Asymptotic Maximum Entropy Principle for Utility Elicitation under High Uncertainty and Partial Information

Rafik Hadfi
Department of Computer Science
School of Techno-Business Administration
Nagoya Institute of Technology, Japan
rafik@itolab.nitech.ac.jp

Takayuki Ito
Department of Computer Science
School of Techno-Business Administration
Nagoya Institute of Technology, Japan
ito.takayuki@nitech.ac.jp

Abstract

Decision making has proposed multiple methods to help the decision maker in his analysis, by suggesting ways of formalization of the preferences as well as the assessment of the uncertainties. Although these techniques are established and proven to be mathematically sound, experience has shown that in certain situations we tend to avoid the formal approach by acting intuitively. Especially, when the decision involves a large number of attributes and outcomes, and where we need to use pragmatic and heuristic simplifications such as considering only the most important attributes and omitting the others. In this paper, we provide a model for decision making in situations subject to a large predictive uncertainty with a small learning sample. The high predictive uncertainty is concretized by a countably infinite number of prospects, making the preferences assessment more difficult. Our main result is an extension of the Maximum Entropy utility (MEU) principle into an asymptotic maximum entropy utility principle for preferences elicitation. This will allow us to overcome the limits of the existing MEU method to the extend that we focus on utility assessment when the set of the available discrete prospects is countably infinite. Furthermore, our proposed model can be used to analyze situations of high-cognitive load as well as to understand how humans handle these problems under *Ceteris Paribus* assumption.

Keywords: Decision Theory, Uncertainty, Maximum Entropy, Ceteris Paribus, Bounded Rationality

1 Introduction

Decision making involves generally two principal components. One dealing with the judgements about the uncertainties in the given world, whereas the other is related to the preferences over a set of possible consequences or outcomes. Several techniques have been proposed to help the decision maker in his analysis, by suggesting ways of formalization of the preferences as well as the assessment of the uncertainties. Although these techniques are established and proven to be mathematically sound, experience has shown that in certain situation we tend to avoid the formal approach by acting intuitively (R.L. Keeney 1994). Especially, when the decision involves a large number of attributes and outcomes. In this case, most of the decision makers tend to use pragmatic and heuristic simplifications such as considering only the most important attributes and omitting the others (Gigerenzer, Todd, and Group 1999).

In this paper, we provide a model for decision making in situations subject to a small learning sample and with a large predictive uncertainty with regards to the outcomes. We provide a formulation of this situation using an Information Theory approach and more precisely through the Maximum Entropy (ME) principle for preferences elicitation (Abbas 2006). The considered situation is characterized by a “High-cognitive Load” (Sterelny 2006; Chow 2011) or Bounded Rationality (Rubinstein 1997). In fact, we think that the bounded rationality or the cognitive limits of the mind (of the decision maker, or the agent) could be re-interpreted as the inability to grasp the large number of alternatives, attributes, outcomes and uncertainties. Thus, *the bounded-rationality is caused by the unbounded possibilities* that the decision maker is facing, which yields a huge amount of information to be considered. This amount of information is intractable and yet infinite for a simple optimization technique that seeks an optimal choice given the available information. Especially, when the decision has to be made in a finite amount of time despite the infinite number of possibilities. It is worth mentioning that this problem could be seen as an instance of the Frame Problem (Mccarthy and Hayes 1969), in the context of preferences elicitation. In such situations, the elicitation of a good utility function is not a realistic option and one should resort to other, less quantitative forms of preference representation. For instance, adopting a *Ceteris Paribus* preferential statements might be an option, as it was widely discussed in Philosophical Logic and Artificial Intelligence (Domshlak 2002). In fact, we ought to focus on what needs to be known (and represented) about a given environment and to omit what can be safely omitted. It is under such *Ceteris Paribus* assumption that we will propose our model to deal with the infinity of outcomes. Our main result could be seen as an extension of the ME principle whenever the set of prospects is countably infinite, and involving uncertainties. We think that using entropy methods enables us to capture the characteristics of such decision problems, as well as to the way solutions are realistically established by humans.

In the context of utility representation, several models were provided. For instance, (Chajewska and Koller 2000) proposed a model which takes into consideration the uncertainties over the utilities by considering a person's utility function as a random variable, with a density function over the possible outcomes. In the work of (Abbas 2006), probability and utility are considered with some analogy, thus yielding the notion of joint utility density function for multiple attributes. The application of Information Theory to Utility Theory gave a new interpretation of the notion of utility dependence, but most importantly, it allowed the elaboration of the MEU principle (Abbas 2006) as a way to assign utility values when only partial information is available. The same author assumed a continuous entropy measure on a continuous bounded domain of outcomes, which is true when the support of the distribution is finite. However, this continuity hypotheses does not hold when the support is countably infinite, which makes the information measure discontinuous in all probability distributions with countably infinite support (Ho and Yeung 2009). In other words, when the number of plausible outcomes that could be elicited by the ME utility function is countably infinite. In our approach, we adopted the same ME utility elicitation, but after establishing the behavior of the entropy with regard to the infinity of the outcomes.

The remainder of the paper is structured as following. In section 2, we provide the necessary concepts concerning the probability-utility formalism as well as the definition of the notion of utility convergence. In section 3, we describe the asymptotic case of the ME utility principle with its underlying continuity propositions. In the same section, we provide an example of realization of such situation, and we provide the solution to the asymptotic case. In section 4, we provide the conclusions and outline the future work.

2 Convergence of preferences

We start by providing our main framework, by adopting the utility-probability analogy and its usage for entropy maximization. We also provide a utility-based interpretation of the notion of convergence. Therefore, we start by stating it probabilistically, and then use utility increment vectors to define the utility convergence. Then, we define the continuity of the *Shannon* entropy with respect to a distance metric D_U. These are the first steps before treating in the next section the ME principal for countably infinite prospects.

2.1 Probability-Utility analogy

The analogy between probability distributions and utility was established in (Abbas 2006). Therefore, we use the same formalism as to name the utility vectors and the utility density functions. As in (Abbas 2006), we assume that the utility values are represented as a vector, namely a utility increment vector ΔU_i referring to a discrete utility function with one attribute i. As in (1), the vector ΔU_i contains the utility values $\{\Delta u_j\}_{j=0}^{k}$ of the $k+1$ ordered and discrete outcome $\{x^j\}_{j=0}^{k}$, starting from the lowest outcome x^0 to the highest outcome x^k, named x^*. We also define the sequence $\Delta U_{(n)}$ of n utility increment vectors as in (2).

$$\Delta U_i = (\Delta u_0, ..., \Delta u_j, ..., \Delta u_k), \tag{1}$$

$$\sum_{j=0}^{k} \Delta u_j = 1,\ \Delta u_j \geq 0\ \forall j \in [0,k]$$

$$\Delta U_{(n)} = \{\Delta U_i\}_{i=1}^{n} \tag{2}$$

A sequence of utility increment vectors can be compared to a sequence of random variables. It is built according to an analogy with a probability mass function $P = (p_1, ..., p_k)$. Each discrete utility increment vector ΔU_i corresponds to a normalized utility $U_i(x)$ function as in (3).

$$U_i(x) = \int_{x^0}^{x} u_i(x)\,\mathrm{d}x \tag{3}$$

$$u_i(x) = \frac{d}{dx} U_i(x) \tag{4}$$

That is, for a given prospect $x \in [x^0, x^*]$, $U_i(x)$ is determined by integrating the utility density function $u_i(x)$ (Abbas 2006) from the least preferred prospect x^0 up to the prospect x (the accumulated welfare from x^0 to x). The normalized utility function $U_i(x)$ has the same mathematical properties as a cumulative distribution function (CDF) as both are non-decreasing and range from zero to one (5).

$$0 \leq U_i(x) \leq 1,\ \frac{d}{dx} U_i(x) \geq 0\ \forall x \tag{5}$$

All along the paper, we will make usage of these notions of utility increment vector ΔU, the sequence of utility increment vectors $\Delta U_{(n)}$ as well as the utility density function $u(x)$.

2.2 Distance measure

The distance between two utility functions is defined based on the notion of similarity that could exist between them. By similarity, we mean the strategic equivalence (R.L. Keeney 1994), *i.e.*, two utility functions u_1 and u_2 are strategically equivalent, written $u_1 \sim u_2$, if and only if they imply the same preference ranking for any two lotteries. For instance, to compare two utility functions we can define the total variational distance that reflects the difference between the accumulated welfare all along the considered prospects. In the discrete case of two utility increment vectors, it is reduced to the $\mathcal{L}_1$-norm as it is shown in (6).

$$D_V(\Delta U_1, \Delta U_2) = \sum_{j} |\Delta U_{1,j} - \Delta U_{2,j}| \tag{6}$$

where D_V stand for the variational distance, and $\Delta U_{i,j}$ is the j^{th} element of ΔU_i. In case we are comparing two utility increment vectors ΔU_1 and ΔU_2 having respectively different dimensions L and M, (6) becomes:

$$D_V(\Delta U_1, \Delta U_2) = \sum_{j=1}^{L} |\Delta U_{1,j} - \Delta U_{2,j}| - \sum_{j=L+1}^{M} |\Delta U_{2,j}| \tag{7}$$

We can also use the divergence between two utility density functions (4) based on the *Kullback-Leibler* divergence given in (8), for both discrete and continuous cases.

$$D_U(\Delta U_1, \Delta U_2) = \sum_j \Delta U_{1,j}\ ln(\frac{\Delta U_{1,j}}{\Delta U_{2,j}}) \tag{8a}$$

$$D_u(u_1, u_2) = D_{KL}(u_1||u_2) = \int_x u_1(x)\ ln(\frac{u_1(x)}{u_2(x)}) \tag{8b}$$

where we adopt the convention $D_u(u_1, u_2) = 0$ if $u_2(x) = 0$ but $u_1(x) > 0$ for some x.
Moreover, based on the *Pinsker*'s inequality (Weissman et al. 2003) we have (9).

$$\frac{1}{2}\,[D_V(\Delta U_1, \Delta U_2)]^2 \leq D_U(\Delta U_1, \Delta U_2) \tag{9}$$

Both divergence (8) and the variational distance (6) can be used as measures of the difference between two utility increment vectors (respectively utility density functions) defined on the same set of prospects. However, once applied to utility functions, *Pinsker*'s inequality has the important implication that for two utility increment vectors ΔU_1 and ΔU_2 defined on the same set of prospects, if $D_U(\Delta U_1, \Delta U_2)$ (or $D_U(\Delta U_2, \Delta U_1)$) is small, then so is $D_V(\Delta U_1, \Delta U_2)$ (or $D_V(\Delta U_2, \Delta U_1)$). Furthermore, for a sequence of utility increment vectors $\Delta U_{(n)}$, as $n \to \infty$, if $D_U(\Delta U, \Delta U_n) \to 0$, then $D_V(\Delta U, \Delta U_n) \to 0$, *i.e.*, the convergence in divergence is a stronger notion of convergence than the convergence in variational distance. Thus, we will use the divergence measures (8) as to define the continuity of the *Shannon* entropy, in the sense that we study the convergence of a sequence of utility increment vectors as well as their entropies. Using (8) fits better with the idea of maximum likelihood estimation we might use in order to use the MEU.

2.3 Utility Convergence

From a probabilistic viewpoint, we say that a sequence P_n, (with a CDF F_n), is converging in distribution to a distribution P (with a CDF F), and noted as in (10).

$$P_n \xrightarrow{d} P \tag{10}$$

$$\text{with} \quad \lim_{n\to\infty} F_n(x) = F(x)\, \forall x \in \mathbb{R} \tag{11}$$

where F is continuous in x. As we mentioned in **2.1.**, a utility function can be represented by a CDF as in (3). If we assume that :

$$F(x) = \int_{x_{min}}^{x} f(x)dx \tag{12}$$

and if (3) is an analogy with (12), then what is the utilitarian significance of the sequence $F_n(x)$? To understand this setting, we rely on Merging Theory (Sorin 1999) (Lehrer 2000), which studies whether the beliefs of an agent, once updated after successive realizations of the process, converge to the true distribution. Now, we can interpret a sequence of preferences $\Delta U_{(n)}$ as a process over time. In fact, we can consider the sequential decision problem as in the case of a Bayesian agent observing the successive realizations of a discrete stochastic process $\{\Delta U_{(n)}|\ n \in T\}$ on the space of outcomes, indexed by n, and where n varies over a time index set T. The evolution of the process is announced round after round to the observer who observes a true distribution and holds an a priori preferences' belief on the process. We consider the preferences merging, namely, the convergence of this sequence of preferences $\Delta U_{(n)}$ to the limiting preferences ΔU. Firstly, let's assume that the decision maker is given the task of assessing a utility increment vector (13) for $k+1$ outcomes at time n.

$$\Delta U_{(n)} = (\Delta u_0{}^{(n)}, ..., \Delta u_k{}^{(n)}) \tag{13}$$

For instance, the assessment of the preferences $\Delta U_{(n+1)}$ is different from $\Delta U_{(n)}$, to the extent that new information have been made available to the decision maker, and used to update the preferences. A sequence $\Delta U_{(n)}$ converging to ΔU can be written as in (14).

$$\Delta U_{(n)} \xrightarrow{U} \Delta U \tag{14}$$

This notion of utility convergence reflects the idea that we expect to see the next outcome in a sequence of utility increment vectors $\Delta U_{(n)}$ to become better and better modeled by ΔU. The yielded convergence is expressed by the limit (15), for the discrete case.

$$\lim_{n\to\infty} \Delta U_{(n)} = \Delta U \tag{15}$$

This notion of convergence will be used to define the continuity of the *Shannon entropy* with regard to utility.

3 Asymptotic MEU

In this section, we provide the MEU principle, in the case of a countably infinite support. We start by providing the definition of the support of utility function (respectively a utility increment vector). Let a utility function u defined on the outcomes set D. The set of all the values that u could take is $\{u(x)|x \in D\}$.

Definition 1. *The support of u denoted by S_u, is the set of all the outcomes x in a set D, such that $u(x) > 0$.*

$$S_u = \{x|x \in D, u(x) > 0\} \tag{16}$$

If $S_u = D$, we say that u is strictly positive. Otherwise, u contain null utility values, which correspond to the undesirable, unwanted outcomes. The support of a discrete utility function will be used in the case of a large number of outcomes, literally approaching infinity. It is the case of utility functions that vanish for a certain number of outcomes ($u(x_j) = 0$), while being strictly positive for another infinite number of outcomes ($u(x_{i\neq j}) > 0$). In the discrete case we have (17).

Definition 2. *The support of ΔU is the set of all the indexes j such as $\Delta u_j \in \Delta U$ and $\Delta u_j > 0$.*

$$S_{\Delta U} = \{j \in \mathbb{N},\ \Delta u_j \in \Delta U \mid \Delta u_j > 0\} \tag{17}$$

3.1 Problem statement and method

Before stating the problem, let's give an example of situations, where the notion of infinity could arise while considering the outcomes.

Let's consider a multi-attribute utility function u over a set of attributes $a = \{a_1, ..., a_j, ..., a_n\}$, with $Domain(a_j) = D_j$. We want to define a preference ranking over the complete assignments on a.

Each complete assignment to a corresponds to a possible outcome of the decision makers action. Given the sizes of the attributes' domains, we can compute the number of possible assignments as $np = \prod_{j=1}^{n} card(D_j)$. Now, let's define another utility function u', over the domain $D = \times_{j=1}^{n} D_j$. It is obvious that the complexity of the assessment of u' grows up exponentially as the domains' sizes of the n attributes a_j are growing up.

Now, consider the case of complex systems subcontracting and manufacturing, for example, a *Boeing 747-400*, which is made in 33 countries and contains 6×10^6 parts (Holzer 2011). Let's consider that each part of the plane is an attribute, therefore, designing the plane is reduced to finding and instantiating the attributes by assigning values from their domains. In the end, the best design will be chosen amongst all the possible instantiations, in other words, the outcomes. While keeping in mind the goal of providing efficient and automated tools for preference elicitation, we can highlight the considerable effort and complexity that will arise if we want to build a utility function over such possible set of outcomes. Given 6×10^6 interdependent attributes with a maximal size domain d, even if $d = 2$ (which is unlikely since we are dealing with complex systems), the number of possible combinations is $np = 2^{6 \times 10^6}$, which is too large to be treated quantitatively ($np \sim \infty$).

It is under such assumptions of infinity that we will adopt a *Ceteris Paribus* preferential statement, notably statements in which "all else being equal", and by varying a number of variables (in our case, $[1, ..., L]$) while holding the others ($[L + 1, ..., M]$) constant (Domshlak 2002). In other words, we will reduce the actual frame of M infinite outcomes (ΔU) to L outcomes ($\Delta U'$), which could be reasonably assessed, under *Ceteris Paribus*.

In our case, we are assessing a utility increment vector ΔU_{meu} containing all the preferences as discrete elements (1). We also take the support $S_{\Delta U}$ as countably infinite, which makes the assessment process more difficult, as we mentioned in the previous example. The idea here is to use another utility increment vector $\Delta U'$ to contain the reasonably assessed preferences of the decision maker, according to his subjective belief. Then, we try to estimate ΔU with a new utility increment vector ΔU_{meu} with respect to the ME principle and by minimizing $D_U(\Delta U, \Delta U')$ (8).

Let $\Delta U = (\Delta u_1, ...\Delta u_M)$ be the true utility increment vector to be assessed, where M is a large number that tends to infinity. We can see M as the number np we provided in the example above. Let $\Delta U' = (\Delta u'_1, ...\Delta u'_L)$ be the utility increment vector that the decision maker was able to assess, due to the reduction of the number of outcomes to L ($L < M$), under *Ceteris Paribus*. We propose then to find the entropy maximization utility increment vector with respect to the minimal possible distance between ΔU and $\Delta U'$, that is, $D_U(\Delta U, \Delta U') \leq \epsilon$. One way to solve this maximization problem is to adopt the approach used in (Abbas 2006), by finding the continuous utility function u^* that maximizes the entropy (18).

$$u^*(x) = \underset{u(x)}{\operatorname{argmax}}\, H(u(x)) \tag{18}$$

Since $S_{\Delta U}$ is countably infinite, (18) cannot be solved with *Lagrange* multipliers and simple derivation methods. In fact $H(\Delta U)$ is discontinuous whenever $S_{\Delta U}$ is countably infinite, and most importantly when the continuity measure is based on the *KL-distance* (Ho and Yeung 2009).

In the next section, we introduce the *Shannon* entropy and define its continuity as well as its discontinuity whenever the considered utility increment vector has an infinite support.

3.2 Continuity of the entropy measures

The *Shannon* entropy measures are functions mapping a probability distribution to a real value. They can be described as the measure of uncertainty about a discrete random variable X having a probability mass function p.

Definition 3. *The entropy H(X) of a random variable X is:*

$$H(X) = -\sum_{x} p(x) \log p(x) \tag{19}$$

We adopt the convention that summation is over the support of the given probability distribution, in order to avoid undefined cases. An important characteristics of *Shannon* entropy, is that it measures the spread of a probability distribution and therefore, achieves its maximum value when the distribution assigns equal probabilities to all the outcomes. This concept was used as a method to assign prior probability distributions that maximize *Shannon* entropy measure under partial information constraints. It is possible to apply *Shannon* entropy measures to a utility increment vector reflecting the spread of the prospects (Abbas 2006) as in (20).

$$H(\Delta u_1, ..., \Delta u_n) = -\sum_{i=1}^{n} \Delta u_i \log(\Delta u_i) \tag{20}$$

In the case where the outcomes are finite, the *Shannon* entropy measures are continuous function. We propose to focus on the case where the entropy measure H is applied to utility increment vectors ΔU with countably infinite elements, situation reflecting the high uncertainty. More precisely we are interested in studying the continuity of H with respect to the distance measures we established in the section **2.2**. For instance, entropy is discontinuous with respect to the *Kullback-Leibler* divergence (Ho and Yeung 2009). We should highlight that the underlying utility functions follow the axioms of normative utility functions (von Neumann and Morgenstern 1947), which gives (21).

$$\int_{x} u(x) = 1 \tag{21}$$

We propose to define the continuity of a function f that will be lately extended into the entropy measure H.

Definition 4. *Let π_X be the set of all utility density functions on the set of outcomes X and let $u \in \pi_X$. $f : \pi_X \to [0, 1]$ is continuous at u if, given any $\epsilon > 0$, $\exists \delta > 0$ such that: $\forall u' \in \pi_X : D_u(u, u') < \delta \Longrightarrow |f(u') - f(u)| < \epsilon$.*

For the discrete case, we have the following definition.

Definition 5. *Let π_k be the set of all utility increment vectors defined for k outcomes, and let $\Delta U \in \pi_k$.*
$f : \pi_k \to [0, 1]$ is continuous at ΔU if, given any $\epsilon > 0$, $\exists \delta > 0$ such that: $\forall \Delta U' \in \pi_k : D_U(\Delta U, \Delta U') < \delta \implies |f(\Delta U') - f(\Delta U)| < \epsilon$.

If f fails to be continuous at the utility density function u (resp. ΔU), then we say that f is discontinuous at u (resp. ΔU). Given the notion of convergence we defined in section **2.3.**, we can provide the following definitions of the discontinuity of the function f.

Definition 6. *Let π_X be the set of all utility density functions on the set of prospects X and let $u \in \pi_X$. A function $f : \pi_X \to [0, 1]$ is discontinuous at u if there exists a sequence of utility density functions $u_{(n)} \in \pi_X$ such that :*

$$\lim_{n\to\infty} D_u(u_{(n)}, u) = 0 \tag{22}$$

but $f(u_{(n)})$ does not converge to $f(u)$, i.e.,

$$\lim_{n\to\infty} f(u_{(n)}) \neq f(u) \tag{23}$$

Similarity, for the discrete case we have:

Definition 7. *Let π_k be the set of all utility increment vectors defined for k prospects. Let $\Delta U \in \pi_k$.*
A function $f : \pi_k \to [0, 1]$ is discontinuous at ΔU if there exists a sequence of utility increment vectors $\Delta U_{(n)} \in \pi_k$ such that :

$$\lim_{n\to\infty} D_U(\Delta U_{(n)}, \Delta U) = 0 \tag{24}$$

but $f(\Delta U_{(n)})$ does not converge to $f(\Delta U)$, i.e.,

$$\lim_{n\to\infty} f(\Delta U_{(n)}) \neq f(\Delta U) \tag{25}$$

3.3 Discontinuity

In this section, we establish the discontinuity of H at any utility increment vector ΔU having a countably infinite support. Let $\Delta U_{(n)}^{(a,b)}$ be a sequence of utility increment vectors with the real parameters a and b. We will use this sequence to show that H is discontinuous at $\Delta U_1 = (1, 0, 0, ...)$.
For fixed real numbers a and b and an integer n, such as $a > 1$ and $b > 0$ and $n > a$. We define $\Delta U_{(n)}^{(a,b)}$ as in (26).

$$\Delta U_{(n)}^{(a,b)} = \{1 - (\frac{\log a}{\log n})^b, \frac{1}{n}(\frac{\log a}{\log n})^b, ..., \frac{1}{n}(\frac{\log a}{\log n})^b, 0, 0, ...\} \tag{26}$$

Based on our definition of convergence in **2.3.**, we show that the sequence $\{\Delta U_{(n)}^{(a,b)}\}$ converges to $\Delta U_1 = (1, 0, 0, ...)$. Computing the distance between the vector ΔU_1 and the parametrized vector $\Delta U_{(n)}^{(a,b)}$ gives (27).

$$D_U(\Delta U_1, \Delta U_{(n)}^{(a,b)}) = -\left(\log(1 - \left(\frac{\log a}{\log n}\right)^b\right) \tag{27}$$

We have $\Delta U_{(n)}^{(a,b)} \xrightarrow{U} \Delta U$, which is given in (28).

$$\lim_{n\to\infty} D_U(\Delta U_{(n)}^{(a,b)}, \Delta U_1) = 0 \tag{28}$$

Then, the entropy of $\Delta U_{(n)}^{(a,b)}$ is given by (29).

$$\begin{aligned} H(\Delta U_{(n)}^{(a,b)}) = &-\left[1 - \left(\frac{\log a}{\log n}\right)^b\right] \log \left[1 - \left(\frac{\log a}{\log n}\right)^b\right] \\ &-n\left[\frac{1}{n}\left(\frac{\log a}{\log n}\right)^b\right] \log \left[\frac{1}{n}\left(\frac{\log a}{\log n}\right)^b\right] \end{aligned} \tag{29}$$

Hence, we have (30). For a complete proof we refer the reader to (Ho and Yeung 2009).

$$\lim_{n\to\infty} H(\Delta U_{(n)}^{(a,b)}) = \begin{cases} 0 & \text{if } b > 1 \\ \log a & \text{if } b = 1 \\ \infty & \text{if } 0 < b < 1 \end{cases} \tag{30}$$

Proposition 1. *Based on Definition 7. and if we take $f = H$, $a > 1$ and $0 < b \leq 1$ in (30), we have:*

$$\lim_{n\to\infty} H(\Delta U_{(n)}^{(a,b)}) = \infty \neq H(\Delta U_1) \tag{31}$$

$$\textit{but} \qquad \lim_{n\to\infty} \Delta U_{(n)}^{(a,b)} = \Delta U_1 \tag{32}$$

Therefore, we can state that H is discontinuous at the utility increment vector $\Delta U_1 = (1, 0, 0, ...)$.

3.4 Bound and majorization

Given (30), we propose to find a bound to η in (33).

$$\eta = |H(\Delta U) - H(\Delta U')| \tag{33}$$

with ΔU and the $\Delta U'$ the utility increment vectors we provided at the beginning of section **3.**. In fact, if the dimension M of ΔU is finite and known, (33) is also finite and we propose to find its upper bound (34).

$$\sup_{\Delta U} |H(\Delta U) - H(\Delta U')| \tag{34}$$

Since the utility increment vector $\Delta U'$ is available to the decision maker (assessed under *Ceteris Paribus* as we mentioned above), we will start by solving (35).

$$\begin{aligned} &\sup_{\Delta U} |H(\Delta U')| \\ &\text{subject to} \qquad D_U(\Delta U, \Delta U') \leq \epsilon \end{aligned} \tag{35}$$

With a finite value of M, (35) is reduced to finding (36).

$$\max_{\substack{\Delta U \\ D_{\Delta U}(\Delta U, \Delta U') \leq \epsilon}} \quad H(\Delta U) \tag{36}$$

Now we can think about the majorization of (36) and thus providing the solution ΔU_{meu}.
Let $\gamma = \sum_{i=1}^{L} \Delta u_i$. We can write ΔU as in (37).

$$\Delta U = (\Delta u_1, ..., \Delta u_M) \tag{37a}$$

$$\Delta U' = (\frac{\Delta u_1}{\gamma}, ..., \frac{\Delta u_L}{\gamma}) \tag{37b}$$

$$\Delta U'' = (\frac{\Delta u_{L+1}}{(1-\gamma)}, ..., \frac{\Delta u_M}{(1-\gamma)}) \tag{37c}$$

Therefore, $H(Q)$ can be written as in (38).

$$H(\Delta U) = h(\gamma) + \gamma * H(\Delta U') + (1-\gamma) * H(\Delta U'') \tag{38}$$

with h the binary entropy function (39).

$$h(x) = -x \log(x) - (1-x)\log(1-x) \tag{39}$$

Since $\#(\Delta U'') = (M-L)$, we can define an upper bound for $H(\Delta U'')$ (Yeung 2008), as in (40).

$$H(\Delta U'') \leq log(M-L) \tag{40}$$

(38) and (40) give (41)

$$\max_{\substack{\Delta U \\ D_{\Delta U}(\Delta U, \Delta U') \leq \epsilon}} H(\Delta U) \leq \tag{41}$$
$$h(\gamma) + \gamma * H(\Delta U') + (1-\gamma) * \log(M-L)$$

Therefore, the maximum value that could be reached by the entropy H is shown in (42).

$$H(\Delta U^*) = h(\gamma) + \gamma * H(\Delta U') + (1-\gamma) * \log(M-L) \tag{42}$$

The utility increment vector that achieves this maximum is ΔU_{meu} (43).

$$\Delta U_{meu} = \left(\Delta u_0, ..., \Delta u_L, \frac{(1-\gamma)}{(M-L)}, ..., \frac{(1-\gamma)}{(M-L)}\right) \tag{43}$$

The solution (43) is the optimal utility increment vector that ensures the ME with respect to the given information (37b). The specified preferences $\Delta U'$ could be characterized by a utility density function, while the rest of the vector ΔU_{meu} will be a uniformly distributed utility increment vector.

4 Conclusion

We consider the ME principle for utility elicitation in the case of high uncertainty, that is, when the decision maker is facing a large number of outcomes, while having a limited knowledge. Despite its practical importance, as to understand how decisions are made in bounded rationality, this problem has not been studied from the perspective of entropy maximization. To address this problem, we assumed that this situation of high uncertainty could be translated into a countably infinite number of outcomes. The decision maker is asked to provide a utility function that maximizes the entropy, given the available information. Solving this type of problems could rely on *Lagrange* multipliers. But, like we have shown, this method, and general derivation-based methods could not be used due to the discontinuity of the entropy measures whenever the support is infinite. Therefore, we proposed another method based on finding a limiting least upper bound of the entropy, and thus giving a utility increment vector that maximizes it.

As an important research issue to be further investigated, we think about considering the case of multi-attribute utility increment vectors, and therefore generalize the univariate asymptotic case to the multivariate case.

References

Abbas, A. E. 2006. Maximum entropy utility. *Operations Research* 54(2):277–290.

Chajewska, U., and Koller, D. 2000. Utilities as random variables: Density estimation and structure discovery. In *In Proceedings of the Sixteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-00)*, 63 – 71.

Chow, S. J. 2011. Heuristics, concepts, and cognitive architecture: Toward understanding how the mind works. *Electronic Thesis and Dissertation Repository. Paper 128.*

Domshlak, C. 2002. Modeling and reasoning about preferences with cp-nets.

Gigerenzer, G.; Todd, P.; and Group, A. R. 1999. *Simple heuristics that make us smart*. Evolution and cognition. Oxford University Press.

Ho, S.-W., and Yeung, R. W. 2009. On the discontinuity of the shannon information measures. *IEEE Trans. Inf. Theor.* 55:5362–5374.

Holzer, M. 2011. Aircraft fasteners: How to select the correct type of fastener. *Aircraft Maintenance Technology*.

Lehrer, E. 2000. Relative entropy in sequential decision problems. *Journal of Mathematical Economics* 33(4):425–439.

Mccarthy, J., and Hayes, P. J. 1969. Some Philosophical Problems from the Standpoint of Artificial Intelligence. In *Machine Intelligence*, volume 4, 463–502.

R.L. Keeney, H. R. 1994. Decisions with multiple objectives preferences and value tradeoffs. *Behavioral Science* 39(2):169–170.

Rubinstein, A. 1997. *Modeling Bounded Rationality*. MIT Press.

Sorin, S. 1999. Merging, reputation, and repeated games with incomplete information. *Games and Economic Behavior* 29(1-2):274–308.

Sterelny, K. 2006. Cognitive load and human decision, or, three ways of rolling the rock up hill. In Carruthers, P.; Laurence, S.; and Stich, S., eds., *The Innate Mind: Culture and Cognition*. Cambridge: Cambridge University Press.

von Neumann, J., and Morgenstern, O. 1947. Theory of games and economic behavior, 2nd ed. princeton university press, princeton, nj.

Weissman, T.; Ordentlich, E.; Seroussi, G.; Verdu, S.; and Weinberger, M. J. 2003. Inequalities for the l 1 deviation of the empirical distribution. *Group* (HPL-2003-97R1):1 – 10.

Yeung, R. 2008. *Information theory and network coding*. Information technology–transmission, processing, and storage. Springer.