An Axiomatic Framework for Influence Diagram Computation with Partially Ordered Utilities

Nic Wilson
Cork Constraint Computation Centre
University College Cork, Ireland
n.wilson@4c.ucc.ie

Radu Marinescu
IBM Research
Dublin, Ireland
radu.marinescu@ie.ibm.com

Abstract

This paper presents an axiomatic framework for influence diagram computation, which allows reasoning with partially ordered values of utility. We show how an algorithm based on sequential variable elimination can be used to compute the set of maximal values of expected utility (up to an equivalence relation). Formalisms subsumed by the framework include decision making under uncertainty based on multi-objective utility, or on interval-valued utilities, as well as a more qualitative decision theory based on order-of-magnitude probabilities and utilities.

1 Introduction

Actions can lead to many different kinds of consequences, for example, financial gain/loss, risk to health, effect on the environment or gain/loss to reputation. It may not be possible to map the various potential consequences of a set of actions to the same scale of utility in a way that avoids making essentially arbitrary choices. It is thus natural to consider notions of imprecise utility and multi-attribute/objective utility, where utility values are only partially ordered.

We consider decision making under uncertainty using influence diagrams, but where we allow more general notions of uncertainty than probability, and more general notions of utility functions, which, in particular, allow utility values to be only partially ordered. We construct an axiomatic framework, listing properties of a formalism that allow maximal (generalised) expected utility to be computed by sequential elimination of all the variables.

In general terms, variable elimination algorithms can be viewed as follows. We have a collection Θ of functions, where each function in Θ only involves a small number of variables. In the case of influence diagram computation, Θ contains both probability functions and utility functions. Θ is used as a compact representation (or decomposition) of a function $\bigotimes\Theta$ on all the variables, equalling a combination of all the functions in Θ. For example, in a Bayesian network, Θ consists of a collection of conditional probability functions and $\bigotimes\Theta$ is the joint probability distribution. Since $\bigotimes\Theta$ involves all the variables, it will be a huge object to represent explicitly.

For a standard influence diagram we have both chance and decision variables, and we eliminate chance variables with a sum operator, and decision variables with a max operator. We can compute the maximum expected utility by applying a sequence of sum and max eliminations to $\bigotimes\Theta$, eliminating all the variables. Performing combinations leads to functions involving larger sets of variables, which is expensive in terms of both computational cost and time. One therefore would like to delay performing computations where possible. Thus, when eliminating a variable X, with, for example, the $\sum$ operator, one transforms Θ to a collection Θ', which includes only functions that don't involve X, and is such that $\sum_X(\bigotimes\Theta) = \bigotimes\Theta'$. Crucially, the functions in Θ that don't involve X are left unchanged, so still appear in Θ'.

We first define (Section 2) our axioms for uncertainty and utility functions, which are the properties that we will use for the variable elimination algorithms. We describe some formalisms that satisfy the axioms, including interval-valued utility, multi-objective utility and order of magnitude probability and utility (Section 3). Section 4 shows how both chance variables and decision variables are eliminated. To eliminate a variable involves replacing the current collection of generalised probability and utility functions with a new set whose combination is equivalent to the marginal of the initial set. Section 5 defines influence diagram systems, which involve the form of uncertainty and utility functions that are generated by an influence diagram. Section 6 shows that one can iteratively eliminate all the variables to obtain the maximum value of expected utility for the case where utility values are totally ordered.

We go on to consider the case where values of utility are only partially ordered (Section 7); then there will often not be a unique maximal value of (expected) utility, but a set of them. To compute this set we need to perform operations on sets of utility values. In Section 8 we give the main result of the paper, which shows how to compute, by sequential variable elimination, a set of utility values that is equivalent, in a particular sense, to the set of maximal values of expected utility. Section 9 discusses related work, and Section 10 concludes. The Appendix includes some of the proofs, with more complete proofs being in a supplementary document (Wilson and Marinescu 2012).

2 Uncertainty And Utility Functions

In standard influence diagrams, probability potentials take non-negative real values, and utility functions take real values. In Section 2.1 we give properties of generalised probability and utility values which will still allow correct variable elimination algorithms.

Most of the properties of positive reals are still assumed for generalised probability values, the most important exception being that we do not assume a cancellation property for addition. The properties assumed for generalised utility values are much weaker than those satisfied by the real numbers; in particular, we allow partially ordered utility values, which will be useful for expressing imprecise information about utilities, or multi-objective utilities.

Section 2.2 defines the generalised probability and utility functions, along with combination and marginalisation operators, generalising definitions for standard probability and utility.

2.1 Uncertainty-Utility Values Structures

An *uncertainty values structure* is defined to be a tuple $\langle Q, +, \times, 0, 1\rangle$ that is a positive commutative semiring with multiplicative inverses. In other words: $+$ and $\times$ are both commutative and associative binary operations on set Q, with additive identity element 0 (so $q+0=q$ for all $q \in Q$), multiplicative identity element 1 (so $q \times 1 = q$ for all $q \in Q$) which also satisfies $q \times 0 = 0$ for all $q \in Q$, and, for all $p, q \in Q$, $p + q = 0$ if and only if $p = q = 0$; also, $\times$ distributes over $+$, i.e., $(p+q) \times r = (p \times r) + (q \times r)$. Furthermore, it is assumed that multiplicative inverses exist for all non-zero elements of Q, so that for all $q \in Q - \{0\}$ there exists some (unique) element $q^{-1} \in Q$ with $q \times q^{-1} = 1$.

A *utility values structure* is defined to be a tuple $\langle U, +, 0\rangle$ such that $+$ is a commutative and associative binary operation U with identity element 0.

A *weak uncertainty-utility values structure* (abbreviated to *weak u.u.v. structure*) is defined to be a tuple $\mathfrak{U} = \langle Q, +_Q, \times_Q, 0_Q, 1, U, +_U, 0_U, \times_{QU}\rangle$, where $\langle Q, +_Q, \times_Q, 0_Q, 1\rangle$ is an uncertainty values structure, $\langle U, +_U, 0_U\rangle$ is a utility values structure, and $\times_{QU}$ is a function from $Q \times U \to U$ satisfying the properties $(*1)$, $(*2)$ and $(*3)$ below, for arbitrary $u, u_1, u_2 \in U$ and $q, q_1, q_2 \in Q$ (dropping the $(\cdot)_Q$ and $(\cdot)_U$ subscripts since there is no ambiguity):

$(*1)$ $1 \times u = u$ and $0 \times u = 0_U$;

$(*2)$ $q_1 \times (q_2 \times u) = (q_1 \times q_2) \times u$;

$(*3)$ $q \times (u_1 + u_2) = (q \times u_1) + (q \times u_2)$.

We further define $\mathfrak{U}$ to be an *uncertainty-utility values structure* (u.u.v. structure) if it is a weak uncertainty-utility values structure and satisfies also:

$(*4)$ $(q_1 + q_2) \times u = (q_1 \times u) + (q_2 \times u)$.

Q will contain the probability-like values, and U will contain the utility-like values.

Disjunctive Operations: When we are eliminating decision variables in an influence diagram computation, we use a $\max$ operator. We generalise this to a disjunctive operation, as defined below. For the totally ordered case (see the end of Section 6), we use $\max$ as the disjunctive operator; for the partially ordered case in Section 7, we use $\max$ over subsets of utility values.

Let $\mathfrak{U} = \langle Q, +_Q, \times_Q, 0_Q, 1, U, +_U, 0_U, \times_{QU}\rangle$ be a weak u.u.v. structure. Let $\vee$ be a binary operation on U. We say that $\vee$ is a disjunctive operation for $\mathfrak{U}$ if $\vee$ is a commutative and associative operation on U such that both $+_U$ and $\times_{QU}$ distribute over $\vee$, so that for any $q \in Q$ and all $u_1, u_2, u_3 \in U$, $u_1 + (u_2 \vee u_3) = (u_1 + u_2) \vee (u_1 + u_3)$, and $q \times (u_1 \vee u_2) = (q \times u_1) \vee (q \times u_2)$.

Operation respecting ordering: Let T be some set and let $\odot$ be some function from $T \times U$ to U. We say that $\odot$ respects binary relation $\succeq$ on U if for all $t \in T$ and all elements u_1, u_2 of U, $u_1 \succeq u_2 \Rightarrow t \odot u_1 \succeq t \odot u_2$. (This can be also be viewed as $\odot$ being monotonic with respect to $\succeq$.)

We say that $\mathfrak{U}$ respects $\succeq$ if both $+_U$ and $\times_{QU}$ respect $\succeq$.

Suppose that $\mathfrak{U}$ respects total order $\succeq$. If we define $\vee$ to be maximum with respect to $\succeq$, then $\vee$ is a disjunctive operation.

2.2 Combining And Marginalising Uncertainty And Utility

We consider a set of variables that is partitioned into $\mathbf{X}$ and $\mathbf{D}$, where the elements of $\mathbf{X}$ are known as *chance variables*, and the elements of $\mathbf{D}$ are known as *decision variables*. Each element $Y \in \mathbf{X} \cup \mathbf{D}$ has an associated set of possible values Ω_Y. We also write, for $S \subseteq \mathbf{X} \cup \mathbf{D}$, Ω_S for the Cartesian product of Ω_Y over $Y \in S$.

Let $\mathfrak{U} = \langle Q, +_Q, \times_Q, 0_Q, 1, U, +_U, 0_U, \times_{QU}\rangle$ be a weak u.u.v. structure. A **$\mathfrak{U}$-uncertainty function** over $\mathbf{X} \cup \mathbf{D}$ is a function $\mathbf{P}$ from Ω_S to Q, for some $S \subseteq \mathbf{X} \cup \mathbf{D}$, known as the *scope* of $\mathbf{P}$, and denoted by $sc(\mathbf{P})$.

A **$\mathfrak{U}$-utility function** over variables $\mathbf{X} \cup \mathbf{D}$ is a function $\mathbf{U}$ from Ω_S to U, for some $S \subseteq \mathbf{X} \cup \mathbf{D}$, where S is the *scope* $sc(\mathbf{U})$ of $\mathbf{U}$.

We say that $\mathbf{P}$ *involves* Y if $sc(\mathbf{P}) \ni Y$. If $T \supseteq S = sc(\mathbf{P})$ and $\mathbf{x} \in \Omega_T$ then we also write $\mathbf{P}(\mathbf{x})$ as an abbreviation for $\mathbf{P}(\mathbf{x}^{\downarrow S})$, where $\mathbf{x}^{\downarrow S}$ is the projection of $\mathbf{x}$ to variables S. Similarly, for $\mathfrak{U}$-utility function $\mathbf{U}$.

Let Φ be a collection (i.e., a multiset) of $\mathfrak{U}$-uncertainty functions over $\mathbf{X} \cup \mathbf{D}$. We define their combination $\prod \Phi = \prod_{\mathbf{P} \in \Phi} \mathbf{P}$ to be the $\mathfrak{U}$-uncertainty function with scope $S = \bigcup_{\mathbf{P} \in \Phi} sc(\mathbf{P})$, given by, for $\mathbf{x} \in \Omega_S$, $(\prod \Phi)(\mathbf{x}) = \prod_{\mathbf{P} \in \Phi}(\mathbf{P}(\mathbf{x}))$, where the last use of $\prod$ refers to repeated application of the (associative and commutative) operation $\times_Q$.

Analogously, we define, for collection Ψ of $\mathfrak{U}$-utility functions over $\mathbf{X} \cup \mathbf{D}$, their combination $\sum \Psi = \sum_{\mathbf{U} \in \Psi} \mathbf{U}$ to have scope $S = \bigcup_{\mathbf{U} \in \Psi} sc(\mathbf{U})$, and to be given by $(\sum \Psi)(\mathbf{x}) = \sum_{\mathbf{U} \in \Psi} \mathbf{U}(\mathbf{x})$, where the last $\sum$ refers to iterative use of $+_U$.

If $\mathbf{P}$ is a $\mathfrak{U}$-uncertainty function and $\mathbf{U}$ is a $\mathfrak{U}$-utility function then $\mathbf{P} \times \mathbf{U}$ is a $\mathfrak{U}$-utility function with scope $S = sc(\mathbf{P}) \cup sc(\mathbf{U})$ given by $(\mathbf{P} \times \mathbf{U})(\mathbf{x}) = \mathbf{P}(\mathbf{x}) \times_{QU} \mathbf{U}(\mathbf{x})$, for any $\mathbf{x} \in \Omega_S$. (Thus $\mathfrak{U}$-utility functions can represent expected utility as well as input utility functions.)

Generating a marginalisation operator from an operation on utility values: Suppose that $\odot$ is some commutative and associative operation on U. Let $\mathbf{U}$ be a $\mathfrak{U}$-utility function and let $Y \in \mathbf{X} \cup \mathbf{D}$ be a variable in the scope of $\mathbf{U}$. We define $\bigodot_Y \mathbf{U}$ to be the $\mathfrak{U}$-utility function with scope $S = sc(\mathbf{U}) - \{Y\}$ given by $(\bigodot_Y \mathbf{U})(\mathbf{x}) = \bigodot_{y \in \Omega_Y} \mathbf{U}(\mathbf{x}y)$, for $\mathbf{x} \in \Omega_S$, where $\mathbf{x}y$ is assignment $\mathbf{x}$ extended with $Y = y$. This defines operation $\sum_Y$, based on operation $\odot = +$, and $\bigvee_Y$, based on disjunctive operation $\vee$.

3 Example Formalisms

We give some examples of formalisms that satisfy the axioms defined above. The results given later in this paper imply that, for any of these formalisms, the maximal values of expected utility can be computed (up to equivalence) by a variable elimination algorithm.

3.1 Upper And Lower Utility

It can sometimes be hard to determine precisely a utility value. To allow a representation of imprecise utility, we can let utility functions assign pairs of utility values instead of single values, representing a lower and an upper value of utility. For an influence diagram computation (see Section 5), each policy π, dynamically assigning the decision variables, has an associated lower expected utility LEU(π) and an upper expected utility UEU(π).

A natural ordering on these (expected) utility pairs is the pointwise one given by $\langle u, v\rangle \succeq \langle u', v'\rangle$ if and only if $u \geq u'$ and $v \geq v'$. That is, π is at least as good as π' if LEU(π) $\geq$ LEU(π') and UEU(π) $\geq$ UEU(π'). We then define U to consist of the set of all pairs $\langle u, v\rangle$ of real numbers with $u \leq v$. The sum of two pairs $\langle u_1, v_1\rangle$ and $\langle u_2, v_2\rangle$ is performed pointwise, i.e., to be $\langle u_1 + u_2, v_1 + v_2\rangle$. The additive identity utility element is the pair $\langle 0, 0\rangle$. For probability value p and utility pair $\langle u, v\rangle$, $p \times \langle u, v\rangle$ is defined to be $\langle p \times u, p \times v\rangle$. This gives rise to an u.u.v. structure $\langle \mathbb{R}^+, +, \times, 0, 1, U, +_U, \langle 0, 0\rangle, \times_{QU}\rangle$ that respects the partial order $\succeq$. Since utility values are only partially ordered, we may well have more than one maximal value of expected utility; we are interested in being able to compute the set of maximal (i.e., undominated) values of $\big\langle \mathrm{LEU}(\pi), \mathrm{UEU}(\pi)\big\rangle$ over all policies π.

Multi-Objective Utility A related system (which is mathematically a generalisation) is based on multi-objective utility. One may have more than one independent scale of utility, for example, one based on monetary gain, and one based on risk to health. Again, scalar multiplication and addition are performed pointwise, and we can use the same idea for the ordering. Alternatively, since the product (Pareto) ordering is a rather weak one, instead one may want to consider imprecise trade-offs between the scales of utility. The ordering can be strengthened to take such trade-offs into account, whilst still maintaining the monotonicity properties.

A further related system is multi-agent probability and utilities. Each of a number m of agents makes a judgement of the probability and utility values, which are each represented as vectors of m real values.

3.2 Order Of Magnitude Calculus

We consider the order of magnitude probability and utility system from (Wilson 1995), which can be viewed as a decision theory for kappa (ranking) functions (Goldszmidt and Pearl 1996). (A very different linear utility theory for kappa functions is given in (Giang and Shenoy 2000).) Let $\mathcal{O} = \{\langle \sigma, n\rangle : n \in \mathbb{Z}, \sigma \in \{+, -, \pm\}\} \cup \{\langle 0, \infty\rangle\}$, where $\mathbb{Z}$ is the set of integers. The element $\langle \pm, \infty\rangle$ will sometimes be written as 0, and element $\langle +, 0\rangle$ as 1. We also define $\mathcal{O}_\pm = \{\langle \pm, n\rangle : n \in \mathbb{Z} \cup \{\infty\}\}$, and $\mathcal{O}_+ = \{\langle +, n\rangle : n \in \mathbb{Z}\}$.

Elements of $\mathcal{O}$ are interpreted in terms of polynomials (or rational functions) in a parameter ε. ε can be considered as an infinitesimal, or, alternatively, a very small unknown number. $\langle +, n\rangle$ represents a function which is positive and of order ε^n, and $\langle -, n\rangle$ is negative and of order ε^n. When we add a positive and a negative value, both of order ε^n, the answer can be positive or negative, and of order ε^m for any $m \geq n$. We write this imprecise value as $\langle \pm, n\rangle$.

Multiplication: For $\langle \sigma, m\rangle, \langle \sigma', n\rangle \in \mathcal{O}$, let $\langle \sigma, m\rangle \times \langle \sigma', n\rangle = \langle \sigma \otimes \sigma', m + n\rangle$, where $\infty + m = m + \infty = \infty$ for $m \in \mathbb{Z} \cup \{\infty\}$, and $\otimes$ is the natural multiplication of signs: it is the commutative operation on $\{+, -, \pm\}$ such that $+ \otimes - = -$, $+ \otimes + = - \otimes - = +$, and for any $\sigma \in \{+, -, \pm\}$, $\sigma \otimes \pm = \pm$.

Addition: Commutative operation $+$ is given by $\langle \sigma, m\rangle + \langle \sigma', n\rangle = \langle \sigma, m\rangle$ if $m < n$, and equals $\langle \sigma \oplus \sigma', m\rangle$ if $m = n$, where $+ \oplus + = +$, $- \oplus - = -$, and otherwise, $\sigma \oplus \sigma' = \pm$.

Ordering: We use a slightly stronger ordering than that defined in (Wilson 1995). Transitive relation $\succeq$ on $\mathcal{O}$ is given by: $\langle +, l\rangle \succeq \langle \sigma, m\rangle \succeq \langle -, n\rangle$, for any l, m, n such that $l, n \leq m$, and any $\sigma \in \{+, -, \pm\}$.

We will define an u.u.v. structure $\mathfrak{U}$ for the order of magnitude case. We define U to be $\mathcal{O}$ and Q to be $\mathcal{O}_+ \cup \{0\}$. The previously stated properties including ($*1$), ($*2$), ($*3$), and ($*4$) all hold, and the operations respect the ordering.

Proposition 1. *Define $\mathfrak{U}^{\mathcal{O}}$ to be the tuple $\langle \mathcal{O}_+ \cup \{0\}, +, \times, 0, 1, \mathcal{O}, +, 0, \times\rangle$. Then $\mathfrak{U}^{\mathcal{O}}$ is an uncertainty-utility values structure that respects partial order $\succeq$, i.e., $\succeq$ is respected by $+$ and the operation $\times : \mathcal{O}_+ \cup \{0\} \times \mathcal{O} \to \mathcal{O}$.*

4 Elimination Of Variables

In this section we will consider an u.u.v. structure $\mathfrak{U}$, and a pair (Φ, Ψ), where Φ (respectively, Ψ) is a collection of $\mathfrak{U}$-uncertainty (respectively, $\mathfrak{U}$-utility) functions over $\mathbf{X} \cup \mathbf{D}$. We assume that each variable in $\mathbf{X} \cup \mathbf{D}$ is involved in some element of $\Phi \cup \Psi$.

A pair (Φ, Ψ) will be considered as a compact representation of the (overall) utility function $\prod_{\mathbf{P} \in \Phi} \mathbf{P} \times \sum_{\mathbf{U} \in \Psi} \mathbf{U}$. We write $\bigotimes(\Phi, \Psi) = \prod_{\mathbf{P} \in \Phi} \mathbf{P} \times \sum_{\mathbf{U} \in \Psi} \mathbf{U}$.

We will be interested in computing a generalised expected utility corresponding to the result of eliminating all variables from $\bigotimes(\Phi, \Psi)$. In Sections 4.1 and 4.2 we will show how to eliminate chance variables and decision variables.

In eliminating a chance variable X we generate a new pair (Φ', Ψ') that doesn't involve X such that $\sum_X \bigotimes(\Phi, \Psi) = \bigotimes(\Phi', \Psi')$. This involves combining functions that involve X; importantly, functions that don't involve X are left as they are. Similar remarks apply for the elimination of a decision variable.

4.1 Elimination Of A Chance Variable

Let $\Phi_{\not\ni X}$ be the multiset containing the elements of Φ not involving variable X, i.e., with $sc(\mathbf{P}) \not\ni X$, and let $\Phi_{\ni X}$ be the other elements in Φ. Let $\mathbf{P}^+ = \prod_{\mathbf{P} \in \Phi_{\ni X}} \mathbf{P}$ be the combination of elements of Φ involving variable X, and let $\mathbf{P}^X_\Phi$ be $\sum_X \mathbf{P}^+$. Let $\Phi' = \Phi_{\not\ni X} \cup \{\mathbf{P}^X_\Phi\}$.

Similarly, let $\Psi_{\not\ni X}$ be the multiset containing the elements of Ψ not involving X, and let $\Psi_{\ni X}$ be the other elements in Ψ. For $\mathbf{U} \in \Psi_{\ni X}$, define $\mathbf{U}^{-X}$ to be $\frac{1}{\mathbf{P}^X_\Phi} \times \sum_X (\mathbf{P}^+ \times \mathbf{U})$ (where, in this equation, we define $\frac{1}{0}$ to be 0). Let $\Psi' = \Psi_{\not\ni X} \cup \{\mathbf{U}^{-X} : \mathbf{U} \in \Psi_{\ni X}\}$.

We define $\sum_X (\Phi, \Psi)$ to be (Φ', Ψ'). Note that no element of Φ' or of Ψ' involves X. Also, in contrast with e.g., (Jensen, Jensen, and Dittmer 1994), we don't combine together the utility functions involving X when eliminating a chance variable X, since it's not necessary.

Theorem 1. *Let $\mathfrak{U}$ be an uncertainty-utility values structure, with associated summation operation $\sum$, let Φ be a collection of $\mathfrak{U}$-uncertainty functions over $\mathbf{X} \cup \mathbf{D}$, let Ψ be a collection of $\mathfrak{U}$-utility functions over $\mathbf{X} \cup \mathbf{D}$. Then for any $X \in \mathbf{X}$ which is involved in some element of Φ,*

$$\sum_X \bigotimes(\Phi, \Psi) = \bigotimes\Big(\sum_X (\Phi, \Psi)\Big).$$

In other words, using the definitions of Φ' and Ψ' given above,

$$\sum_X \Big(\prod_{\mathbf{P} \in \Phi} \mathbf{P} \times \sum_{\mathbf{U} \in \Psi} \mathbf{U}\Big) = \prod_{\mathbf{P} \in \Phi'} \mathbf{P} \times \sum_{\mathbf{U} \in \Psi'} \mathbf{U}.$$

4.2 Elimination Of A Decision Variable

We say that $\mathbf{P}$ *does not depend on variable* Y if for all $y, y' \in \Omega_Y$ and $\mathbf{x} \in \Omega_S$ where $S = sc(\mathbf{P}) - \{Y\}$, $\mathbf{P}(\mathbf{x}y) = \mathbf{P}(\mathbf{x}y')$. If so, we define $\mathbf{P}^{-Y}$ to have scope S and be given by $\mathbf{P}^{-Y}(\mathbf{x}) = \mathbf{P}(\mathbf{x}y)$ (for any $y \in \Omega_Y$).

Theorem 2. *Let $\vee$ be a disjunctive operation for weak uncertainty-utility values structure $\mathfrak{U}$. Let D be a decision variable in $\mathbf{D}$, let Φ be a collection of $\mathfrak{U}$-uncertainty functions over $\mathbf{X} \cup \mathbf{D}$, none of which depend on variable D, and let Ψ be a collection of $\mathfrak{U}$-utility functions over $\mathbf{X} \cup \mathbf{D}$. Define $\bigvee_D (\Phi, \Psi) = (\Phi^{-D}, \Psi'')$, where $\Phi^{-D} = \{\mathbf{P}^{-D} : \mathbf{P} \in \Phi\}$, and $\Psi'' = \Psi_{\not\ni D} \cup \{\bigvee_D \sum_{\mathbf{U} \in \Psi_{\ni D}} \mathbf{U}\}$. Then,*

$$\bigvee_D \bigotimes(\Phi, \Psi) = \bigotimes(\bigvee_D (\Phi, \Psi)),$$

i.e.,

$$\bigvee_D \Big(\prod_{\mathbf{P} \in \Phi} \mathbf{P} \times \sum_{\mathbf{U} \in \Psi} \mathbf{U}\Big) = \prod_{\mathbf{P} \in \Phi^{-D}} \mathbf{P} \times \sum_{\mathbf{U} \in \Psi''} \mathbf{U},$$

5 Influence Diagram Systems

Theorems 1 and 2 show how to eliminate chance variables and decision variables, respectively, using marginalisation operators $\sum_X$ and $\bigvee_D$ on a pair (Φ, Ψ). We would like to iteratively apply these to eliminate all variables, leading to the maximum expected utility. However, Theorem 2 requires that, before eliminating decision variable D, the current set of uncertainty functions does not depend on D. To ensure that this condition holds for the iterative computation, we require restrictions on the elimination ordering, as well as additional structure on the input collection of uncertainty functions Φ.

Let $\mathfrak{U}$ be a weak u.u.v. structure, and let $\mathbf{P}$ be an $\mathfrak{U}$-uncertainty function with scope $S \subseteq \mathbf{X} \cup \mathbf{D}$. We say that $\mathbf{P}$ is a *constant function* if it does not depend on any variable, i.e., there exists some value $q \in Q$ such that for all $\mathbf{x} \in \Omega_{sc(\mathbf{P})}$, $\mathbf{P}(\mathbf{x}) = q$. We say that $\mathbf{P}$ is a *conditional $\mathfrak{U}$-uncertainty function on* X if $X \in sc(\mathbf{P})$ and $\sum_X \mathbf{P}$ is a constant function.

An Influence Diagram system (ID-system) over $\mathfrak{U}$ is a pair $\langle G, (\Phi, \Psi)\rangle$, such that:

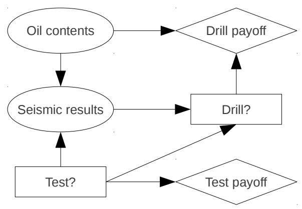
- G is a directed acyclic graph on $\mathbf{X} \cup \mathbf{D}$.
- $\Phi = \{\mathbf{P}_X : X \in \mathbf{X}\}$, where $\mathbf{P}_X$ is a conditional $\mathfrak{U}$-uncertainty function on X with scope $\{X\} \cup pa_G(X)$; ($pa_G(X)$ means the set of parents of X, i.e., $\{Y : (Y, X) \in G\}$);
- each element of Ψ is a $\mathfrak{U}$-utility function (whose scope is a subset of $\mathbf{X} \cup \mathbf{D}$).
- G restricted to $\mathbf{D}$ is a total order, which we write as $D_1, \ldots, D_m$.
- If $(X, D_i) \in G$ then $(X, D_j) \in G$ for any $j > i$ (this is the *no forgetting* condition).

Collection Φ is a Bayesian network-style decomposition of a global uncertainty distribution, namely $\prod \Phi$. Also, $\sum \Psi$ represents the overall utility function. Hence the function $\bigotimes(\Phi, \Psi)$ represents the (generalised) probability times utility function.

$(Y, D_i) \in G$ means that the value of Y is known when choosing the value of decision variable D_i. Let S_i be $pa_G(D_i)$ for $i = 1, \ldots, m$. The choice of value of D_i can therefore depend on the values of variables in S_i but no others.

A *policy* for this ID-system is a sequence $(\pi_1, \ldots, \pi_m)$ where π_i is a function from Ω_{S_i} to Ω_{D_i}; this represents what value of D_i is chosen given the available information: the already observed chance variables, and previous choices of decision variables.

A policy π determines a value for each decision variable D_i (which depends on the parents set S_i). Given a utility function $\mathbf{U}$ involving all the chance variables $\mathbf{X}$, a policy π determines a utility function $[\mathbf{U}]_\pi$ that involves no decision variables, by assigning their values using π. The expected utility given policy π is given by $EU_\pi = \sum_{\mathbf{X}} [\bigotimes(\Phi, \Psi)]_\pi$. If the utility values set U is totally ordered with relation $\succeq$ then maximum expected utility is $\max_\pi EU_\pi$, where $\max$ is taken with respect to $\succeq$.



Oil contents	P(O)		
	dry	wet	soak
	0 5	0.3	0.2

Test payoff	U_1(T)
Test?	
yes	-10
no	0

Seismic results		P(S | O,T)			
Oil cnt.	Test?	closed	open	diffuse	notest
dry	yes	0.1	0.3	0 6	0
dry	no	0	0	0	1
wet	yes	0.3	0.4	0 3	0
wet	no	0	0	0	1
soak	yes	0.5	0.4	0.1	0
soak	no	0	0	0	1

Drill payoff		U_2(O,D)
Oil cnt.	Drill?	
dry	yes	-70
dry	no	0
wet	yes	50
wet	no	0
soak	yes	200
soak	no	0

Figure 1: The oil wildcatter influence diagram.

For $k = 0, \ldots, m-1$, let $\mathbf{I}_k$ be the set of chance node parents of D_{k+1} that are not parents of any D_i, for $i \leq k$. Hence $\mathbf{I}_k$ is the set of chance nodes that D_{k+1} depends on, but no earlier D_i depends on. We also let $\mathbf{I}_m$ be the other chance nodes, that are not parents of any decision node.

A **legal elimination sequence** for an ID-system is a permutation $Y_1, \ldots, Y_n$ of the variables in $\mathbf{X} \cup \mathbf{D}$ that extends the relation $<$ given by: $\mathbf{I}_0 < D_1 < \mathbf{I}_1 < \cdots < D_m < \mathbf{I}_m$, so that, for $i = 0, \ldots, m$, each element X of $\mathbf{I}_i$ comes after D_i (if $i \geq 1$) and before D_{i+1} (if $i < m$). For example, if $Y_j \in D_1$ and $Y_k \in \mathbf{I}_1$ then we must have $j < k$. (Note that the sequence is not necessarily compatible with G, so we could have $i < j$ and $(Y_j, Y_i) \in G$, when Y_i and Y_j are both chance variables.)

Let τ be a sequence $Y_1, \ldots, Y_n$ of different elements in $\mathbf{X} \cup \mathbf{D}$, and let $\vee$ be a disjunctive operation for $\mathfrak{U}$. We define $\mathbb{M}_\tau^{+,\vee}(\mathbf{U})$ to mean the result of iterative application of the marginalisations corresponding to sequence τ, i.e., $\mathbb{M}_{Y_1}(\mathbb{M}_{Y_2}(\cdots(\mathbb{M}_{Y_n}\mathbf{U})\cdots))$ where $\mathbb{M}_Y$ is $\sum_Y$ if Y is a chance variable, and $\mathbb{M}_Y$ is $\bigvee_Y$ if Y is a decision variable. Note that the marginalisation operations are applied from right to left, so that the Y_n-marginalisation is performed first.

Suppose that $\mathfrak{U}$ respects total order $\succeq$ on U. As mentioned in Section 2.1, if we define $\vee$ to be maximum with respect to $\succeq$, then $\vee$ is a disjunctive operation. Then, for legal elimination sequence τ, $\mathbb{M}_\tau^{+,\vee}\big(\bigotimes(\Phi,\Psi)\big)$ can be shown (e.g., using Proposition 5 below) to be the maximum value of expected utility over all possible policies, i.e., $\max_\pi EU_\pi$.

Example 1. *Figure 1 shows the influence diagram of the oil wildcatter problem (adapted from (Raiffa 1968)). An oil wildcatter must decide either to* drill *or* not to drill *for oil at a specific site. Before drilling, they could perform a* seismic test *that will help determine the* geological structure *of the site. The test results can show a* closed *reflection pattern (indication of significant oil), an* open *pattern (indication of some oil), or a* diffuse *pattern (almost no hope of oil). The special value* notest *indicates that the test results will not be available if the seismic test is not done. There are therefore two decision variables, T (Test) and D (Drill), and two chance variables S (Seismic results) and O (Oil contents).*

The probabilistic knowledge consists of the conditional probability distributions $P(O)$ and $P(S|O,T)$, while the utility function is the sum of $U_1(T)$ and $U_2(D,O)$; thus we set $\Phi = \{P(O), P(S|O,T)\}$ and $\Psi = \{U_1(T), U_2(D,O)\}$. There is a unique legal elimination sequence: $\tau = T, S, D, O$ (so that O is eliminated first). The maximum expected utility, which equals $\mathbb{M}_\tau^{+,\vee}\big(\bigotimes(\Phi,\Psi)\big)$, is thus equal to $\bigvee_T \sum_S \bigvee_D \sum_O \big(\bigotimes(\Phi,\Psi)\big)$, where $\vee$ here means $\max$. The optimal policy is to perform the seismic test and to drill only if the test results show an open or a closed pattern. The expected utility of this policy is 22.5.

$$\pi_T = \textit{yes}\ ; \qquad \pi_D = \begin{cases} \textit{yes} & \textit{if } S = \textit{open} \\ \textit{yes} & \textit{if } S = \textit{closed} \\ \textit{no} & \textit{if } S = \textit{diffuse} \end{cases}$$

6 Elimination Of All Variables

We show (Theorem 3 below) how all variables can be iteratively eliminated for an ID-system. The proof uses iterative application of Theorems 1 and 2, where the main difficulty is in showing that, when eliminating a decision variable D, none of the uncertainty functions $\mathbf{P}$ depend on D (see the conditions of Theorem 2). The conditions assumed on Φ in an ID-system imply this.

Theorem 3. *Let $\mathfrak{U}$ be an uncertainty-utility values structure with operation $+$ on utility values, and let $\vee$ be a disjunctive operation for $\mathfrak{U}$. Let $\mathfrak{I} = \langle G, (\Phi,\Psi)\rangle$ be an ID-system over $\mathfrak{U}$, and let τ be a legal elimination sequence for $\mathfrak{I}$. Then*

$$\mathbb{M}_\tau^{+,\vee}\big(\bigotimes(\Phi,\Psi)\big) = \bigotimes\big(\mathbb{M}_\tau^{+,\vee}(\Phi,\Psi)\big).$$

The right-hand-side $\bigotimes\big(\mathbb{M}_\tau^{+,\vee}(\Phi,\Psi)\big)$ involves iterative local computations, based on sequential elimination of variables. When eliminating (marginalising out) a variable Y we only deal with functions involving Y: the ones not involving Y are just left as they are. (The operator $\bigotimes$ on the right-hand-side will be easy and often be redundant, since we'll typically just have a single utility value, representing expected utility, when we've eliminated all the variables.)

Application for Totally Ordered Case: Suppose that $\succeq$ is a total order on U such that $+_U$ and $\times_{QU}$ both respect $\succeq$, and again define $\vee$ to be maximum with respect to $\succeq$, which is a disjunctive operation. As observed above in Section 5, the left-hand-side $\mathbb{M}_\tau^{+,\vee}\big(\bigotimes(\Phi,\Psi)\big)$ in Theorem 3 is equal to maximum value of expected utility over all possible policies, i.e., $\max_\pi EU_\pi$. Hence Theorem 3 implies the correctness of an iterative variable elimination algorithm to compute maximum expected utility. This applies to standard probability and utility functions, and also to the simplified order of magnitude systems defined in Section 8.5 of (Wilson and Marinescu 2012).

The complexity is exponential in the largest scope of a function generated by the algorithm, which is no more than the induced width of G' (ordered according to the elimination sequence), where G' is G with an additional clique S for each input utility function with scope S.

7 Sets Of Utility Values For Partially Ordered Case

Let $\mathfrak{U}$ be an uncertainty-utility values structure and let $\mathfrak{I} = \langle G, (\Phi, \Psi) \rangle$ be an ID-system over $\mathfrak{U}$. Suppose that the set of utility values U is only partially ordered, by relation $\succeq$. For finite set A of utility values we can consider the set of maximal ones $\max_{\succeq}(A)$ consisting of all $u \in A$ such that there does not exist a different element $v \in A$ with $v \succeq u$. We are interested in policies that generate a maximal value of expected utility, i.e., values of utility in $\max_{\succeq} \{EU_\pi : \text{policies } \pi\}$. Towards this aim, we consider operations on sets of utility values (in an analogous way to the approach taken in Section 4.3 of (Fargier, Rollon, and Wilson 2010) for the much simpler case where there are just utility functions, and in Section 3 of (Perny, Spanjaard, and Weng 2005)).

7.1 Operations On Sets Of Utility Values

Consider an uncertainty-utility values structure $\mathfrak{U} = \langle Q, +_Q, \times_Q, 0_Q, 1, U, +_U, 0_U, \times \rangle$. We extend addition and scalar multiplication of utilities to 2^U, the set of subsets of U, in the obvious way. For any $A, B \subseteq U$ and $q \in Q$, define:

$$A + B = \{a + b \,:\, a \in A, b \in B\};$$
$$q \times A = \{q \times u : u \in A\}.$$

Given the uncertainty-utility values structure $\mathfrak{U} = \langle Q, +_Q, \times_Q, 0_Q, 1, U, +_U, 0_U, \times_{QU} \rangle$, we define $\mathfrak{U}^*$ to be the tuple $\mathfrak{U}^* = \langle Q, +_Q, \times_Q, 0_Q, 1, 2^U, +, \{0_U\}, \times \rangle$ (using the operations on sets as just defined). We have the following result:

Proposition 2. *For any u.u.v. structure $\mathfrak{U}$, the associated tuple $\mathfrak{U}^*$ is a weak uncertainty-utility values structure.*

Unfortunately, property $(*4)$ (see Section 2.1) does not hold in general for sets, so $\mathfrak{U}^*$ is not necessarily an uncertainty-utility values structure. $(q_1 + q_2) \times A$ is equal, using the property $(*4)$ for $\mathfrak{U}$, to $\{(q_1 \times a) + (q_2 \times a) \,:\, a \in A\}$. On the other hand, $(q_1 \times A) + (q_2 \times A)$ equals $\{(q_1 \times a_1) + (q_2 \times a_2) \,:\, a_1, a_2 \in A\}$. This implies that $(q_1 + q_2) \times A$ is a subset of $(q_1 \times A) + (q_2 \times A)$. However, they will very often not be equal. To give a simple example with bi-objective utility, let $q_1 = q_2 = 0.5$, and let $A = \{(1,0),(0,1)\}$, using the pointwise operations on pairs of real numbers. $(q_1 + q_2) \times A = 1 \times \{(1,0),(0,1)\} = \{(1,0),(0,1)\}$, whereas $(q_1 \times A) + (q_2 \times A) = \{(0.5,0),(0,0.5)\} + \{(0.5,0),(0,0.5)\} = \{(1,0),(0.5,0.5),(0,1)\}$.

However, in Section 7.3 we will define an equivalence relation that ensures that property $(*4)$ holds up to equivalence.

7.2 Ordering On Sets Of Utilities

We assume a partial ordering $\succeq$ that is respected by $\mathfrak{U}$, i.e., $\succeq$ is a partial order on U and operations $+_U$ and $\times_{QU}$ respect $\succeq$. If $a \succeq b$ then we say that *a dominates b.* We write $\succ$ for the strict part of $\succeq$, so that $a \succ b$ if and only if $a \succeq b$ and $a \neq b$.

For $A \subseteq U$, define $\max_{\succeq}(A)$, *the maximal elements of* A, to consist of all $a \in A$ such that there does not exist $b \in A$ with $b \succ a$. Hence, $\max_{\succeq}(A)$ is the set of undominated elements of A.

For $A, B \subseteq U$ we say that $A \succcurlyeq B$ if every element of B is dominated by some element of A (so that A contains as least as large elements as B), i.e., if for all $b \in B$ there exists $a \in A$ with $a \succeq b$. Relation $\succcurlyeq$ on 2^U is a reflexive and transitive relation. We define equivalence relation $\approx$ by $A \approx B$ if and only if $A \succcurlyeq B$ and $B \succcurlyeq A$.

7.3 The Equivalence Relation $\equiv$ Between Utility Sets

The *convex closure* $\mathcal{C}(A)$ of a (finite or infinite) subset A of U is defined to consist of every element of the form $\sum_{i=1}^{k}(q_i \times a_i)$, where k is an arbitrary natural number, each a_i is in A, each q_i is in Q, where $\sum_{i=1}^{k} q_i = 1$.

We will argue that certain different sets of (expected) utility values can reasonably be considered as equivalent. First of all, if A contains elements u and v with $u \succ v$, then we can consider that A and $A - \{v\}$ are equivalent. The second consideration is based on convex closure. For clarity, let's consider the case where the uncertainty values are probability values (but that the utility values may be partially ordered, such as for multi-objective utility or interval-valued utilities). If an agent can generate an expected utility value u with policy π, and utility value u' with policy π' then they may choose an independent auxiliary event E (e.g. based on a random number generator such as rolling a die) with chance p, and choose π if E holds and π' otherwise. (From the outside it may not even be possible to tell that they are doing this, since we only see the choices they make.) The expected utility is then $pu + (1-p)u'$. More generally, if one can achieve any of a set A of expected utility values, one can generate any element of $\mathcal{C}(A)$ by using the same kind of procedure. Thus, A and $\mathcal{C}(A)$ are equivalent in a natural sense.

We define equivalence relation $\equiv$ on subsets of U by:

$$A \equiv B \text{ if and only if } \mathcal{C}(A) \approx \mathcal{C}(B).$$

Two sets of utility values are therefore considered equivalent if, for every convex combination of elements of one, there is a convex combination of elements of the other which is at least as good (with respect to the partial order $\succeq$ on U). We have $A \equiv \mathcal{C}(A)$, and, if A is finite then $A \equiv \max_{\succeq}(A)$. Moreover, if $A \equiv B$ then $\max_{\succeq}(\mathcal{C}(A)) = \max_{\succeq}(\mathcal{C}(B))$ (and the converse often holds). The equivalence relation is respected by scalar multiplication, addition and union of subsets of utility values:

Proposition 3. *Let A, B and C be subsets of U, and let q be an element of Q. Suppose that $A \equiv B$. Then (i) $q \times A \equiv q \times B$; (ii) $A + C \equiv B + C$; (iii) $A \cup C \equiv B \cup C$.*

As observed above, Property $(*4)$ does not necessarily hold for sets of utility values. However, it does hold for convex sets, and a corresponding property based on $\equiv$ holds generally:

Proposition 4. *Consider u.u.v. structure $\mathfrak{U}$, written as $\langle Q, +_Q, \times_Q, 0_Q, 1, U, +_U, 0_U, \times\rangle$, which respects the partial order $\succeq$. Then the associated weak u.u.v. structure $\mathfrak{U}^*$ satisfies the following variant of Property (*4): for all $q_1, q_2 \in Q$ and for all $A \in 2^U$, $(q_1 + q_2) \times A \equiv (q_1 \times A) + (q_2 \times A)$.*

8 Variable Elimination Based On Sets Of Utilities

Let $\mathfrak{U}$ be an u.u.v. structure that respects partial order $\succeq$ on U, and let $\mathfrak{U}^* = \langle Q, +_Q, \times_Q, 0_Q, 1, 2^U, +, \{0_U\}, \times\rangle$ be the induced weak u.u.v. structure. As well as the operation $+$ on sets of utility, we also define operation $+'$ on finite sets of utility values by $A +' B = \max_{\succeq}(A + B)$. Define operation $\vee$ on finite subsets of U by $A \vee B = \max_{\succeq}(A \cup B)$. $\vee$ is commutative and associative.

A $\mathfrak{U}$-utility function $\mathbf{U}$ can be mapped in the obvious way to a $\mathfrak{U}^*$-utility function $\mathbf{U}^*$ with the same scope, defined by $\mathbf{U}^*(\mathbf{x}) = \{\mathbf{U}(\mathbf{x})\}$. For collection Ψ of $\mathfrak{U}$-utility functions, define collection $\Psi^* = \{\mathbf{U}^* : \mathbf{U} \in \Psi\}$ of $\mathfrak{U}^*$-utility functions.

The following result shows that $u \in \mathbb{M}_\tau^{+,\cup}\big(\bigotimes(\Phi, \Psi^*)\big)$ if and only if there exists some policy whose expected utility is u. (The *no forgetting* condition, that the choice of value of a decision variable can depend on all the earlier chance variables, is crucial here.)

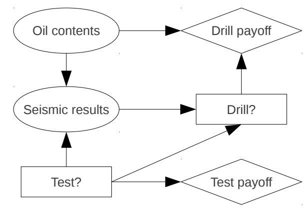
Proposition 5. *Let $\mathfrak{I} = \langle G, (\Phi, \Psi)\rangle$ be an $\mathfrak{U}$-ID-system, and let τ be a legal elimination sequence for $\mathfrak{I}$. Then $\mathbb{M}_\tau^{+,\cup}\big(\bigotimes(\Phi, \Psi^*)\big)$ is equal to the set of all possible values of expected utility over all policies for $\mathfrak{I}$, i.e., $\{\sum_{\mathbf{X}}[\bigotimes(\Phi, \Psi)]_\pi$: policies $\pi\}$.*

We now give the main result of the paper which shows how to use an iterative variable elimination algorithm to compute a set of utility values that is equivalent to the set of maximal values of expected utility. The idea of the proof is that, because of Propositions 2 and 4, $\mathfrak{U}^*$ is an u.u.v. structure modulo the equivalence relation. Scalar multiplication, addition and union respect equivalence (Proposition 3). Theorem 3 then can be applied for $\mathfrak{U}^*$ modulo equivalence; also union is $\equiv$-equivalent to $\vee$, and addition is $\equiv$-equivalent to $+'$.

Theorem 4. *Let $\mathfrak{U}$ be be an uncertainty-utility values structure (with operation $+$ on utility values) that respects partial order $\succeq$. As above, let $\vee$ be the operation induced from $\succeq$ on finite sets of utility values. Let $\mathfrak{I} = \langle G, (\Phi, \Psi)\rangle$ be an $\mathfrak{U}$-ID-system, and let τ be a legal elimination sequence for $\mathfrak{I}$. Then,*

$$\mathrm{Max}_{\succcurlyeq}\big(\mathbb{M}_\tau^{+,\cup}\big(\bigotimes(\Phi, \Psi^*)\big)\big) \equiv \bigotimes\big(\mathbb{M}_\tau^{+',\vee}(\Phi, \Psi^*)\big).$$

By Proposition 5, the left-hand-side is the set of all optimal (i.e., undominated) values of expected utility. The right-hand-side gives an iterative variable elimination algorithm for computing an equivalent set of utility values. This therefore applies for influence diagrams based on the systems described in Section 3, involving multi-objective utility theory, interval-valued utilities, or the order of magnitude system.



Oil contents	P(O)		
	dry	wet	soak
	0.5	0 3	0 2

Test payoff	U_1(T)
Test?	
yes	{(-10,10)}
no	{(0,0)}

Seismic results		P(S | O,T)			
Oil cnt.	Test?	closed	open	diffuse	notest
dry	yes	0.1	0 3	0.6	0
dry	no	0	0	0	1
wet	yes	0.3	0.4	0.3	0
wet	no	0	0	0	1
soak	yes	0.5	0.4	0.1	0
soak	no	0	0	0	1

Drill payoff		U_2(O,D)
Oil cnt.	Drill?	
dry	yes	{(-70,18)}
dry	no	{(0,0)}
wet	yes	{(50,12)}
wet	no	{(0,0)}
soak	yes	{(200,8)}
soak	no	{(0,0)}

Figure 2: A bi-objective influence diagram.

Example 2. *Figure 2 displays a bi-objective influence diagram corresponding to the oil wildcatter decision problem from Example 1. For our purpose, we add a second utility scale representing the* damage to environment *in addition to the one that represents the decision maker's* profit*. Therefore, our aim is to find optimal policies that simultaneously maximize profit and minimize the environmental impact. Given two utility vectors $\vec{u} = (u_1, u_2)$ and $\vec{v} = (v_1, v_2)$, the dominance relation in this case is defined by $\vec{u} \succcurlyeq \vec{v} \Leftrightarrow u_1 \geq v_1$ and $u_2 \leq v_2$. For example $(10, 2) \succcurlyeq (8, 4)$ but $(10, 2)$ does not dominate $(8, 1)$. We compute $\bigvee_T \sum'_S \bigvee_D \sum'_O \big(\bigotimes(\Phi, \Psi^*)\big)$, where $\Phi^* = \{U_1^*(T), U_2^*(D, O)\}$, and, e.g., $U_1^*(T = yes) = \{(-10, 10)\}$. The structure of the computation is just the same as for Example 1, but instead of standard utility functions (which assign a real value to each appropriate tuple) we have functions that assign a set of real pairs to each tuple. The Pareto set $\max_{\succcurlyeq}\{EU_\pi \mid$ policies $\pi\}$ contains four elements, namely $\{(22.5, 17.56), (20, 14.2), (11, 12.78), (0, 0)\}$, corresponding to the four optimal policies shown below.*

	π	π	π	π
Test?	*yes*	*no*	*yes*	*no*
Drill?	*yes (S = closed)* *yes (S = open)* *no (S = diffuse)*	*yes (S = notest)*	*yes (S = closed)* *no (S = open)* *no (S = diffuse)*	*no (S = notest)*
EU_π	$\{(22.5, 17.56)\}$	$\{(20, 14.2)\}$	$\{(11, 12.78)\}$	$\{(0, 0)\}$

Computation time and size of utility values sets: As for the totally ordered case (see Section 6), the complexity is exponential in the largest scope of a function generated by the algorithm. However, the main difference between the computation for the partially ordered case and that for standard influence diagrams is that, instead of a single utility values being stored (for each assignment to the scope of a utility function), we have a set of utility values. For the order of magnitude calculus case, this isn't a serious problem: for each finite set A of utility values, there exists set B containing at most two utility values with $B \equiv A$ (see Section 8.3 of (Wilson and Marinescu 2012)).

The order of magnitude influence diagrams have been implemented and experimentally tested, as described in (Marinescu and Wilson 2011). Problems with, for example, 50 bi-valued chance variables, 5 bi-valued decision variables and a single utility function defined over 5 variables, were solved in less than 1 minute by a dedicated variable elimination based algorithm. Additional experiments on selected classes of influence diagrams also showed that as the probabilistic and utility information becomes more precise, the qualitative decision process via order of magnitude influence diagrams becomes closer to the standard one.

For the multi-objective (or interval-valued) utility case, the cardinality of the utility sets can get very large. We experimented with random multi-objective influence diagrams with 5 decisions, n chance variables and involving 3, 5 and 10 objectives, respectively. The variable elimination based algorithm was able to solve only relatively small problem instances with n up to 25 and ran out of memory on larger instances. For example, problems with 25 chance variables, 5 decisions and 3 objectives were solved in about 15 minutes and had a Pareto set containing on average around 25,000 elements. Our experiments showed clearly that producing the entire Pareto set in this case is intractable even for relatively small problems. To overcome this difficulty, we also implemented an approximation algorithm that relies on the concept of ϵ-dominance between utility values as a relaxation of the Pareto dominance relation (Papadimitriou and Yannakakis 2000). Subsequent experiments on similar problem classes showed that the approximation algorithm scaled to much larger problem instances (with up to 60 chance variables) compared with the exact version, while producing significantly smaller sets of undominated utility values.

9 Related Work

The variable elimination approach we use in this paper is based on that for standard influence diagrams, in particular, in (Jensen, Jensen, and Dittmer 1994), which builds on previous work such as (Shachter and Peot 1992; Shenoy 1992). The work that is closest in spirit to the current work is that by Pralet, Schiex and Verfaillie (2009), who also consider an axiomatic framework for generalised influence diagrams (and other sequential decision making problems), involving a form of generalised expected utility (Chu and Halpern 2003). In contrast with our framework, Pralet et al.'s work does not assume division (multiplicative inverses) for the uncertainty values, which allows some more qualitative uncertainty formalisms to be reasoned with. (The existence of multiplicative inverses is used to generate properties we need for convex sets of utility values.) However, Pralet et al. focus on the case of totally ordered utility values, with the major contribution of our paper being for the partially ordered case.

Another general computation framework is Valuation Algebra/Networks (Kohlas 2003), building on work by Shenoy and Shafer (1990). Since it involves only one marginalisation operator it doesn't apply directly for solving influence diagrams, which require a marginalisation for eliminating chance nodes, and a different one for eliminating decision nodes. Prakash Shenoy (1992; 2000) has developed this kind of computational structure for solving influence diagrams based on standard probability and utility. A general axiomatic framework for solving Markov Decision Processes (which have a different and somewhat simpler structure than influence diagrams) is described in (Perny, Spanjaard, and Weng 2005); this framework also allows utilities to be only partially ordered. (Kikuti and Cozman 2007; Kikuti, Cozman, and Filho 2011) allow interval probabilities (which are not covered by our framework), and focus on precise utility; similarly, (de Campos and Ji 2008).

(Diehl and Haimes 2004) consider influence diagrams with multiple objectives (with just a single multi-objective value node), with the solution methods involving propagation of sets of utility vectors, as in the current paper, but based on influence diagram transformations (Shachter 1986). (López-Díaz and Rodríguez-Muñiz 2007) consider generalised influence diagrams based on fuzzy random variables.

10 Discussion

The major contribution of the paper is to show how a variable elimination method can be used to compute the set of optimal values of expected utility (up to equivalence) for influence diagrams based on formalisms which involve partially ordered utilities. The formalisms covered include all those discussed in Section 3: decision making under uncertainty based on multi-objective utility, a system of interval-valued utilities, and of multi-agent expected utility, as well as the Order of Magnitude system. Our implementations of these show that the approach is practical for problems of substantial size. More generally, Theorem 4 shows that variable elimination is sound for any formalism that satisfies the axioms given in Section 2, i.e., that the formalism forms an uncertainty-utility values (u.-u.v.) structure which respects the partial order on utility values. Thus to show that the variable elimination method applies for a formalism, one just has to verify these axioms (which are simple to check for a given formalism). Proposition 1 shows that the OOM system verifies the axioms (see also (Marinescu and Wilson 2011)). Propositions 18 and 19 in Sections 9.1 and 9.2 at the end of the (Wilson and Marinescu 2012) prove that the systems of multi-agent expected utility, multi-objective utility and of interval-valued utilities, satisfy the axioms. Hence the variable elimination algorithm is correct for all of these formalisms.

Our axiomatic framework complements that of (Pralet, Schiex, and Verfaillie 2009), in that, although it requires stronger conditions on the uncertainty functions, it allows utility values to be only partially ordered. As mentioned above, we have already implemented and tested our algorithm approach for the order of magnitude probability and utility case; and also for the the case of interval-valued utility and multi-objective utility without trade-offs. We are currently working on implementing the approach for multi-objective utility with trade-offs.

We consider in this paper a straight-forward variable elimination algorithm. One might improve this to efficiently make use of constraints (zero values of the uncertainty and utility functions), building, for instance, on the work of (Pralet, Schiex, and Verfaillie 2009).

Acknowledgements

This material is based upon works supported by the Science Foundation Ireland under Grant No. 08/PI/I1912.

Appendix

The appendix includes proofs of the four theorems, based on some auxiliary results. The proofs of the latter, and other results in the paper, can be found in the supplementary document (Wilson and Marinescu 2012). In particular, the full proofs of Theorems 1, 2, 3 and 4 are in Sections 3, 4, 6 and 7, respectively, of (Wilson and Marinescu 2012).

Proving the Chance Variable Elimination Result (Theorem 1)

We will consider an uncertainty-utility values structure $\mathfrak{U}$, and a pair (Φ, Ψ), where Φ (respectively, Ψ) is a collection of $\mathfrak{U}$-uncertainty (respectively, $\mathfrak{U}$-utility) functions over $\mathbf{X} \cup \mathbf{D}$. We assume that each variable in $\mathbf{X} \cup \mathbf{D}$ is involved in some element of $\Phi \cup \Psi$.

We use two lemmas.

Lemma 1. *Let $\mathbf{P}$, $\mathbf{P}_1$ and $\mathbf{P}_2$ be $\mathfrak{U}$-uncertainty functions over $\mathbf{X} \cup \mathbf{D}$ and let $\mathbf{U}$, $\mathbf{U}_1$ and $\mathbf{U}_2$ be $\mathfrak{U}$-utility functions over $\mathbf{X} \cup \mathbf{D}$. Let Φ be a finite (multi-)set of $\mathfrak{U}$-uncertainty functions, and let Ψ be a finite (multi-)set of $\mathfrak{U}$-utility functions. Let X be a variable.*

(i) $(\mathbf{P}_1 \times \mathbf{P}_2) \times \mathbf{U} = \mathbf{P}_1 \times (\mathbf{P}_2 \times \mathbf{U})$*;*

(ii) $\mathbf{P} \times (\mathbf{U}_1 + \mathbf{U}_2) = (\mathbf{P} \times \mathbf{U}_1) + (\mathbf{P} \times \mathbf{U}_2)$*; and, more generally,* $\mathbf{P} \times \sum_{\mathbf{U} \in \Psi} \mathbf{U} = \sum_{\mathbf{U} \in \Psi} (\mathbf{P} \times \mathbf{U})$*;*

(iii) $\sum_X (\mathbf{U}_1 + \mathbf{U}_2) = \sum_X \mathbf{U}_1 + \sum_X \mathbf{U}_2$*; more generally,* $\sum_X \sum_{\mathbf{U} \in \Psi} \mathbf{U} = \sum_{\mathbf{U} \in \Psi} \sum_X \mathbf{U}$*;*

(iv) $\sum_X (\mathbf{P} \times \mathbf{U}) = \left(\sum_X \mathbf{P}\right) \times \mathbf{U}$ *if sc*$(\mathbf{U}) \not\ni X$*, i.e., if* $\mathbf{U}$ *doesn't involve variable* X*;*

(v) $\sum_X (\mathbf{P} \times \mathbf{U}) = \mathbf{P} \times \sum_X \mathbf{U}$ *if sc*$(\mathbf{P}) \not\ni X$.

As well as the notation from Section 4.1, we use the following notation. Let $\mathbf{U}^- = \sum_{\mathbf{U} \in \Psi_{\not\ni X}} \mathbf{U}$, and let $\mathbf{U}^+ = \sum_{\mathbf{U} \in \Psi_{\ni X}} \mathbf{U}$. We further define

$$\mathbf{U}' = \frac{1}{\mathbf{P}_\Phi^X} \times \sum_X (\mathbf{P}^+ \times \mathbf{U}^+).$$

Lemma 2. *With the notation defined above, we have:*

(i) $\mathbf{U}' = \sum_{\mathbf{U} \in \Psi_{\ni X}} \frac{1}{\mathbf{P}_\Phi^X} \times \sum_X (\mathbf{P}^+ \times \mathbf{U})$*;*

(ii) $\mathbf{U}^- + \mathbf{U}' = \sum_{\mathbf{U} \in \Psi'} \mathbf{U}$.

(iii) $\sum_X (\mathbf{P}^+ \times \mathbf{U}^-) = \mathbf{P}_\Phi^X \times \mathbf{U}^-$.

(iv) $\sum_X (\mathbf{P}^+ \times \mathbf{U}^+) = \mathbf{P}_\Phi^X \times \mathbf{U}'$.

Proof: (i): By definition of $\mathbf{U}^+$, $\mathbf{U}'$ equals $\frac{1}{\mathbf{P}_\Phi^X} \times \sum_X (\mathbf{P}^+ \times \sum_{\mathbf{U} \in \Psi_{\ni X}} \mathbf{U})$. Using Lemma 1(ii) and then (iii), this equals $\frac{1}{\mathbf{P}_\Phi^X} \times \sum_{\mathbf{U} \in \Psi_{\ni X}} \sum_X (\mathbf{P}^+ \times \mathbf{U})$. Applying Lemma 1(ii) again gives the result.

(ii): By part (i), $\mathbf{U}'$ equals $\sum_{\mathbf{U} \in \Psi_{\ni X}} \mathbf{U}^{-X}$. Hence, $\sum_{\mathbf{U} \in \Psi'} \mathbf{U} = \sum_{\mathbf{U} \in \Psi_{\not\ni X}} \mathbf{U} + \sum_{\mathbf{U} \in \Psi_{\ni X}} \mathbf{U}^{-X} = \mathbf{U}^- + \mathbf{U}'$.

(iii) follows from Lemma 1(iv). (iv) follows immediately from the definition of $\mathbf{U}'$. □

Proof of Theorem 1 Let $\mathbf{P}^- = \prod_{\mathbf{P} \in \Phi_{\not\ni X}} \mathbf{P}$. Using Lemmas 1 and 2 we have:

$$\sum_X \left(\prod_{\mathbf{P} \in \Phi} \mathbf{P} \times \sum_{\mathbf{U} \in \Psi} \mathbf{U} \right) =^a \sum_X \Big(\mathbf{P}^- \times \big(\mathbf{P}^+ \times (\mathbf{U}^- + \mathbf{U}^+) \big) \Big)$$

$$=^b \mathbf{P}^- \times \sum_X \big(\mathbf{P}^+ \times (\mathbf{U}^- + \mathbf{U}^+) \big).$$

$$\sum_X \big(\mathbf{P}^+ \times (\mathbf{U}^- + \mathbf{U}^+) \big) =^c \sum_X \big(\mathbf{P}^+ \times \mathbf{U}^- \big) + \sum_X \big(\mathbf{P}^+ \times \mathbf{U}^+ \big)$$

$$=^d \mathbf{P}_\Phi^X \times (\mathbf{U}^- + \mathbf{U}').$$

Putting things together:

$$\sum_X \left(\prod_{\mathbf{P} \in \Phi} \mathbf{P} \times \sum_{\mathbf{U} \in \Psi} \mathbf{U} \right) = \mathbf{P}^- \times \mathbf{P}_\Phi^X \times (\mathbf{U}^- + \mathbf{U}')$$

$$=^e \prod_{\mathbf{P} \in \Phi'} \mathbf{P} \times \sum_{\mathbf{U} \in \Psi'} \mathbf{U}.$$

Equality $=^a$ uses Lemma 1 (i). Equality $=^b$ uses Lemma 1 (v). Equality $=^c$ uses Lemma 1 (ii) and (iii). Equality $=^d$ uses Lemma 2(iii) and (iv), and Lemma 1 (ii). Equality $=^e$ uses Lemma 2(ii).

Proving the Decision Variable Elimination Result (Theorem 2)

We'll use the following lemma. Part (i) follows easily. The proofs of parts (ii) and (iii) are very similar to the proofs of (v) and (iv), respectively, of Lemma 1, using the two distributivity properties of disjunctive operation $\vee$.

Lemma 3. *Let D be a decision variable in $\mathbf{D}$, let Φ be a collection of $\mathfrak{U}$-uncertainty functions over $\mathbf{X} \cup \mathbf{D}$, and let $\mathbf{P}$ be an $\mathfrak{U}$-uncertainty function over $\mathbf{X} \cup \mathbf{D}$. Let $\mathbf{U}_1$ and $\mathbf{U}_2$ be $\mathfrak{U}$-utility functions over $\mathbf{X} \cup \mathbf{D}$.*

(i) If for all $\mathbf{P} \in \Phi$, $\mathbf{P}$ *does not depend on D then* $(\prod \Phi)$ *does not depend on D and* $(\prod \Phi)^{-D} = \prod_{\mathbf{P} \in \Phi^{-D}} \mathbf{P}$.

(ii) If $\mathbf{P}$ *does not depend on D then* $\bigvee_D (\mathbf{P} \times \mathbf{U}) = \mathbf{P}^{-D} \times \bigvee_D \mathbf{U}$.

(iii) If $\mathbf{U}_1$ *does not involve D then* $\bigvee_D (\mathbf{U}_1 + \mathbf{U}_2) = \mathbf{U}_1 + \bigvee_D \mathbf{U}_2$.

Proof of Theorem 2 Applying Lemma 3 (i) and (ii) gives

$$\bigvee_D \Big(\prod_{\mathbf{P} \in \Phi} \mathbf{P} \times \sum_{\mathbf{U} \in \Psi} \mathbf{U} \Big) = \prod_{\mathbf{P} \in \Phi^{-D}} \mathbf{P} \times \bigvee_D \sum \Psi.$$

Applying Lemma 3 (iii) to $\mathbf{U}_1 = \sum \Psi_{\not\ni D}$ and $\mathbf{U}_2 = \sum \Psi_{\ni D}$ gives $\bigvee_D \sum \Psi = \sum \Psi_{\not\ni D} + \bigvee_D \sum \Psi_{\ni D}$, which equals $\sum \Psi''$, completing the proof.

Proving Elimination of all Variables Result (Theorem 3)

We use the following result.

Proposition 6. *Let $\mathfrak{I} = \langle G, (\Phi, \Psi) \rangle$ be an ID-system over $\mathfrak{U}$, and let τ be a legal elimination sequence for $\mathfrak{I}$. Let τ_k be the part of τ starting just after D_k, i.e., if τ is $Y_1, \ldots, Y_n$ and $Y_j = D_k$ then τ_k is the sequence $Y_{j+1}, \ldots, Y_n$. Let (Φ_k, Ψ_k) be $\mathbb{M}^{+,\vee}_{\tau_k}(\Phi, \Psi)$ (where if τ_k is empty, $\Phi_k = \Phi$).*

Then, $\mathbb{M}^{+,\vee}_{\tau}(\Phi, \Psi)$ is well defined, that is, for each $k = m, \ldots, 1$ and each $\mathbf{P} \in \Phi_k$, $\mathbf{P}$ does not depend on D_k.

Proof of Theorem 3: Let us write $\mathbb{M}^{+,\vee}_{\tau}(\cdot)$ as $\mathbb{M}_{Y_1}(\mathbb{M}_{Y_2}(\cdots(\mathbb{M}_{Y_n}(\cdot))\cdots))$. The left-hand-side of the theorem involves a sequence of $n+1$ operators, with $\bigotimes$ being the rightmost (i.e., the one applied first). The right-hand-side involves the same operators, but with $\bigotimes$ being applied at the end, i.e., leftmost. The idea of the proof is to use Theorems 1 and 2 to move the $\bigotimes$ incrementally from right to left.

For i, j with $1 \leq i \leq j \leq$ let $\mathbb{M}^{+,\vee}_{\tau[i:j]}$ be an abbreviation for the operator: $\mathbb{M}_{Y_i}(\mathbb{M}_{Y_{i+1}}(\cdots(\mathbb{M}_{Y_j}(\cdot))\cdots))$. We also define $\mathbb{M}^{+,\vee}_{\tau[1:0]}$ to be the identity operator (i.e., that makes no change to its operand).

For $i = 0, \ldots, n-1$, let (Φ^i, Ψ^i) be $\mathbb{M}^{+,\vee}_{\tau[i+1:n]}(\Phi, \Psi)$, and so $(\Phi^0, \Psi^0) = \mathbb{M}^{+,\vee}_{\tau[1:n]}(\Phi, \Psi)$, which equals $\mathbb{M}^{+,\vee}_{\tau}(\Phi, \Psi)$. We also define $(\Phi^n, \Psi^n) = (\Phi, \Psi)$.

We'll show, for any $i = 1, \ldots, n$, $\mathbb{M}_{Y_i}(\bigotimes(\Phi^i, \Psi^i)) = \bigotimes(\mathbb{M}_{Y_i}(\Phi^i, \Psi^i))$.

If Y_i is a chance variable, so that $\mathbb{M}_{Y_i}$ is $\sum_{Y_i}$, then this follows by Theorem 1. Otherwise, Y_i is a decision variable, and $\mathbb{M}_{Y_i}$ is $\bigvee_{Y_i}$. By Proposition 6, Φ^i doesn't depend on variable Y_i, so we can apply Theorem 2 to give $\mathbb{M}_{Y_i}(\bigotimes(\Phi^i, \Psi^i)) = \bigotimes(\mathbb{M}_{Y_i}(\Phi^i, \Psi^i))$ in this case also.

Applying operator $\mathbb{M}^{+,\vee}_{\tau[1:i-1]}$ to both sides of this equation, and using the fact that $\mathbb{M}_{Y_i}(\Phi^i, \Psi^i) = (\Phi^{i-1}, \Psi^{i-1})$, we have, for all $i = 1, \ldots, n$,

$$\mathbb{M}^{+,\vee}_{\tau[1:i]}(\bigotimes(\Phi^i, \Psi^i)) = \mathbb{M}^{+,\vee}_{\tau[1:i-1]}(\bigotimes(\Phi^{i-1}, \Psi^{i-1})).$$

Chaining the equalities for $i = n, \ldots, 1$, we obtain $\mathbb{M}^{+,\vee}_{\tau[1:n]}(\bigotimes(\Phi^n, \Psi^n)) = \mathbb{M}^{+,\vee}_{\tau[1:0]}(\bigotimes(\Phi^0, \Psi^0)) = \bigotimes(\Phi^0, \Psi^0)$, i.e., $\mathbb{M}^{+,\vee}_{\tau}(\bigotimes(\Phi, \Psi)) = \bigotimes(\mathbb{M}^{+,\vee}_{\tau}(\Phi, \Psi))$, proving the result.

Proving Theorem 4

We first need to extend operations $+'$ and $\vee$, to enable Lemma 4 below.

Extending operations $+'$ and $\vee$ to infinite sets: For $A \subseteq U$ we define subset $\mathcal{R}_{\succeq}(A)$ of A to consist of all elements of A which are not strictly dominated by some maximal element of A, that is:

$$\mathcal{R}_{\succeq}(A) = \{a \in A \,:\, \nexists b \in \max\nolimits_{\succeq}(A) \quad \text{such that} \quad b \succ a\}.$$

Clearly, we always have $\max_{\succeq}(A) \subseteq \mathcal{R}_{\succeq}(A)$. If A is such that every element of A is dominated by some maximal element of A (in particular, this is the case if A is finite), then $\mathcal{R}_{\succeq}(A) = \max_{\succeq}(A)$.

Proposition 7. *For any subset A of U, $A \equiv \mathcal{R}_{\succeq}(A)$ and $A \equiv \mathcal{C}(A)$. If A is finite then $A \equiv \max_{\succeq}(A)$.*

We define operation $\vee$ on subsets of U by $A \vee B = \mathcal{R}_{\succeq}(A \cup B)$. Analogously, we define operation $+'$ on subsets of U by $A +' B = \mathcal{R}_{\succeq}(A + B)$.

If A and B are finite, $\mathcal{R}_{\succeq}(A \cup B) = \max_{\succeq}(A \cup B)$ and $\mathcal{R}_{\succeq}(A + B) = \max_{\succeq}(A + B)$, so this agrees with the definitions of $\vee$ and $+'$ given in Section 8 for finite sets.

The following lemma is an immediate consequence of the fact that for any $A \subseteq U$, $\mathcal{R}_{\succeq}(A) \equiv A$ (see Proposition 7).

Lemma 4. *For any subsets A, B of U, $A +' B \equiv A + B$, and $A \vee B \equiv A \cup B$.*

Factoring $\mathfrak{U}^*$ by equivalence $\equiv$: Let us write $\mathfrak{U}^*$ as $\langle Q, +_Q, \times_Q, 0_Q, 1, 2^U, +, \{0_U\}, \times \rangle$. We define $2^U_{\equiv}$ to be the set of all $\equiv$-equivalence classes of 2^U. For $A \in 2^U$, we can write $[A]$ to mean the equivalence class containing A. The operations $+$ and $\cup$ on 2^U, and the scalar multiplication $\times$ all respect the equivalence relation $\equiv$, by Proposition 3. Hence they give rise to well-defined operations on $2^U_{\equiv}$ (which we use the same symbols for), and we have, for $A, B \subseteq U$, and $q \in Q$, $[A] + [B] = [A + B]$; $[A] \cup [B] = [A \cup B]$; and $q \times [A] = [q \times A]$.

Let us define the quotient $\mathfrak{U}^*/\equiv$ to be the tuple $\langle Q, +_Q, \times_Q, 0_Q, 1_Q, 2^U_{\equiv}, +_U, [\{0_U\}], \times \rangle$.

The key result below follows using Propositions 2, 4 and the fact that the union operation $\cup$ is a disjunctive operation for $\mathfrak{U}^*$.

Proposition 8. *$\mathfrak{U}^*/\equiv$ is an uncertainty-utility values structure, and $\cup$ is a disjunctive operation for $\mathfrak{U}^*/\equiv$.*

The result below follows immediately from Lemma 4.

Lemma 5. *For any subsets A, B of U, $[A] +' [B] = [A] + [B]$, and $[A] \vee [B] = [A] \cup [B]$. Hence $+'$ and $+$ are the same operation on $2^U_{\equiv}$, and $\vee$ and $\cup$ are the same operation on $2^U_{\equiv}$.*

Proof of Theorem 4: Probability-utility functions collection (Φ, Ψ^*) over $\mathfrak{U}^*$ maps to a probability utility functions collection over $\mathfrak{U}^*/\equiv$, which we write as $([\Phi], [\Psi^*])$.

Proposition 8 and Theorem 3 imply that

$$\mathbb{M}^{+,\cup}_{\tau}\big(\bigotimes\big([\Phi], [\Psi^*]\big)\big) = \bigotimes\big(\mathbb{M}^{+,\cup}_{\tau}\big([\Phi], [\Psi^*]\big)\big).$$

Now, by Lemma 5, $+'$ over $\mathfrak{U}^*/\equiv$ is exactly the same operation as $+$ over $\mathfrak{U}^*/\equiv$, and, similarly, $\cup$ and $\vee$ are the same operation over $\mathfrak{U}^*/\equiv$. So we have:

$$\mathbb{M}^{+,\cup}_{\tau}\big(\bigotimes\big([\Phi], [\Psi^*]\big)\big) = \bigotimes\big(\mathbb{M}^{+',\vee}_{\tau}\big([\Phi], [\Psi^*]\big)\big).$$

This implies that

$$\mathbb{M}^{+,\cup}_{\tau}\big(\bigotimes\big(\Phi, \Psi^*\big)\big) \equiv \bigotimes\big(\mathbb{M}^{+',\vee}_{\tau}\big(\Phi, \Psi^*\big)\big).$$

The left-hand-side L of this is a finite subset of U, and so $\mathrm{Max}_{\succeq}(L) \equiv L$, completing the proof. (Note that the use of $\vee$ in the right-hand-side is always on finite sets so, for these, $A \vee B = \max_{\succeq}(A \cup B)$.)

References

Chu, F. C., and Halpern, J. Y. 2003. Great expectations. Part I: On the customizability of Generalized Expected Utility. In *Proc. IJCAI'03*, 291–296.

de Campos, C. P., and Ji, Q. 2008. Strategy selection in influence diagrams using imprecise probabilities. In *UAI-08*, 121–128.

Diehl, M., and Haimes, Y. Y. 2004. Influence diagrams with multiple objectives and tradeoff analysis. *IEEE Transactions On Systems, Man, and Cybernetics Part A* 34(3):293–304.

Fargier, H.; Rollon, E.; and Wilson, N. 2010. Enabling local computation for partially ordered preferences. *Constraints* 15(4):516–539.

Giang, P. H., and Shenoy, P. 2000. A qualitative linear utility theory for spohn's theory of epistemic beliefs. In *Proc. UAI'00*, 220–229.

Goldszmidt, M., and Pearl, J. 1996. Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artif. Intell.* 84(1-2):57–112.

Jensen, F.; Jensen, F.; and Dittmer, S. 1994. From influence diagrams to junction trees. In *UAI-94*, 367–363.

Kikuti, D., and Cozman, F. G. 2007. Influence diagrams with partially ordered preferences. In *Proceedings of 3rd Multidisciplinary Workshop on Advances in Preference Handling*.

Kikuti, D.; Cozman, F. G.; and Filho, R. S. 2011. Sequential decision making with partially ordered preferences. *Artificial Intelligence, in press, doi:10.1016/j.artint.2010.11.017*.

Kohlas, J. 2003. *Information Algebras: Generic Structures for Inference*. Springer-Verlag.

López-Díaz, M., and Rodríguez-Muñiz, L. J. 2007. Influence diagrams with super value nodes involving imprecise information. *European Journal of Operational Research* 179(1):203–219.

Marinescu, R., and Wilson, N. 2011. Order-of-magnitude influence diagrams. In *UAI-11*, 489–496.

Papadimitriou, C., and Yannakakis, M. 2000. On the approximability of trade-offs and optimal access to web sources. In *IEEE Symp. on FOCS*, 86–92.

Perny, P.; Spanjaard, O.; and Weng, P. 2005. Algebraic Markov Decision Processes. In *Proc. IJCAI'05*, 1372–1377.

Pralet, C.; Schiex, T.; and Verfaillie, G. 2009. *Sequential Decision-Making Problems—Representation and Solution*. Wiley.

Raiffa, H. 1968. *Decision analysis*. Addison-Wesley.

Shachter, R., and Peot, M. 1992. Decision making using probabilistic inference methods. In *UAI-92*, 276–283.

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

Shenoy, P. P., and Shafer, G. 1990. Axioms for probability and belief function propagation. In *Uncertainty in Artificial Intelligence 4*, 575–610.

Shenoy, P. 1992. Valuation-based systems for Bayesian decision analysis. *Operations Research* 40(1):463–484.

Shenoy, P. P. 2000. Valuation network representation and solution of asymmetric decision problems. *European Journal of Operational Research* 121(3):579–608.

Wilson, N., and Marinescu, R. 2012. *An axiomatic framework for Influence Diagram computation for partially ordered utilities: proofs and auxiliary material*. Unpublished report, available at http://www.4c.ucc.ie/∼rmarines/kr2012-proofs.pdf.

Wilson, N. 1995. An order of magnitude calculus. In *UAI-95*, 548–555.