Directional Decomposition of Multiattribute Utility Functions

Ronen I. Brafman[1] and Yagil Engel[2]

[1] Department of Computer Science, Ben Gurion University, Israel
brafman@cs.bgu.ac.il
[2] Industrial Engineering & Management, Technion, Israel
yagile@ie.technion.ac.il

Abstract. Several schemes have been proposed for compactly representing multiattribute utility functions, yet none seems to achieve the level of success achieved by Bayesian and Markov models for probability distributions. In an attempt to bridge the gap, we propose a new representation for utility functions which follows its probabilistic analog to a greater extent. Starting from a simple definition of marginal utility by utilizing reference values, we define a notion of conditional utility which satisfies additive analogues of the chain rule and Bayes rule. We farther develop the analogy to probabilities by describing a directed graphical representation that relies on our concept of conditional independence. One advantage of this model is that it leads to a natural structured elicitation process, very similar to that of Bayesian networks.

1 Introduction

Specifying a multi-variate utility function is known to be a difficult task, and often considered a bottleneck in implementation of intelligent systems. It requires quantifying one's preferences – a non-trivial cognitive task which involves contemplating a large number of questions about the relative desirability of uncertain outcomes, or gambles. Furthermore, the very personal and subjective nature of utility information makes it harder to reuse and learn, unlike probabilistic knowledge, which can often be learned from data and reused for various instances of a system. Yet, the preference and utility elicitation tasks must be carried out when analyzing decision problems. A number of attempts have been made to aid this elicitation process by structuring it so that either the type of questions that must be answered is simpler and/or the number of questions is smaller. Often, this process is aided by some graphical structure that captures some properties of the utility function.

The level of success of current formalisms is not clear, partly because assessing the benefit of various utility elicitation processes is difficult. Even given that current models provide significant theoretical simplifications, the cognitive burden imposed by the elicitation process may still be prohibitive for practical applications. Given the well recognized practical benefits yielded by probabilistic graphical models, it is likely that much more can be done for utilities, too. In this work we attempted to follow the footsteps of probability theory more closely than before, by defining a notion of conditional utility that is closer in form to its probabilistic analog. We then show how this concept leads naturally to milestones such as the chain rule and Bayes rule analogies, and finally to a graphical representation based on a directed acyclic graph. While our new method of representing and eliciting utilities bears certain similarities to existing methods, as detailed below, it offers an elicitation process – both for qualitative structure and numeric values – that is clear, simple, and intuitive. Furthermore, it provides immediate computational benefits, and several promising direction for future research that are based on the close resemblance to probabilistic models. We believe that this method can become an essential part of the toolkit of decision analysts and an important component in real-world decision support systems.

In the remainder of this paper we define a new notion of conditional utility and utilize it to define *utility difference networks*. We explain their elicitation process and compare them to existing formalism for representing structured utility functions. Finally, we discuss a few open questions.

2 Background and Related Work

Let Θ denote the space of possible outcomes, with $\preceq$ a preference relation (weak total order) over Θ. Let $\Gamma = \{a_1, \ldots, a_n\}$ denote a set of attributes describing Θ. Each attribute $a \in \Gamma$ has a domain $\mathcal{D}(a)$, so that $\Theta \subseteq \prod_{i=1}^{n} \mathcal{D}(a_i)$. We use prime signs and superscripts to denote specific assignment for an attribute, and a concatenation of assignment symbols (as in $a'_i a''_j$) means that each of the attributes gets a respective value. We use γ and γ_i to denote subsets of Γ, and the same notation as before to denote assignments to all the attributes in the set. For example, if $\gamma_1 = \{a_i, a_j\}$, then $\gamma_1^0 = a_i^0 a_j^0$. Finally, we use $\mathcal{D}(\gamma)$ to denote the set of all possible assignments to γ, that is the projection of Θ over $\prod_{a_i \in \gamma} \mathcal{D}(a_i)$.

Definition 1. *Let $\gamma_1, \gamma_2 \subset \Gamma$. γ_1 and γ_2 are* conditionally additive independent (CAI) *given their complement $\Gamma \setminus (\gamma_1 \cup \gamma_2)$, if preferences over lotteries on Γ depend only on their marginal conditional probability distributions over γ_1 and γ_2.*

Graphical models have been employed for the representation of decomposed utility, as early as by (Gorman, 1968; Tatman and Shachter, 1990). However, the first representation that relies on *conditional independence*, and thus follow the footsteps of probabilistic models, can be attributed to Bacchus and Grove (1995). These authors show that conditional additive independence has a perfect map, meaning that given a set of attributes and a preference order, there exists a graph whose node separation expresses the exact set of independence conditions. Further, they show that the utility function decomposes to a sum over lower dimensional functions, each defined over a maximal clique of the graph. This decomposition is a special type of *generalized additive independence (GAI)*, a global independence condition introduced originally by Fishburn (1967).

Definition 2. *Let $\gamma_1, \ldots, \gamma_g \subseteq \Gamma$ such that $\bigcup_{i=1}^{g} \gamma_i = \Gamma$. $\gamma_1, \ldots, \gamma_g$ are called* generalized additive independent *(GAI) if preferences over lotteries on Γ depend only on their marginal distributions over $\gamma_1, \ldots, \gamma_g$.*

An (expected) utility function $u(\cdot)$ can be decomposed additively according to its (possibly overlapping) GAI sub-configurations.

Theorem 1 (Fishburn (1967)). *Let $\gamma_1, \ldots, \gamma_g$ be GAI. Then there exist functions $f_1, \ldots, f_g$ such that*

$$u(a_1, \ldots, a_m) = \sum_{r=1}^{g} f_r(\gamma_r). \tag{1}$$

Bacchus and Grove revived this notion and named it GAI. This opened the way to an increasing body of research on representation and reasoning with GAI. Boutilier et al. (2001) introduce UCP networks, which is a directed form of CAI-maps. The directionality though is obtained from identifying preferential independence conditions over sets of attributes, that is exogenously to the GAI decomposition. Gonzales and Perny (2004) introduce GAI nets, which is a graphical representation for GAI, where nodes represent subsets of attributes, and nodes are connected if their respective subsets intersect. Braziunas and Boutilier (2005) provide a method of elicitation that takes advantage of the locality property of GAI.

CAI and GAI require comparisons of probability distributions and preferences over lotteries. In applications in which uncertainty is not a crucial element (e.g., electronic commerce applications), it is not required and usually not desired to involve probabilities in user interaction. Engel and Wellman (2007) extend the work of Dyer and Sarin (1979) and introduce *conditional difference independence (CDI)*. Intuitively, attributes x and y are CDI of each other if any difference in value over assignments to x does not depend on the current assignment of y, for any possible assignment to the rest of the variables. CDI is very similar to CAI, and therefore has a perfect map as well.

Definition 3. [3] *Let $\gamma_1, \gamma_2 \subset \Gamma$. γ_1 and γ_2 are* conditionally difference independent *given $\gamma_3 = \Gamma \setminus (\gamma_1 \cup \gamma_2)$, denoted as CDI$(X, Y)$, if*

$$\forall \textit{ assignments } \hat{\gamma}_3, \gamma_1', \gamma_1'', \gamma_1', \gamma_2'' \quad u(\gamma_1'\gamma_2'\hat{\gamma}_3) - u(\gamma_1''\gamma_2'\hat{\gamma}_3) = u(\gamma_1'\gamma_2''\hat{\gamma}_3) - u(\gamma_1''\gamma_2''\hat{\gamma}_3)$$

Our new concept of independence and graphical model most closely resemble CDI. However, in comparison to CDI, it introduces several benefits: (i) it is directional, allowing for a more intuitive elicitation process and (ii) the independence condition is weaker, meaning it can be applied in some cases wherein which CDI does not hold.

Another direction of research relied on other types of utility independence. CUI networks (Engel and Wellman, 2008) is a graphical model that relies on the concept of *conditional utility independence* (Keeney, 1971), which intuitively requires the (cardinal) preference order over a subset of the attributes to be independent of another subset of attributes. Earlier works by Shoham (1997) and La Mura and Shoham (1999) are also seeking utility representation that is similar to a probability distribution. Shoham (1997) proposes a redefinition of utility function as a set function, over additive factors

[3] Difference independence and CDI are defined given a preference order over preference differences, and its numeric representation is a measurable value function. For brevity of presentation we describe it in terms of utilities.

in the domain that together contribute to the decision maker's well being. La Mura and Shoham (1999) propose only a redefinition of the utility independence concept, which is a multiplicative version of difference independence (that is, refers to utility ratios rather than differences). In non-probabilistic settings, and especially in situations in which decision outcomes can be measured against monetary differences (as in purchasing), we believe that utility differences are more natural to elicit than ratios.

A common drawback of most previous models is that most focus is given to the process of data elicitation, whereas the process of *structure elicitation*, in which the independence structure is identified, is usually left to domain experts. This is particularly true for GAI based representations, as UCP and GAI networks, because there is no explicit and intuitive process for identifying and/or verifying GAI conditions. Our novel model, in contrast, has the benefit of an intuitive and incremental structure elicitation process.

3 Reference and Conditional Utility

There are inherent differences between probability distributions and utility functions, which make any analogy between the two problematic. Arguably, the most primal difference is the fact that probability distribution is a set function, defined over events that encapsulate a set of atomic outcomes. In contrast, there is no meaning for the utility of a set of atomic outcomes. For probability distributions, there is a natural definition for a function over a subspace of the world on which the problem is defined. Technically, if the world is represented by a set of attributes $\mathit{\Gamma}$, one can define a probability distribution over some $\gamma \subset \mathit{\Gamma}$ by summing over the atomic outcomes that hold for any assignment to γ, thus marginalize out the irrelevant parameters (namely, $\mathit{\Gamma} \setminus \gamma$).

Whereas there is no meaning for marginalizing parameters of a utility function, a similar effect can be achieved by fixing those parameters on some *reference value*. For probabilities, we ask the question *what is the probability of outcomes in* γ *when we don't know the value of* $\mathit{\Gamma} \setminus \gamma$. While we do not have an exact analogy for utilities, with reference values we get ask: *what is the utility of outcomes in* γ *when the value of* $\mathit{\Gamma} \setminus \gamma$ *is fixed on the reference*. The idea of using a reference value has been exploited in previous works (Fishburn, 1967; Braziunas and Boutilier, 2005; Engel and Wellman, 2008), however it was never taken quite as far in driving the analogy to probabilities.

Let $a_1^0 \dots a_n^0 \in \mathit{\Theta}$ denote a predetermined complete assignment, which we call *the reference assignment*. The reference assignment allows us to define a utility function over a subspace of the joint domain. Let $\overline{\gamma} = \mathit{\Gamma} \setminus \gamma$.

Definition 4. *The* reference utility function *is defined as follows*

$$u_r(\gamma) = u(\gamma\overline{\gamma}^0)$$

The next step is to define the notion of conditioning, within a subspace of the domain.

Definition 5. *The* conditional utility function *is defined as follows*

$$u_r(\gamma_1|\gamma_2) = u_r(\gamma_1\gamma_2) - u_r(\gamma_1^0\gamma_2)$$

where $\overline{\gamma} = \mathit{\Gamma} \setminus \{\gamma_1 \cup \gamma_2\}$.

This definition has a direct rooting in the definition of conditional probabilities. The definition of the latter is

$$p(\gamma_1|\gamma_2) = \frac{p(\gamma_1\gamma_2)}{\gamma_2}$$

As common in probabilistic reasoning, we take a log of the definition in order to replace multiplication with additivity. This results exactly in Definition 5.

Given that, it is not surprising that the utility function exhibits an additive decomposition which is similar to the multiplicative decomposition of a probability function. We first have to normalize the utility function (henceforth) such that $u(\Gamma^0) = 0$.

Theorem 2 (The chain rule).

$$u(\Gamma) = \sum_{i=1}^{n} u_r(a_i|\{a_j\}_{j=1}^{i-1})$$

Proof. By definitions of conditional utility and reference utility,

$$u_r(a_i|\{a_j\}_{j=1}^{i-1}) = u_r(a_1 \dots a_{i-1}a_i) - u_r(a_1 \dots a_{i-1}a_i^0) = \\ u(a_1 \dots a_{i-1}a_i a_{i+1}^0 \dots a_n^0) - u(a_1 \dots a_{i-1}a_i^0 a_{i+1}^0 \dots a_n^0)$$

Summing over $i = 1, \dots, n$ on both sides yields the desired result, because: (1) the negative term for $i = 1$ is $u(a_1^0 \dots a_n^0) = 0$, (2) the negative term for i cancels out with the positive term for $i - 1$ (both are $u(a_1 \dots a_{i-1}a_i^0 a_{i+1}^0 \dots a_n^0)$), and (3) the positive term for $i = n$ is $u(\Gamma)$. □

Finally, it is easy to see that this definition obeys an additive adaptation of Bayes rule. Again, taking log over the probabilistic equation we obtain the following

Theorem 3 (Bayes Rule Analog).

$$u_r(\gamma_1|\gamma_2) = u_r(\gamma_2|\gamma_1) + u_r(\gamma_1) - u_r(\gamma_2)$$

3.1 Conditional Independence

The chain rule by itself does not provide significant computational value, because the last term ($i = n$) includes the left-hand side of the equation $u(\Gamma)$. The idea, similar to the one employed to achieve compact probability functions, is that the conditional utility function $u_r(a_i|a_1, \dots, a_{i-1})$ may not depend on all of the attributes $a_1, \dots, a_{i-1}$, but only on some subset of them, in which case the terms considered by the chain rule have lower dimensionality. This is formalized as follows.

Definition 6. γ_1 *is said to be* conditionally independent *of* γ_2 *given* γ_3 *(CDI*$_r(\gamma_1, \gamma_2|\gamma_3)$*) if for any* $\gamma_3' \in \mathcal{D}(\gamma_3)$*,*

$$u_r(\gamma_1|\gamma_2\gamma_3') = u_r(\gamma_1|\gamma_3')$$

When $\gamma_3 = \Gamma \setminus \gamma_1 \cup \gamma_2$, then CDI$_r(\gamma_1, \gamma_2|\gamma_3)$ is equivalent to γ_1 and γ_2 being CDI. Therefore, CDI$_r$ is a generalization of CDI. The novelty of this definition is that it refers to a subset of the attributes. Whereas in previous independence concepts the conditional set must always be "the rest of the attributes", here we specifically select a conditional set, and can ignore the attributes which are not relevant to γ_1.

As an example, consider the values in Table 1, which provides the value for the eight different instantiations of three boolean attributes, x, y, and z. The difference between the two values in each column corresponds to the difference in x given difference instantiations of yz.

x^0yz	x^0y^0z	x^0yz^0	$x^0y^0z^0$
9	6	6	3
xyz	xy^0z	xyz^0	xy^0z^0
12	7	8	5

Table 1: Utility for each assignment to attributes x, y, and z.

We see that CDI(x, y) does not hold because $u(xyz) - u(x^0yz) \neq u(xy^0z) - u(x^0y^0z)$ (according to the two left columns). In our terms it means that CDI$_r(x, y|z)$ does not hold. However, CDI$_r(x, y|)$ does hold, because the difference is equal for the reference value z^0 (see the two right columns): $u(xyz^0) - u(x^0yz^0) = u(xy^0z^0) - u(x^0y^0z^0)$ (or, equivalently, $u_r(x|y) = u_r(x|)$).

4 Utility Difference Networks

Loyal to the Bayes-net analogy, we seek a directed graphical structure, with a node for each attribute, and the following property: each attribute is conditionally directional independent, given its parents, of all its other non-descendants. Let $Pa(a)$ denote the parents of a node a in a graph, and let $Dn(a)$ denote its descendants. Furthermore, let $Co(a) = \Gamma \setminus \{a\} \cup Pa(a) \cup Dn(a)$.

Definition 7. *A* utility difference network *is a DAG $G = (V, E)$, with V corresponding to a set of attributes Γ, and for any $a \in \Gamma$, CDI$_r(a, Pa(a)|Co(a))$.*

The utility computation from the directed graph is again very similar to how probabilities are computed from a Bayes-net. The following theorem is a direct result of the chain rule and Definition 7.

Theorem 4. *The utility function can be computed from the utility difference network as follows*

$$u(\Gamma) = \sum_{i=1}^{n} u_r(a_i|Pa(a_i))$$

Previous graphical models usually assume that the model is given, obtained by some domain expert. In particular, how to identify a GAI decomposition remains an unsolved question (except for the case that the GAI structure is a result of a collection of CAI or CDI conditions). Note also that while each pairwise CDI condition requires in theory the verification of order of $\exp(n)$ equalities for utility differences (because a verification is required for each instantiation of the rest of the attributes), with our new notion of conditional independence we only need to consider the independent attributes and the conditioning set. However, we note that when creating a full network this is not a significant advantage, because the number of queries for the *last* variable in the ordering will reach the same order of magnitude as in CDI.

4.1 Elicitation

The process of obtaining a utility difference network structure is similar in spirit to that used for Bayesian networks. It is summarized by the following procedure. As is the case in Bayesian network, the result depends on the *variable ordering* that is used by the procedure, and the choice of variable ordering is usually based on heuristic assessments. Intuitively, we would like to place the most important variables first, because these are the variables that are likely to have many connections to other variables. By keeping them on top we avoid having to represent all of these dependencies as parents of the same variable. Furthermore, it makes intuitive sense to have the important variables first, so the dependence between other variables are conditioned on them.

For each variable in its turn, we find a set of parents: those attributes that are required in order to render the current variable independent of the rest. We use the notation $\Gamma^i = \{a_1, \ldots, a_i\}$, and (x, y) refers to a directed arc from $x \in V$ to $y \in V$.

algorithm $\underline{ProcGetStructure}(\Gamma)$
input: Γ, ordered as $\{a_1, \ldots, a_n\}$
output: a utility difference network over Γ
for i=1 to n:
find minimal $\hat{\Gamma}^i \subseteq \Gamma^{i-1}$ such that $\text{CDI}_r(a_i, \Gamma^{i-1} \setminus \hat{\Gamma}^i | \hat{\Gamma}^i)$
For each $x \in \hat{\Gamma}^i$, Add (x, a_i) to E
return $G = (\Gamma, E)$

The data in the nodes of a utility difference network is in form of conditional utility function, that is obtained by querying a user for preference differences. For example, the node a with parents γ requires the function $u_r(a|\gamma)$, which is obtained by queries for the differences $u_r(a\gamma) - u_r(a^0\gamma)$.

4.2 Example

In order to demonstrate the difference between CDI and CDI_r, we consider the hard-drive example used by Engel and Wellman (2007), and we show their CDI-map of the problem in Figure 1a. The example describes various decision criteria that a procurement department of a company evaluates when purchasing some quantity of new hard-drives. Below, each attribute is listed with a designated attribute name (the first letter), and its (sometimes arbitrary) domain.

RPM (**R**) 3600, 4200, 5400 RPM
Transfer rate (**T**) 3.4, 4.3, 5.7 MBS
Volume (**V**) 60, 80, 120, 160 GB
Supplier ranking (**S**) 1, 2, 3, 4, 5
Quality rating (**Q**) (of the HD brand) 1, 2, 3, 4, 5
Delivery time (**D**) 10, 15, 20, 25, 30, 35 days
Warranty (**W**) 1, 2, 3 years
Insurance (**I**) (for the case the deal is signed but not implemented) α, β, γ
Payment timeline (**P**) 10, 30, 90 days

Intuitively, the most salient criterion is volume, and the other important ones are RPM, warranty, and quality; hence the ordering prefix V, R, W, Q seems sensible. The rest of the ordering is less crucial, and we use I, D, P, T. First, V is placed as a root. When R is considered next, we find that it depends on V because for high volume hard-drives the marginal utility of improving RPM is higher. We find that W depends on V, because larger hard-drives tend to fail more. Now when considering Q, we might find that there is similar dependence between Q and V. However, if the reference value of W is the maximum value (3 years), we might find that given the reference value of W (and the rest of the variables), Q and V are independent because the longer warranty alleviates (substitutes) quality concerns. Further down in the network we might see a similar effect: a convenient value for insurance alleviates the dependency of delivery terms on the payment terms. In addition, we find the CDI conditions as described by Engel and Wellman (2007). For example, given any fixed value of V the marginal value of improving the quality rating does not depend on the RPM. Also, S is CDI and therefore CDI$_r$ of all the rest. We obtain the DAG depicted in Figure 1b. The utility difference network does not achieve lower dimension than the corresponding CDI-map, however it provides directionality that can be exploited for a more natural elicitation process.

5 Discussion

CDI is a stronger condition than CDI$_r$, and as such the locality property it achieves is stronger. To see this, consider the elicitation of data for an attribute x. In a CDI-map, the marginal utility over x is independent of the value of any node outside its local neighborhood. Therefore, the marginal utility of x can be elicited using local queries that involve differences over x and given a fixed value of the neighbors, without even knowing the value of the rest of the variables. In utility difference network the marginal utility of x may in some cases depend on the value of a non-neighbor y. For example, if y is an ancestor of x, the marginal value of x is independent of y given any value of x's parents, but only *given that the rest of the variables are fixed on the reference value*. It is possible that there is some instantiation of the rest of the variables, under which the independence is lost. Therefore, elicitation must specify explicitly that the the rest of the variables are fixed on the reference value.

In fact, this can be seen as an advantage of utility difference network. In CDI-map, such x and y will necessarily have an edge connecting them, whereas in a utility difference network such edge can be omitted. In that sense, utility difference network refines

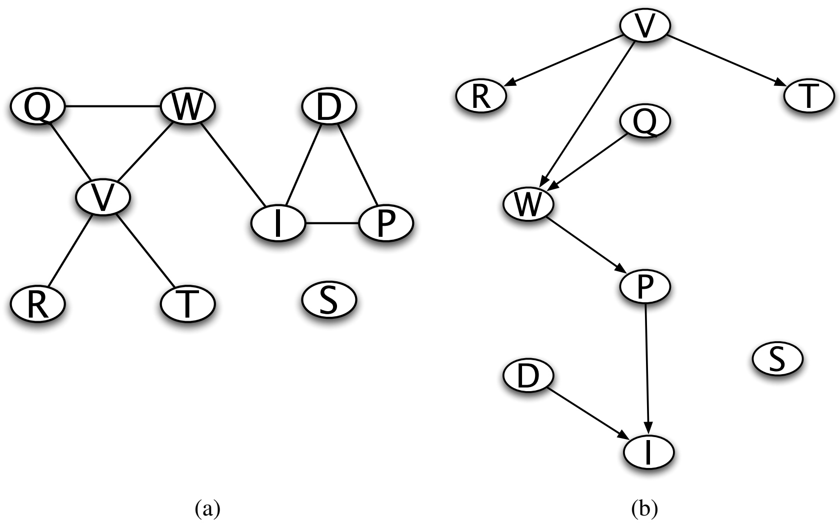

Fig. 1: Networks for the example: (a) CDI-map (b) utility difference network.

CDI. Furthermore, if x is fully CDI of y, this can be exploited in utility difference networks as well; as long as a local query makes sense to the user there is no need to indicate explicitly that the rest of the variables are fixed on the reference value.

A promising direction coming out of this representation is in introducing a form of Bayesian Learning for utilities. Consider a digital camera manufacturer that wishes to obtain information about customers' preferences. The company may be able to observe some limited set of choices made by customers. Perhaps we can also assume that single dimensional utilities (e.g., how much worths an improvement in a single attribute, all else being equal) is easy to estimate, or elicit. The company can now use the evidence (customer's choices) and Theorem 3, in order to obtain data regarding future choices of the customer. For example, the customer may have chosen to pay an extra $\$60$ for a camera with $10_{\times}$ zoom and 6_{mgp}, over one with $7_{\times}$ zoom and 6_{mgp}. Now (given the single dimensional data, in the form of reference utilities over each attribute) the company can compute the amount that the customer is willing to pay to get 6_{mgp} over 4_{mgp}, given that the zoom is $10_{\times}$.

Practical problem with this direction are yet to be resolved: this assumes that the outcomes above differ only in these two attributes, and in addition the rest of the attributes are fixed on the reference value (or, alternatively, difference independence holds between the two attributes we considered and any other attribute). Furthermore, we should theoretically be able to infer information about a customer only according to choices made by that customer. It is possible though that in some cases heuristic information can be inferred across different customers.

6 Conclusions

We propose a new representation scheme for utility functions. Starting from a definition of utility for a subspace of the domain, with respect to reference value of the rest of the attributes, we proceed with a definition of conditional utility as the marginal utility of an attribute, conditioned on some other attributes, and relative to the reference value. We show that conditional utility accommodates the logarithmic adaptations of the chain rule and Bayes rule, and develop the analogy to probabilities further by describing a directed graphical representation that relies on a concept of conditional independence.

In comparison with previous directed models (Boutilier et al., 2001; Engel and Wellman, 2008), we believe that our representation is simpler and easier to construct. Utility Difference Networks can be considered an adaptation of CDI-maps into a DAG, and though does not provide reduction of dimensionality, we believe that it has the potential to benefit the field in a similar way to how Bayesian Networks facilitated probabilistic reasoning in comparison to Markov Networks.

There are several direction to explore following this work. One is the form of Bayesian Learning proposed in Section 5. Furthermore, the fact that conditional utilities satisfy the chain rule and Bayes rule, implies that it may be possible to perform utility inference using algorithms similar to those that are used for belief propagation. Partial information obtained from observing the agent behavior, possibly coupled with observations about single dimensional preferences of similar users, can be used to infer other preferences. Value conditioning stemming from partial user choices or product constraints can be reasoned with, much like evidence in belief propagation, yielding estimates of the utility of various choices for other attributes.

Acknowledgements

Ronen Brafman is partially supported by ISF grant 8254320, by the Paul Ivanier Center for Robotics Research and Production Management, and by the Lynn and William Frankel Center for Computer Science. Yagil Engel is supported in part by an Aly Kaufman fellowship at the Technion.

Bibliography

Fahiem Bacchus and Adam Grove. Graphical models for preference and utility. In *Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 3–10, Montreal, 1995.

Craig Boutilier, Fahiem Bacchus, and Ronen I. Brafman. UCP-networks: A directed graphical representation of conditional utilities. In *Seventeenth Conference on Uncertainty in Artificial Intelligence*, pages 56–64, Seattle, 2001.

Darius Braziunas and Craig Boutilier. Local utility elicitation in GAI models. In *Twenty-first Conference on Uncertainty in Artificial Intelligence*, pages 42–49, Edinburgh, 2005.

James S. Dyer and Rakesh K. Sarin. Measurable multiattribute value functions. *Operations Research*, 27:810–822, 1979.

Yagil Engel and Michael P. Wellman. Generalized value decomposition and structured multiattribute auctions. In *Eighth ACM Conference on Electronic Commerce*, pages 227–236, 2007.

Yagil Engel and Michael P. Wellman. CUI networks: A graphical representation for conditional utility independence. 31:83–112, 2008.

Peter C. Fishburn. Interdependence and additivity in multivariate, unidimensional expected utility theory. *International Economic Review*, 8:335–342, 1967.

Christophe Gonzales and Patrice Perny. GAI networks for utility elicitation. In *Ninth International Conference on Principles of Knowledge Representation and Reasoning*, pages 224–234, Whistler, BC, Canada, 2004.

William M. Gorman. The structure of utility functions. *Review of Economic Studies*, 35:367–390, 1968.

Ralph L. Keeney. Utility independence and preferences for multiattributed consequences. *Operations Research*, 19(4):875–893, 1971.

Piero La Mura and Yoav Shoham. Expected utility networks. In *Fifteenth Conference on Uncertainty in Artificial Intelligence*, pages 366–373, Stockholm, 1999.

Yoav Shoham. A symmetric view of probabilities and utilities. In *Fifteenth International Joint Conference on Artificial Intelligence*, pages 1324–1329, Nagoya, Japan, 1997.

Joseph A. Tatman and Ross D. Shachter. Dynamic programming and influence diagrams. 20:365–379, 1990.