprises the vertices that are utility independent of at least one variable conditionally on all others.

Example 12. Figure 9 reports two utility diagrams. In Figure 9(a), $K = \{3, 4, 5\}$, while in Figure 9(b), $K = \{4, 5\}$. The first diagram is not canonical since there is an edge connecting 4 to 3. Conversely, the second diagram is canonical since there is no edge between 4 and 5.

Figure 9. Illustration of the Difference Between Canonical and Noncanonical Utility Diagrams

(a) A noncanonical utility diagram (b) A canonical utility diagram



These diagrams entertain very convenient properties for utility elicitation (Abbas 2010, for a discussion, see). Importantly, for such diagrams, the utility factorization is equal for any expansion order comprising the vertices in K. Notice that this is not true for directional utility diagrams since they are not canonical in general. From Abbas (2010), the utility factorization of any canonical utility diagram can be written as

$$u(\mathbf{y}) = \sum_{\mathbf{y}_K^{0*} \in \mathbb{Y}_K^{0*}} u(\mathbf{y}_K^{0*}, \mathbf{y}_{-K}) \prod_{i \in K} g_i(y_i \mid \mathbf{y}_{\Pi_i^u}, \mathbf{y}_{iP}^{0*}),$$

where we recall that iP denotes the set of vertices that precede i in the order. Notice that in contrast to directional utility diagrams, canonical utility diagrams have utility factorizations including functions of more than one attribute. For a discussion of the elicitation process for such terms we refer to Abbas (2010).

Example 13. The utility factorization associated to the canonical utility diagram in Figure 9(b) is

$$\begin{aligned} u(\mathbf{y}) = {} & u(\mathbf{y}_{[3]}, y_4^0, y_5^0)(1 - u(y_5 \mid y_1))(1 - u(y_4)) \\ & + u(\mathbf{y}_{[3]}, y_4^*, y_5^0)(1 - u(y_5 \mid y_1))u(y_4) \\ & + u(\mathbf{y}_{[3]}, y_4^0, y_5^*)u(y_5 \mid y_1)(1 - u(y_4)) \\ & + u(\mathbf{y}_{[3]}, y_4^*, y_5^*)u(y_5 \mid y_1)u(y_4). \end{aligned}$$

We are now ready to define a model comprising both probabilistic and utility edges embedding utility diagrams that are not necessarily directional.

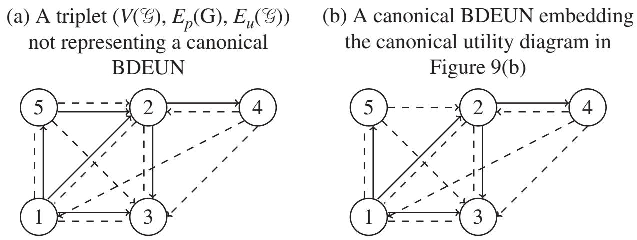
Definition 9. A *canonical bidirected expected utility network* $\mathcal{G}$ consists of a set of vertices $V(\mathcal{G}) = [n]$, a probabilistic edge set $E_p(\mathcal{G})$, denoted by solid arrows, and a utility edge set $E_u(\mathcal{G})$, denoted by dashed arrows, such that

- $(V(\mathcal{G}), E_u(\mathcal{G}))$ is a canonical utility diagram;
- $(V(\mathcal{G}), E_p(\mathcal{G}))$ is a BN model such that $i \notin \Pi_j^p$ for any $i \in K$ and any $j \in [n]\backslash K$.

A BDEUN is such that there is no probabilistic edge from any vertex in K to another one not in this set.

Example 14. Figure 10 reports the graphical representation of two triplets $(V(\mathcal{G}), E_p(\mathcal{G}), E_u(\mathcal{G}))$, where in both cases $(V(\mathcal{G}), E_u(\mathcal{G}))$ is the canonical utility diagram in Figure 9(b). (For simplicity, we replaced two

Figure 10. Graphical Representations of Probabilitistic and (Canonical) Utility Structures



dashed directed edges between the same two vertices with an undirected one.) The diagram in Figure 10(a) is not a canonical BDEUN since $(5,2) \in E_p(\mathcal{G})$ with $5 \in K$ and $2 \in [n]\backslash K$. Conversely, the diagram in Figure 10(b) is a canonical BDEUN.

Notice that, just as for DEUNs, any triplet ($V(\mathcal{G})$, $E_p(\mathcal{G})$, $E_u(\mathcal{G})$), where ($V(\mathcal{G})$, $E_u(\mathcal{G})$) is a canonical utility diagram and ($V(\mathcal{G})$, $E_p(\mathcal{G})$) is a BN, can be transformed into a canonical BDEUN via edge reversals.

Example 15. The diagram in Figure 10(a) can be converted into a canonical BDEUN by reversing the edge $(5,2) \in E_p(\mathcal{G})$. For this example, this operation does not require the addition of any new probabilistic edge.

We are now ready to introduce an evaluation algorithm for canonical BDEUNs.

Proposition 3. *For a canonical BDEUN $\mathcal{G}$, let $j = \max\{[n]\backslash K\}$. The expected utility score $\bar{u}$ associated to $\mathcal{G}$ can be computed according to the following algorithm*:

1. *follow the algorithm in Theorem 1 until vertex $j+1$ included*;
2. *compute*

$$\bar{\mathbf{u}}_j = \int_{\mathbb{Y}_j} (\bar{\mathbf{u}}_{j+1} \circ \mathbf{u}(\mathbf{y}_{[n]\backslash K}, \mathbf{y}_K^{0*}))p(y_j \mid \mathbf{y}_{\Pi_j^p})\,\mathrm{d}y_j;$$

3. *for i from $j-1$ to 1, compute*

$$\bar{\mathbf{u}}_i = \int_{\mathbb{Y}_i} \bar{\mathbf{u}}_{i+1} p(y_i \mid \mathbf{y}_{\Pi_i^p})\,\mathrm{d}y_i;$$

4. *return* $|\bar{\mathbf{u}}_1|$.

The proof of this result follows steps almost identical to those of Theorem 1 and is thus not reported here. The intuition behind the above algorithm is that it exactly follows the steps of Theorem 1 until it reaches a vertex outside the set K, say, j. At that point, it computes the $\circ$ product between $\bar{\mathbf{u}}_{j+1}$ and those terms $\mathbf{u}_j(\mathbf{y}_{[n]\backslash K}, \mathbf{y}_K^{0*})$ that would be associated to the criterion weights in a DEUN. The contribution to the overall utility function of all the terms not in K is then already accounted for by $\mathbf{u}(\mathbf{y}_{[n]\backslash K}, \mathbf{y}_K^{0*})$. Consequently, the algorithm can be completed by simple marginalizations over the variables Y_i for $i \in [n]\backslash K$.

Example 16. For the canonical BDEUN in Figure 10(b), the algorithm in Proposition 3 follows the same steps as in Theorem 1 for the computation of $\bar{\mathbf{u}}_5$ and $\bar{\mathbf{u}}_4$, since for this BDEUN $K = \{4,5\}$. At this stage, the algorithm computes first $\mathbf{u}_3(\mathbf{y}_{[3]}, \mathbf{y}_K^{0*}) \circ \bar{\mathbf{u}}_4$ and then a marginalization over Y_3, since 3 is the first vertex not in K. The algorithm then concludes by marginalizing over Y_2 and Y_1 and finally computing the sum of the entries of $\bar{\mathbf{u}}_1$.

9. Discussion

Graphical representations of both probabilistic and preferential independences have received great attention in the literature. However, so far very little effort has been applied to the study of how probabilistic and preferential graphical models could be combined to provide a graphical representation of the expected utility structure of a decision problem. In this paper we presented one of the first attempts to formally define network models depicting both the probabilistic and the utility relationships for a random vector of attributes. We have demonstrated here how such graphical representations then provide a framework for the fast computation of the overall expected utility through various distributed routines.

While the constraint of having only directed probabilistic edges is very often met in practice, and indeed BNs are the most common probabilistic graphical model, restricting the class of underlying utility diagrams to only directional ones may be unreasonable in some applications. We demonstrated here that it is still possible to compute the expected utility of a decision in a distributed fashion for a specific subclass of bidirected utility diagrams, those usually called canonical. However, for noncanonical utility diagrams, evaluation algorithms are still to be developed. Such utility structures could be included by representing the probabilistic structure via a probabilistic chain graph (Lauritzen 1996). Propagation algorithms also exist for this model class, and therefore adaptations of these could enable the computation of expected utilities in this more general class of models.

Finally, both DEUNs and canonical BDEUNs could also be generalized to include decision nodes and therefore fully represent the structure of a DM's decision problem, just as influence diagrams extend BN models. We envisage that the evaluation of such a network could be performed by algorithms that share many features with the ones presented here, but also equipped with optimization steps over decision spaces.

Appendix A. Proof of Theorem 1

Define for $i \in [n]$

$$\tilde{\mathbf{u}}_i = \bigcirc_{j\in[i]} \mathbf{u}_j(y_j \mid \mathbf{y}_{\Pi^u_j}),$$

and note that

$$\bar{u} = \int_{\mathbb{Y}} u(\mathbf{y})p(\mathbf{y})\,\mathrm{d}\mathbf{y} = \left|\mathbf{u}_0(\mathbf{y}^{0*}) \circ \int_{\mathbb{Y}} \tilde{\mathbf{u}}_n p(\mathbf{y})\,\mathrm{d}\mathbf{y}\right|. \tag{A.1}$$

Now consider the second integral in Equation (A.1). We have that

$$\begin{aligned}\int_{\mathbb{Y}} \tilde{\mathbf{u}}_n p(\mathbf{y})\,\mathrm{d}\mathbf{y} &= \int_{\mathbb{Y}} (\tilde{\mathbf{u}}_{n-1} \circ \mathbf{u}(y_n \mid \mathbf{y}^{0*}_{\Pi^u_n}))p(y_n \mid \mathbf{y}_{\Pi^p_n})p(\mathbf{y}_{[n-1]})\,\mathrm{d}\mathbf{y}\\ &= \int_{\mathbb{Y}_{[n-1]}} \tilde{\mathbf{u}}_{n-1}p(\mathbf{y}_{[n-1]}) \circ \int_{\mathbb{Y}_n} \mathbf{u}(y_n \mid \mathbf{y}^{0*}_{\Pi^u_n})p(y_n \mid \mathbf{y}_{\Pi^p_n})\,\mathrm{d}y_n\,\mathrm{d}\mathbf{y}_{[n-1]}\\ &= \int_{\mathbb{Y}_{[n-1]}} \tilde{\mathbf{u}}_{n-1}p(\mathbf{y}_{[n-1]}) \circ \bar{\mathbf{u}}_n\,\mathrm{d}\mathbf{y}_{[n-1]}\\ &= \int_{\mathbb{Y}_{[n-1]}} (\tilde{\mathbf{u}}_{n-1} \circ \bar{\mathbf{u}}_n)p(\mathbf{y}_{[n-1]})\,\mathrm{d}\mathbf{y}_{[n-1]},\end{aligned} \tag{A.2}$$

where $\bar{\mathbf{u}}_n$ is defined in Equation (7). By marginalizing out y_{n-1}, we can then deduce from Equation (A.2) that

$$\begin{aligned}\int_{\mathbb{Y}} \tilde{\mathbf{u}}_n p(\mathbf{y})\,\mathrm{d}\mathbf{y} &= \int_{\mathbb{Y}_{[n-2]}}\int_{\mathbb{Y}_{n-1}} (\tilde{\mathbf{u}}_{n-2} \circ \mathbf{u}(y_{n-1} \mid \mathbf{y}^{0*}_{\Pi^u_{n-1}}) \circ \bar{\mathbf{u}}_n)\\ &\quad \cdot p(y_{n-1} \mid \mathbf{y}_{\Pi^p_{n-1}})p(\mathbf{y}_{[n-2]})\,\mathrm{d}y_{n-1}\,\mathrm{d}\mathbf{y}_{[n-2]} \qquad \text{(A.3)}\\ &= \int_{\mathbb{Y}_{[n-2]}} \tilde{\mathbf{u}}_{n-2}p(\mathbf{y}_{[n-2]}) \circ \int_{\mathbb{Y}_{n-1}} (\mathbf{u}(y_{n-1} \mid \mathbf{y}^{0*}_{\Pi^u_{n-1}}) \circ \bar{\mathbf{u}}_n)\\ &\quad \cdot p(y_{n-1} \mid \mathbf{y}_{\Pi^p_{n-1}})\,\mathrm{d}y_{n-1}\,\mathrm{d}\mathbf{y}_{[n-2]}. \qquad \text{(A.4)}\end{aligned}$$

From Equation (8) of Theorem 1, it then follows that

$$\begin{aligned}\int_{\mathbb{Y}} \tilde{\mathbf{u}}_n p(\mathbf{y})\,\mathrm{d}\mathbf{y} &= \int_{\mathbb{Y}_{[n-2]}} \tilde{\mathbf{u}}_{n-2}p(\mathbf{y}_{[n-2]}) \circ \bar{\mathbf{u}}_{n-1}\,\mathrm{d}\mathbf{y}_{[n-2]}\\ &= \int_{\mathbb{Y}_{[n-2]}} (\tilde{\mathbf{u}}_{n-2} \circ \bar{\mathbf{u}}_{n-1})p(\mathbf{y}_{[n-2]})\,\mathrm{d}\mathbf{y}_{[n-2]}.\end{aligned} \tag{A.5}$$

By sequentially repeating the steps in Equations (A.3)–(A.5), we can now deduce that after the marginalization of Y_2,

$$\begin{aligned}\int_{\mathbb{Y}} \tilde{\mathbf{u}}_n p(\mathbf{y})\,\mathrm{d}\mathbf{y} &= \int_{\mathbb{Y}_{[1]}} (\tilde{\mathbf{u}}_1 \circ \bar{\mathbf{u}}_2)p(\mathbf{y}_{[1]})\,\mathrm{d}\mathbf{y}_{[1]}\\ &= \int_{\mathbb{Y}_1} (\mathbf{u}(y_1) \circ \bar{\mathbf{u}}_2)p(y_1)\,\mathrm{d}y_1 = \bar{\mathbf{u}}_1.\end{aligned} \tag{A.6}$$

Therefore by plugging in Equation (A.6) into (A.1), we can conclude that Equation (9) holds.

Appendix B. Proof of Theorem 2

Call

$$\bar{\Phi}_L = \prod_{C\in\mathscr{C}\setminus L} \Phi_C, \quad \text{and} \quad \bar{\Psi}_L = \bigcirc_{C\in\mathscr{C}\setminus L}\Psi_C.$$

We have that

$$\begin{aligned}&\int_{\mathbb{Y}_{L\setminus S}} \Phi_{\mathcal{T}}\Psi_{\mathcal{T}}\,\mathrm{d}\mathbf{y}_{L\setminus S}\\ &= \int_{\mathbb{Y}_{L\setminus S}} \bar{\Phi}_L\Phi_L(|\Psi_L \circ \bar{\Psi}_L|)\,\mathrm{d}\mathbf{y}_{L\setminus S} = \left|\int_{\mathbb{Y}_{L\setminus S}} \bar{\Phi}_L\Phi_L(\Psi_L \circ \bar{\Psi}_{L\setminus S})\,\mathrm{d}\mathbf{y}_{L\setminus S}\right|\\ &= \left|\bar{\Phi}_L \int_{\mathbb{Y}_{L\setminus S}} (\Phi_L\Psi_L \circ \bar{\Psi}_L)\,\mathrm{d}\mathbf{y}_{L\setminus S}\right| = \left|\bar{\Phi}_L\bar{\Psi}_L \circ \int_{\mathbb{Y}_{L\setminus S}} \Phi_L\Psi_L\,\mathrm{d}\mathbf{y}_{L\setminus S}\right|.\end{aligned} \tag{B.1}$$

Writing $\bar{\Psi}_L = \Psi_{\Pi_L} \circ \hat{\Psi}_L$, where Ψ_{Π_L} is the utility potential of the parent clique of L and $\hat{\Psi}_L = \bigcirc_{C\in\mathscr{C}\setminus L\setminus\Pi_L}\Psi_C$, it then follows from Equation (B.1) that

$$\begin{aligned}\int_{\mathbb{Y}_{L\setminus S}} \Phi_{\mathcal{T}}\Psi_{\mathcal{T}}\,\mathrm{d}\mathbf{y}_{L\setminus S} &= \left|\bar{\Phi}_L\hat{\Psi}_L \circ \Psi_{\Pi_L} \circ \int_{\mathbb{Y}_{L\setminus S}} \Phi_L\Psi_L\,\mathrm{d}\mathbf{y}_{L\setminus S}\right|\\ &= \bar{\Phi}_L(|\hat{\Psi}_L \circ \Psi^L_{\Pi_L}|) = \Phi_{\mathcal{T}\setminus L}\Psi_{\mathcal{T}\setminus L}.\end{aligned}$$

Appendix C. Expected Utility Vectors in Section 7.2

C.1. Entries of $\bar{\mathbf{u}}_H$

$$\begin{aligned}\bar{\mathbf{u}}_H(1) &= \big[M_H\overline{ES}^0_d - \overline{ESH}^0_d + m_S(\overline{EH}^0_d - M_H\bar{E}^0_d)\\ &\quad + m^0_E(\overline{SH}_d - M_H\bar{S} + m_S(M_H - \overline{H}_d))\big]\\ &\quad \cdot [(M^0_E - m^0_E)(M_S - m_S)(M_H - m_H)]^{-1},\\ \bar{\mathbf{u}}_H(2) &= \big[M^0_E(M_H\bar{S} - \overline{SH}_d + m_S(\overline{H}_d - M_H)) - M_H\overline{ES}^0_d\\ &\quad + \overline{ESH}^0_d + m_S(M_H\bar{E}^0_d - \overline{EH}^0_d)\big]\\ &\quad \cdot \big[(M^0_E - m^0_E)(M_S - m_S)(M_H - m_H)\big]^{-1},\\ \bar{\mathbf{u}}_H(3) &= \big[\overline{ESH}^*_d - m_H\overline{ES}^*_d + m_S(m_H\bar{E}^*_d - \overline{EH}^*_d)\\ &\quad + m^*_E(m_S\overline{H}_d - \overline{SH}^*_d + m_H(\bar{S} - m_S))\big]\\ &\quad \cdot[(M^*_E - m^*_E)(M_S - m_S)(M_H - m_H)]^{-1},\\ \bar{\mathbf{u}}_H(4) &= [M^*_E(\overline{SH}_d - m_H\bar{S} + m_S(m_H - \overline{H}_d)) - \overline{ESH}^*_d\\ &\quad + m_H\overline{ES}^*_d + m_S(\overline{EH}^*_d - m_H\bar{E}^*_d)\big]\\ &\quad \cdot [(M^*_E - m^*_E)(M_S - m_S)(M_H - m_H)]^{-1},\end{aligned}$$

$$\begin{aligned}
\bar{\mathbf{u}}_H(5) &= \big[M_H(M_S\bar{E}^0_d - \overline{ES}^0_d) + \overline{ESH}^0_d - M_S\overline{EH}^0_d \\
&\qquad + m^0_E(M_S(\overline{H}_d - M_H) + M_H\overline{S} - \overline{SH}_d)\big] \\
&\qquad \cdot \big[(M^0_E - m^0_E)(M_S - m_S)(M_H - m_H)\big]^{-1}, \\
\bar{\mathbf{u}}_H(6) &= \big[M^0_E(M_S(M_H - \overline{H}_d) - M_H\overline{S} + \overline{SH}_d) \\
&\qquad + M_S(\overline{EH}^0_d - M_H\bar{E}^0_d) + M_H\overline{ES}^0_d - \overline{ESH}^0_d\big] \\
&\qquad \cdot \big[(M^0_E - m^0_E)(M_S - m_S)(M_H - m_H)\big]^{-1}, \\
\bar{\mathbf{u}}_H(7) &= \big[M_S(\overline{EH}^*_d - m_H\bar{E}^*_d) - \overline{ESH}^*_d + m_H\overline{ES}^*_d \\
&\qquad + m^*_E(\overline{SH} - M_S\overline{H} + m_H(M_S - \overline{S}))\big] \\
&\qquad \cdot \big[(M^*_E - m^*_E)(M_S - m_S)(M_H - m_H)\big]^{-1}, \\
\bar{\mathbf{u}}_H(8) &= \big[M^*_E(M_S(\overline{H}_d - m_H) - \overline{SH}_d + m_H\overline{S}) + \overline{ESH}^*_d \\
&\qquad - m_H\overline{ES}^*_d + M_S(m_H\bar{E}^*_d - \overline{EH}^*_d)\big] \\
&\qquad \cdot \big[M^*_E - m^*_E)(M_S - m_S)(M_H - m_H)\big]^{-1}.
\end{aligned}$$

C.2. Definition of the Indeterminates in Section C.1, Where $\bar{E}^*_d$, $\overline{EH}^*_d$, $\overline{ES}^*_d$, and $\overline{ESH}^*_d$ are Similarly Defined

$$\begin{aligned}
\overline{S} &= \exp\big(\delta_S(\theta_{0S} + \theta_{HS}\theta^d_{0H}) + 0.5\delta^2_S(\sigma^2_S + (\theta_{HS}\sigma^d_H)^2)\big) \\
\bar{E}^0_d &= \exp\big(\delta^0_E(\theta^d_{0E} + \theta^d_{HE}\theta^d_{0H}) \\
&\qquad + 0.5(\delta^0_E)^2((\sigma^d_E)^2 + (\theta^d_{HE}\sigma^d_H)^2)\big) \\
\overline{H}_d &= \exp\big(\delta_H\theta^d_{0H} + 0.5(\delta_H\sigma^d_H)^2\big) \\
\overline{SH}_d &= \exp\big(\delta_S\theta_{0S} + 0.5\delta^2_S\sigma^2_S + (\delta_S\theta_{HS} + \delta_H)\theta^d_{0H} \\
&\qquad + 0.5(\delta_S\theta_{HS} + \delta_H)^2(\sigma^d_H)^2\big) \\
\overline{EH}^0_d &= \exp\big(\delta^0_E\theta^d_{0E} + 0.5(\delta^0_E\sigma^d_E)^2 + (\delta^0_E\theta^d_{HE} + \delta_H)\theta^d_{0H} \\
&\qquad + 0.5(\delta^0_E\theta^d_{HE} + \delta_H)^2(\sigma^d_H)^2\big) \\
\overline{ES}^0_d &= \exp\big(\delta_S\theta_{0S} + \delta^0_E\theta^d_{0E} + 0.5(\delta^2_S\sigma^2_S + (\delta^0_E\sigma^d_E)^2) \\
&\qquad + (\delta_S\theta_{HS} + \delta^0_E\theta^d_{HE})\theta^d_{0H} + 0.5(\delta_S\theta_{HS} + \delta^0_E\theta^d_{HE})^2(\sigma^d_H)^2\big) \\
\overline{ESH}^0_d &= \exp\big(\delta_S\theta_{0S} + \delta^0_E\theta^d_{0E} + 0.5(\delta^2_S\sigma^2_S + (\delta^d_E\sigma^d_E)^2) \\
&\qquad + (\delta_S\theta_{HS} + \delta^0_E\theta^d_{HE} + \delta_H)\theta^d_{0H} \\
&\qquad + 0.5(\delta_S\theta_{HS} + \delta^0_E\theta^d_{HE} + \delta_H)^2(\sigma^d_H)^2\big).
\end{aligned}$$

Appendix D. Numerical Specifications for the Food Security Example

Table D.1. Parameters Dependent on Decisions

	θ^d_{0H}	σ^d_H	θ^d_{0C}	σ^d_C	θ^d_{0E}	σ^d_E	θ^d_{HE}
d_0	1.5	5	30	8	5	40	7
d_1	−2	4	−5	5	−6	20	2
d_2	−0.5	3	10	4	3	15	7

Table D.2. Parameters Not Dependent on Decisions

$\delta^{00}_C = 0.05$,	$\delta^0_E = 0.01$,	$\theta_{0S} = 5$,
$\delta^{0*}_C = 0.005$,	$\delta^*_E = 0.005$,	$\theta_{HS} = 17$,
$\delta^{*0}_C = 0.001$,	$\delta_S = 0.01$,	$\sigma_S = 20$,
$\delta^{**}_C = 0.02$,	$\delta_H = 0.02$.	

Table D.3. Criterion Weights

$u(y^0_E, y^0_S, y^0_H, y^0_C) = 0$,	$u(y^*_E, y^0_S, y^0_H, y^0_C) = 0.25$,
$u(y^0_E, y^*_S, y^0_H, y^0_C) = 0.2$,	$u(y^*_E, y^*_S, y^0_H, y^0_C) = 0.5$,
$u(y^0_E, y^0_S, y^*_H, y^0_C) = 0.5$,	$u(y^*_E, y^0_S, y^*_H, y^0_C) = 0.75$,
$u(y^0_E, y^*_S, y^*_H, y^0_C) = 0.7$,	$u(y^*_E, y^*_S, y^*_H, y^0_C) = 0.85$,
$u(y^0_E, y^0_S, y^0_H, y^1_C) = 0.05$,	$u(y^*_E, y^0_S, y^0_H, y^1_C) = 0.3$,
$u(y^0_E, y^*_S, y^0_H, y^1_C) = 0.25$,	$u(y^*_E, y^*_S, y^0_H, y^1_C) = 0.55$,
$u(y^0_E, y^0_S, y^*_H, y^1_C) = 0.55$,	$u(y^*_E, y^0_S, y^*_H, y^1_C) = 0.8$,
$u(y^0_E, y^*_S, y^*_H, y^1_C) = 0.75$,	$u(y^*_E, y^*_S, y^*_H, y^1_C) = 1$.

Appendix E. Expected Utility Vectors in Section 7.3

E.1. Entries of the Potential $\Psi^{C_2}_{C_1}$ for the Evaluation of the DEUN in Figure 7

$$\begin{aligned}
&u(y^*_S, y^*_E, y^*_H, y^*_C)u(y_H)u(y_E \mid y^*_H)u(y_C \mid y^*_E, y^*_H)(S - m_S)/(M_S - m_S), \\
&u(y^*_S, y^*_E, y^*_H, y^0_C)u(y_H)u(y_E \mid y^*_H)\hat{u}(y_C \mid y^*_E, y^*_H)(S - m_S)/(M_S - m_S), \\
&u(y^*_S, y^*_E, y^0_H, y^*_C)\hat{u}(y_H)u(y_E \mid y^*_H)u(y_C \mid y^*_E, y^*_H)(S - m_S)/(M_S - m_S), \\
&u(y^*_S, y^*_E, y^0_H, y^0_C)\hat{u}(y_H)u(y_E \mid y^*_H)\hat{u}(y_C \mid y^*_E, y^*_H)(S - m_S)/(M_S - m_S), \\
&u(y^*_S, y^0_E, y^*_H, y^*_C)u(y_H)\hat{u}(y_E \mid y^*_H)u(y_C \mid y^*_E, y^*_H)(S - m_S)/(M_S - m_S), \\
&u(y^*_S, y^0_E, y^*_H, y^0_C)u(y_H)\hat{u}(y_E \mid y^*_H)\hat{u}(y_C \mid y^*_E, y^*_H)(S - m_S)/(M_S - m_S), \\
&u(y^*_S, y^0_E, y^0_H, y^*_C)\hat{u}(y_H)\hat{u}(y_E \mid y^*_H)u(y_C \mid y^*_E, y^*_H)(S - m_S)/(M_S - m_S), \\
&u(y^*_S, y^0_E, y^0_H, y^0_C)\hat{u}(y_H)\hat{u}(y_E \mid y^*_H)\hat{u}(y_C \mid y^*_E, y^*_H)(S - m_S)/(M_S - m_S), \\
&u(y^0_S, y^*_E, y^*_H, y^*_C)u(y_H)u(y_E \mid y^*_H)u(y_C \mid y^*_E, y^*_H)(M_S - S)/(M_S - m_S), \\
&u(y^0_S, y^*_E, y^*_H, y^0_C)u(y_H)u(y_E \mid y^*_H)\hat{u}(y_C \mid y^*_E, y^*_H)(M_S - S)/(M_S - m_S), \\
&u(y^0_S, y^*_E, y^0_H, y^*_C)\hat{u}(y_H)u(y_E \mid y^*_H)u(y_C \mid y^*_E, y^*_H)(M_S - S)/(M_S - m_S), \\
&u(y^0_S, y^*_E, y^0_H, y^0_C)\hat{u}(y_H)u(y_E \mid y^*_H)\hat{u}(y_C \mid y^*_E, y^*_H)(M_S - S)/(M_S - m_S), \\
&u(y^0_S, y^0_E, y^*_H, y^*_C)u(y_H)\hat{u}(y_E \mid y^*_H)u(y_C \mid y^*_E, y^*_H)(M_S - S)/(M_S - m_S), \\
&u(y^0_S, y^0_E, y^*_H, y^0_C)u(y_H)\hat{u}(y_E \mid y^*_H)\hat{u}(y_C \mid y^*_E, y^*_H)(M_S - S)/(M_S - m_S), \\
&u(y^0_S, y^0_E, y^0_H, y^*_C)\hat{u}(y_H)\hat{u}(y_E \mid y^*_H)u(y_C \mid y^*_E, y^*_H)(M_S - S)/(M_S - m_S), \\
&u(y^0_S, y^0_E, y^0_H, y^0_C)\hat{u}(y_H)\hat{u}(y_E \mid y^*_H)\hat{u}(y_C \mid y^*_E, y^*_H)(M_S - S)/(M_S - m_S).
\end{aligned}$$

References

Abbas AE (2009) From Bayes' nets to utility nets. Goggans PM, Chan C-Y, eds. *Proc. 29th Internat. Workshop Bayesian Inference Maximum Entropy Methods Sci. Engrg.* (American Institute of Physics, Oxford, MS), 3–12.

Abbas AE (2010) General decompositions of multiattribute utility functions with partial utility independence. *J. Multicriteria Decision Anal.* 17(1–2):37–59.

Abbas AE (2011) The multiattribute utility tree. *Decision Anal.* 8(3):180–205.

Abbas AE, Howard RA (2005) Attribute dominance utility. *Decision Anal.* 2(4):185–206.

Barons M, Wright SK, Smith JQ (2016) Eliciting probabilistic judgements for integrating decision support systems. Technical report No. 16-05, University of Warwick, Coventry, UK.

Bickel G, Nord M, Price C, Hamilton W, Cook J (2000) Guide to measuring household food security. Technical report, Food and Nutrition Service, U.S. Department of Agriculture, Washington, DC.

Braziunas D, Boutilier C (2005) Local utility elicitation in GAI models. *Proc. 21st Conf. Uncertainty Artificial Intelligence* (AUAI Press, Edinburgh), 42–49.

Cowell RG, Dawid AP, Lauritzen SL, Spiegelhalter DJ (2007) *Probabilistic Networks and Expert Systems* (Springer, New York).

Dawid AP (1979) Conditional independence in statistical theory. *J. Roy. Statist. Soc. Series B* 41(1):1–31.

Engel Y, Wellman MP (2008) CUI networks: A graphical representation for conditional utility independence. *J. Artificial Intelligence Res.* 31(1):83–112.

Fishburn PC (1967) Interdependence and additivity in multivariate, unidimensional expected utility theory. *Internat. Econom. Rev.* 8(3):335–342.

Giang PH, Shenoy PP (2005) Two axiomatic approaches to decision making using possibility theory. *Eur. J. Oper. Res.* 162(2): 450–467.

Gómez M (2004) Real-world applications of influence diagrams. Gámez JA, Cerdan AS, Moral S, eds. *Advances in Bayesian Networks* (Springer, Berlin), 161–180.

Gonzales C, Perny P (2004) GAI networks for utility elicitation. *Proc. 9th Internat. Conf. Principles Knowledge Representation Reasoning* (AAAI Press, Whistler, British Columbia), 224–233.

Heckerman D, Mamdani A, Wellman MP (1995) Real-world applications of Bayesian networks. *Comm. Assoc. Computing Machinery* 38(3):24–26.

Hong DH, Choi CH (2000) Multicriteria fuzzy decision-making problems based on vague set theory. *Fuzzy Sets Systems* 114(1):103–113.

Howard RA, Matheson JE (2005) Influence diagrams. *Decision Anal.* 2(3):127–143.

Jensen F, Jensen FV, Dittmer SL (1994) From influence diagrams to junction trees. *Proc. 10th Conf. Uncertainty Artificial Intelligence, Seattle* (Morgan Kaufmann Publishers Inc., San Francisco), 367–373.

Keeney RL (1974) Multiplicative utility functions. *Oper. Res.* 22(1): 22–34.

Keeney RL, Raiffa H (1993) *Decisions with Multiple Objectives: Preferences and Value Trade-Offs* (Cambridge University Press, Cambridge, UK).

Kitchen S, Tanner E, Brown V, Colin P, Crawford C, Deardon L, Greaves E, Purdon S (2013) Evaluation of the free school meals pilot: Impact report. Technical report DFERR227, Department for Education, Manchester, UK.

La Mura P, Shoham Y (1999) Expected utility networks. *Proc. 15th Conf. Uncertainty Artificial Intelligence, Stockholm* (Morgan Kaufmann Publishers Inc., San Francisco), 366–373.

Lauritzen SL (1996) *Graphical Models* (Clarendon Press, Oxford, UK).

Leonelli M, Smith JQ (2015) Bayesian decision support for complex systems with many distributed experts. *Ann. Oper. Res.* 235(1):517–542.

Leonelli M, Riccomagno E, Smith JQ (2017) Using computer algebra to symbolically evaluate discrete influence diagrams. Technical report No. 15-07, University of Warwick, Coventry, UK.

Nielsen TD, Jensen FV (2009) *Bayesian Networks and Decision Graphs* (Springer, New York).

Pearl J (1988) *Probabilistic Reasoning in Intelligent Systems* (Morgan Kaufmann Publishers, Los Altos, CA).

Shachter RD (1986) Evaluating influence diagrams. *Oper. Res.* 34(6):871–882.

Smets P (2002) Decision making in a context where uncertainty is represented by belief functions. Srivastava RP, Mock TJ, eds. *Belief Functions in Business Decisions* (Springer, Heidelberg), 17–61.

Smith JQ (2010) *Bayesian Decision Analysis: Principles and Practice* (Cambridge University Press, Cambridge, UK).

Smith JQ, Thwaites P (2008) Influence diagrams. Melnick EL, Everitt BS, eds. *Encyclopedia of Quantitative Risk Analysis and Assessment* (Wiley, Chichester, UK), 897–910.

Smith JQ, Barons MJ, Leonelli M (2015a) Coherent frameworks for statistical inference serving integrating decision support systems. Technical report No. 15-10, University of Warwick, Coventry, UK.

Smith JQ, Barons MJ, Leonelli M (2015b) Decision focused inference on networked probabilistic systems: With applications to food security. *Amer. Statist. Assoc. Joint Statist. Meeting, Seattle*, 3220–3233.

Sullivant S, Talaska K, Draisma J (2010) Trek separation for Gaussian graphical models. *Ann. Statist.* 38(3):1665–1685.

Manuele Leonelli is a lecturer in statistics in the School of Mathematics and Statistics at the University of Glasgow and holds a Ph.D. in statistics from the University of Warwick. His primary research interests lie in Bayesian inference and decision theory, probabilistic graphical models, sensitivity analysis, and extreme value theory. Prior to joining the University of Glasgow, he was a Capes Postdoctoral Fellow in the Department of Mathematics at the Federal University of Rio de Janeiro.

Jim Q. Smith is a full professor of statistics at Warwick University and received a Ph.D. from there in 1977. His research mainly focuses on the development of dynamic, high-dimensional probabilistic models to support decision makers and big data. He is currently a fellow of the Alan Turing Institute and part of the Applied Statistics and Risk Unit at Warwick. His most recent book, *Bayesian Decision Analysis: Principles and Practice*, was published by Cambridge University Press in 2010.