ISAHP 1999, Kobe, Japan, August 12-14, 1999

On MAUT, AHP, and PFM

Jonathan Barzilai
Dept. of Industrial Engineering, Dalhousie University

1 Introduction

Mainstream decision theory is based on the misconception that von Neumann and Morgenstern's utility theory is an appropriate (or the only) paradigm for *applied* decision making (as opposed to a *theoretical* framework). There is more than sufficient evidence to indicate that as a framework for *applied* decision making, utility theory has been a failure. Perhaps the clearest indication of this failure is the success of a competing decision methodology, Saaty's Analytic Hierarchy Process (AHP, [17]). The AHP is being applied in practice because it offers an easy-of-use alternative to utility theory, that is, not because it is founded on proper theoretical foundations but despite its theoretical shortcomings.

In section 2 we highlight some fundamental misconceptions related to utility and measurement theory, while section 3 of the paper lists some of the reasons why the AHP is not a valid methodology. A deeper understanding of the theory of measurement has provided the foundations for constructing a new methodology, Preference Function Modelling (PFM, see [5 and 3]). PFM is based on sound theoretical foundations and is easy to use. Practitioners do not need to become experts in measurement or decision theory to use a software package which implements PFM (see [9]).

2 Utility Theory and Related Issues

Modern decision theory is founded on von Neumann and Morgenstern's (single-criterion) utility axioms (see [22]). This is not the only possible *theoretical* framework for utility (that is, preference) theory and it is not a satisfactory framework for the *practical* measurement of preference. Surprisingly, over more than fifty years, many generations of researchers have convinced themselves of the misconception that the set of axioms underlying a preference theory determines what is being measured. The following basic principles clarify this point (see Barzilai [3] for details).

Basic Principle 1: Measurement procedures are not unique -- a property of a set of objects may be measured in multiple ways using various tools and procedures.

Basic Principle 2: *Which* (valid) procedure is used to measure a property (i.e., *how* it is measured) does not determine *what* is being measured.

Basic Principle 3: Measurement can be described by various sets of axioms.

Basic Principle 4: The measurement of a property of a set of objects which is described by a given set of axioms, may be carried out using *any* valid procedure for measuring that property; such procedures do not have to be a translation of the set of axioms into an operational recipe.

It follows that the distinction between utility and value functions is based on a misconception and that utility, value, and preference functions are different names for the same functions representing the underlying property of preference. It also follows that there is no reason for employing the unnecessarily complicated paradigm of gambling with its non-intuitive notions of lotteries, subjective probability and the so-called "basic reference lottery ticket" when simpler, intuitive measurement rules are available. It may be surprising to note that already in 1959, Stevens [19, p. 49] asked with regard to utility: "if the subject is really being asked about the equality of intervals on his scale of subjective value, why do we not ask him to judge the intervals directly? Why introduce the complicating factor of risk and probabil-

This is a paper for **Tutorial, "Preference Function Modeling."** (Program Committee) ity?" Indeed, PFM measures preference differences directly without the artificial introduction of risk and probability into decisions under certainty.

In addition, Stevens [19, p. 60] notes that utility elicitation questions are difficult to answer and concludes: "Since the answer is difficult, it is probably wrong." At the very least, we can say that although PFM and the "standard gamble" method of utility may be theoretically equivalent, probability in general, and "subjective probability" in particular, are not intuitive concepts. Therefore, utility elicitation via hypothetical lotteries introduces extraneous bias and errors into the measurement process.

More recently, utility theory has been intertwined into measurement theory which seems to have developed its own peculiarities. As an example, consider the following. Stevens [20, p. 37] states "... the magnitude of a person's sensation may be discovered by a systematic study of what he does in a controlled experiment in which he operates on other systems. He may, for instance, adjust the loudness in his ears to match the apparent intensity of various concentrations of salt solution applied to his tongue." If we accept the logic behind this experiment, we can design another one in which the subject adjusts the weight of his wallet to match the length of his pen. Thus "cross modality matching" accepts statements of the type "the length of my pen equals the weight of my wallet." More peculiar than "cross modality matching" itself is its mathematical modelling by Krantz *et al.* [10, §4.6] and Roberts [16, §4.4].

The current state of measurement theory is not very satisfactory. Basic concepts such as scale type and meaningfulness are not fully understood – see for example Luce *et al.* [11, §22, §20 and in particular §20.1.2]; the peculiar discussion of equation (2.25) in Roberts [16, p. 74] and the "several versions of meaningfulness" which follow [16, p. 74]; the comment – "something is bizarrely wrong" in Krantz *et al.* [10, footnote, p. 455] and the concluding remarks in Krantz *et al.* [10, §10.15, pp. 535–536]. These and related issues will be resolved in a forthcoming paper.

Moreover, despite the great amount of technical discussion such as can be found in the *Foundations of Measurement* by Krantz, Luce, Suppes and Tversky published in three volumes and spanning close to twenty years, nowhere in these volumes will the reader find the basic model of measurement of length of two objects, say the sides of an isosceles right triangle. This and related types of measurement which form the basic building blocks of PFM are modelled in Barzilai [1].

3 The AHP

Hierarchical decomposition of multi-attribute preference functions was first proposed by Miller in 1966 (see [12] as well as [13–14]) as part of what may be properly named Miller's Hierarchy Process (MHP). Miller's concept of the *decomposition of the criteria* into a sub-criteria tree (i.e., the generation of operational sub-criteria), has been incorporated into a large number of methodologies including Multi-Attribute Utility Theory, which was proposed by Raiffa in 1969 (cf. [15, p. 17]). Unfortunately, the *computational* aspects of MHP, i.e., the assignment of numbers to the elements of the tree, are fundamentally flawed and any methodology which employs these components is flawed as well. The most well known variant of MHP which employs its computational procedure is Saaty's Analytic Hierarchy Process (AHP – see [17]).

The AHP claims to be different from utility (see the title of [21]); to measure priorities as opposed to utilities; and to have its own logic (see the title of [18]). Previous attempts to analyze these and other AHP claims (see e.g. Belton and Gear [7] and Dyer [8]) have mis-identified the problem, proposed incorrect revisions, and failed to address the fundamental issues (see Barzilai [4] for details). The AHP is not a valid methodology for various reasons including the following.

The AHP attaches numbers to alternatives in accordance with their "priorities." It follows that measurement theory (rather than the "logic of priorities") applies to the AHP. It also follows from the discussion in section 2 that priorities, utilities and preferences are different labels for the same property and that the AHP is a utility theory (not a valid one, but still, a utility theory).

In Barzilai [2], we established that the AHP consists of a one-dimensional (single-criterion) measurement procedure and a procedure for combining single-criteria measurements into an overall preference function $f(x_1, \ldots, x_n)$ which, we proved, is always linear in the measured single-criteria variables: $f(x_1, \ldots, x_n) = \sum_{j=1}^{n} w_j x_j$.

The AHP *single-criterion* measurement procedure is flawed for the following reasons. The scale type used for "priority" measurement is incorrect. As a result, the Decision Maker is asked meaningless questions and provides meaningless responses. Pairwise comparisons are not applicable for measurement of preferences – the minimum number of alternatives to be compared is three. The eigenvector method for reconciling inconsistent input is not valid for a variety of reasons – the AHP's claims concerning the eigenvector method are incorrect, circular or meaningless (see Barzilai [6]).

The AHP's *hierarchical decomposition* rule for the construction of an overall preference function from single-criteria measurements superimposes the eigenvector procedure on Miller's rule. Consequently, it shares with MHP all the flaws of the hierarchical decomposition rule and is not valid (see [4 and 2] for details). In particular:

- The ratios of criteria weights are exchange (substitution) rates. They depend on the units in which criteria variables are measured and cannot be determined before criteria variables are measured or independently of these variables.
- It follows that the top-down computation of criteria weights on the criteria tree is not valid.
- It also follows that the ratios of criteria weights cannot be interpreted as "relative importance" and that deriving these weights from questions concerning "relative importance" results in meaningless responses.
- The number of normalizations used to compute criteria weights is incorrect.
- The data collected in the process are insufficient to construct the model $\sum_{j=1}^{n} w_j x_j$.
- The AHP generates non-equivalent preference functions and rankings from equivalent decompositions.

Finally, the "axioms" underlying the AHP are meaningless as well. If they do not properly characterize the AHP, they are of no interest. On the other hand, if they do, they cannot be meaningful either, since they characterize a methodology which suffers from multiple fatal flaws.

4 Preference Function Modelling

Preference Function Modelling (PFM) is a new methodology for preference measurement (see Barzilai [5]). It is based on new results and deeper understanding of the theory of measurement (see Barzilai [3 and 1]).

In particular, we have constructed a homomorphic modelling framework for measurement including a new characterization of uniqueness (scale type) via a system of relational equations (Barzilai [3, Equation (9)]). We also outlined basic principles of measurement in Barzilai [3] and constructed the assignment rules as well as sets of axioms for measurement on affine scales in Barzilai [1]. We proved the very strong result that these sets of axioms are fundamental in the sense that they are implied by any other set of axioms for such measurement.

PFM provides conditions and procedures for measurement on affine scales and allows the user to provide partial and multiple inconsistent estimates which are reconciled through properly justified procedures. The methodology is based on firm theoretical foundations, yet is easy to use.

Notes. PFM is a trademark of Scientific Decisions Inc. (www.ScientificDecisions.com). The author can be reached by e-mail at Jonathan.Barzilai@dal.ca

References

[1] J. Barzilai, "Decision Theory Misconceptions and Remedies," *INFOR*, to appear.

[2] J. Barzilai, "On the Decomposition of Value Functions," *Operations Research Letters*, Vol. 22, Nos. 4–5, pp. 159–170, 1998.

[3] J. Barzilai, "Measurement Foundations for Preference Function Modelling," *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 4038–4044, 1998.

[4] J. Barzilai, "Understanding Hierarchical Processes," *Proceedings of the 19th Annual Meeting of the American Society for Engineering Management*, pp. 1–6, 1998.

[5] J. Barzilai, "A New Methodology for Dealing with Conflicting Engineering Design Criteria," *Proceedings of the 18th Annual Meeting of the American Society for Engineering Management*, pp. 73–79, 1997.

[6] J. Barzilai, "Deriving Weights from Pairwise Comparison Matrices," *Journal of the Operational Society*, Vol. 48, No. 12, pp. 1226–1232, 1997.

[7] V. Belton and A.E. Gear, "On a Shortcoming of Saaty's Method of Analytic Hierarchies," *Omega*, Vol. 11, No. 3, 228—230, 1983.

[8] J.S. Dyer, "Remarks on the Analytic Hierarchy Process," *Management Science*, Vol. 36, No. 3, 249—258, 1990.

[9] E.W. Grundke and J. Barzilai, *SDI Web Site*, www.ScientificDecisions.com, Scientific Decisions Inc.

[10] D.H. Krantz, R.D. Luce, P. Suppes and A. Tversky, *Foundations of Measurement*, Vol. 1, Academic Press, 1971.

[11] R.D. Luce, D.H. Krantz, P. Suppes and A. Tversky, *Foundations of Measurement*, Vol. 3, Academic Press, 1990.

[12] J.R. Miller, *The Assessment of Worth: A Systematic Procedure and Its Experimental Validation*, doctoral dissertation, M.I.T., 1966.

[13] J.R. Miller, *Assessing Alternate Transportation Systems*, Memorandum RM-5865-DOT, The RAND Corporation, 1969.

[14] J.R. Miller, *Professional Decision-Making*, Praeger, 1970.

[15] H. Raiffa, *Preferences for Multi-Attributed Alternatives*, Memorandum RM-5868-DOT/RC, The RAND Corporation, 1969.

[16] F.S. Roberts, *Measurement Theory*, Addison-Wesley, 1979.

[17] T.L. Saaty, *The Analytic Hierarchy Process*, McGraw-Hill, 1980.

[18] T.L. Saaty and L.G. Vargas, *The Logic of Priorities*, Kluwer-Nijhoff, 1982.

[19] S.S. Stevens, "Measurement, Psychophysics and Utility," in C.W. Churchman and P. Ratoosh (Eds.), *Measurement: Definitions and Theories*, Wiley, pp. 18–63, 1959.

[20] S.S. Stevens, *Psychophysics*, Wiley, 1975.

[21] L.G. Vargas, "Why the AHP is not Like MultiAttribute Utility Theory," in *Multiple Criteria Decision Support*, pp. 53—60, P. Korhonen, A. Lewandowski and J. Wallenius (Eds.), Springer-Verlag, 1991.

[22] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, 1944.