Expected Utility and Case-Based Reasoning*

Akihiko Matsui
University of Tokyo, University of Tsukuba†

Preliminary Draft: June, 1993
This Draft: December, 1998

Abstract

This paper presents an equivalence result between expected utility theory (EU) and a modified version of case-based decision theory (CBDT). To be precise, it shows that a model constructed in EU can be embedded in a CBDT model, and vice versa. CBDT, proposed and axiomatized by Gilboa and Schmeidler (1995), is related to case-based reasoning in psychology and artificial intelligence and considered as a descriptive theory of human behavior. In CBDT a decision maker remembers situations similar to the current problem and uses them to help solve it. This idea stems from bounded rationality and is similar in spirit to the satisficing theory of March and Simon (1958) in the sense that the decision maker tends to satisfice rather than optimize.

*The author is indebted to Itzhak Gilboa for many insightful comments and discussions. Discussions at the conference on decision making under uncertainty held at Johns Hopkins University were stimulating and helpful. Thanks are also given to seminar participants at Keio University, University of Pennsylvania and University of Pittsburgh as well as to Francis Diebold, Ed Green, Kim Katz, Hitoshi Matsushima, Stephen Morris, David Schmeidler, Barry Sopher, anonymous referees, and especially, Mamoru Kaneko and associate editor of this journal for useful discussions, comments, and references. The author is responsible for any error.

†Faculty of Economics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan; Institute of Policy and Planning Sciences, University of Tsukuba, Ibaraki 305, Japan (amatsui@e.u-tokyo.ac.jp, amatsui@shako.sk.tsukuba.ac.jp).

1 Introduction

Expected utility theory (henceforth, EU) has been a dominant tool in the analysis of human behavior under uncertainty in economics and game theory (von Neumann and Morgenstern(1947) and Savage (1954) for its axiomatization). While this theory is a powerful analytical tool, some observations have been made to show that EU sometimes fails to explain human behavior in the real world. Some pieces of evidence suggest two tendencies of human behavior. The first is that people tend to look for a better alternative as opposed to the best one. The second is that people often rely on past cases in making decisions. These sets of evidence led to models of bounded rationality. Among them, the best known among economists is the satisficing theory of March and Simon (1958). This theory uses a conceptually different framework. For example, in the satisficing theory, a decision maker is "satisficing" instead of optimizing as in EU.

Recently, Gilboa and Schmeidler (1995) propose *case-based decision theory* (CBDT), which is similar in spirit to the satisficing theory. According to CBDT, decisions under uncertainty are made by analogies to previously-encountered problems. The theory postulates a similarity function over decision problems. An action is evaluated by a similarity-weighted sum of the utility the action yielded in the past cases. CBDT is similar to the satisficing theory of March and Simon in the sense that a decision maker in CBDT tends to be satisfied with the present choice once an "aspiration level" is attained.

The relationship between EU and CBDT has been unclear, which leads to speculations and conjectures about their relationships and the validity of one theory in light of the other. Indeed, it seems, at a first glance, that EU and CBDT are not reconcilable. The only reconciliation takes effect when the identical situation is repeated sufficiently many times in an independent manner. In this case, the frequency of the outcomes converges to the true distribution. Using this fact, Gilboa and Schmeidler (1996) show that an "ambitious" agent, whose aspiration level is adjusted over time, tends to optimize on condition that the same problem is repeated sufficiently many times.[1]

The present paper modifies CBDT to show that a model of decision making based on EU, or an EU model, can be embedded into a model based on CBDT, or a CBDT model, and vice versa. Our embedding does not rely on any asymptotic argument. Any pair of models constructed this way prescribe the same actions at all decision nodes except the first one.[2]

Roughly speaking, the equivalence result takes advantage of the linear structures of the two theories. In the embedding of EU models into CBDT models (resp. CBDT models into EU models), a similarity value (resp. a conditional probability) is given by a linear combination of conditional probabilities (resp. similarity values). It turns out that there is a loose correspondence between conditional probabilities and similarity values: the higher the correlation between two cases in terms of the induced payoffs, the more similar the two cases are, and vice versa.

In addition to the translation between conditional belief systems and similarity functions, embedding involves the construction of other ingredients. Only two of them are mentioned here. First, in case of the embedding of CBDT into EU, constructing a state space is required. The present paper defines a state as a mapping from the set of all possible combinations of problems, actions and histories into the set of results. That is, if a state is known, one can tell what happens if a certain

[1] Gilboa and Schmeidler (1995) note that EU models can be embedded in CBDT models if a decision maker includes hypothetical cases in his memory.

[2] One cannot control the action in the first period in a CBDT model since there is no experience at all. However, this does not pose a serious problem since, in the real problem, one can always assume that there are some previous experiences.

action is taken given the past history and the present problem. The state space is the set of all such mappings.

Second, in case of the embedding of EU into CBDT, the set of problems should be specified. The present paper identifies a problem with a personal history. That is, all the observations the decision maker has ever made constitute a problem for him. In particular, if the past experiences are different, so are the problems.

The rest of the paper is organized as follows. Section 2 describes EU and CBDT. Section 3 embeds CBDT models into EU models and EU models into CBDT models. Section 4 concludes the paper.

2 Two Theories of Sequential Decision Making Problems

The present work considers a single decision maker who faces a sequence of decision nodes. Time is discrete, and the decision maker is faced with decision node in each period and takes an action in a finite set A. In the end of each period, he is informed of the result of that period.[3] In this environment, the decision maker uses either an EU model or a CBDT model to solve the problem.

The axiomatization of the version of CBDT used in the present paper is given in Appendix. The following is a description of the two theories used for the equivalence result.

[3] Note that this model is very general. What is essentially required is that there are at most countably many decision nodes. Though it is assumed that DM's decision node and information arrive alternately, one can always introduce dummy information to make a certain situation consistent with the above description.

2.1 Expected Utility Theory

This subsection considers a situation in which the decision maker analyzes the past experiences based on EU. A model based on EU, or an EU model, is given by:

$$< (\Omega, \mathcal{F}), A, R, f, u, \mu >,$$

where Ω is a state space, $\mathcal{F}$ is a σ-algebra on Ω, A is the set of available actions introduced above, R is a countable set of possible results, or equivalently, pieces of information, $u : R \to \mathbb{R} \equiv (-\infty, \infty)$ is a utility function, and μ is a probability measure on $(\Omega, \mathcal{F})$. The function $f : \cup_{T=1}^{\infty} A^T \times \Omega \to \mathbb{R}$ is an outcome function where A^T is the cross product of A. Write $f_T(\vec{a}, \omega) = f(\vec{a}, \omega)$ if $\vec{a} = (a^1, \cdots, a^T)$ $(T = 1, 2, \cdots)$. A value $f_T(\vec{a}, \omega) \in R$ is the result in the Tth period at $\omega \in \Omega$ if the decision maker has taken $\vec{a} = (a^1, \cdots, a^T)$. Let $\bar{h}_T = ((a^1, r^1), \cdots, (a^{T-1}, r^{T-1}))$ be a history at period $T = 2, 3, \cdots$. The null history is denoted $\bar{h}_1$. Denote by $\bar{h}(\omega, \vec{a})$ the history induced by ω and a sequence of actions $\vec{a} = (a^1, \cdots, a^{T-1})$. Let $\bar{H}_T$ denote the set of histories at time T. In particular, $\bar{H}_1 = \{\bar{h}_1\}$. Write $\bar{H} = \cup_{T=1}^{\infty} \bar{H}_T$. The conditional measure given $\bar{h} \in \bar{H}$, denoted by $\mu_{\bar{h}}$, is a measure on $(\Omega, \mathcal{F})$ and satisfies

$$\mu_{\bar{h}}(E) = \frac{\mu(E \cap \Omega_{\bar{h}})}{\mu(\Omega_{\bar{h}})}, \forall E \in \mathcal{F},$$

if $\mu(\Omega_{\bar{h}}) > 0$, where $\Omega_{\bar{h}} = \{\omega \in \Omega : \bar{h}(\omega, \vec{a}) = \bar{h}\}$.

In each period, the decision maker chooses an action to maximize the expected utility in that period. His expected utility if he takes $a \in A$ conditional on history $\bar{h} = ((a^1, r^1), \cdots, (a^{T-1}, r^{T-1})) \in \bar{H}$ is

$$V(a|\bar{h}) = \sum_{r \in R} \mu_{\bar{h}}(f_{\bar{h}}(a, \cdot) = r) u(r), \tag{1}$$

where $f_{\bar{h}}(a, \cdot) = f_T(\vec{a}, \cdot)$ with $\vec{a} = (a^1, \cdots, a^{T-1}, a)$. The decision maker takes $a \in A$ that maximizes (1). Note that the present specification is general enough to include the case where the future is affected by the current action, and he cares about the future payoff as well in making a decision today. In such a case, $u(r)$ is interpreted as the expected present value of the (future) payoff as opposed to an instantaneous payoff.[4] Throughout the rest of the paper, it is assumed that $V(a|\bar{h})$ is bounded.

2.2 Case-Based Decision Theory

The present work uses a modified version of the original CBDT in Gilboa and Schmeidler (1995).[5] The current model of CBDT is given by:

$$< P, A, R, s, u >$$

where P is a set of decision problems, A is a set of *actions*, R is a set of *results*, $s : (P \times A)^2 \times R \to [0, 1]$ is a *similarity function*, and $u : R \to \mathbb{R}$ is a *utility function*. A value $s((p, a), (p', a'); r')$ is a *similarity measure* between (p, a) and (p', a'), which may depend on r'. It is assumed that there exists a result, denoted by r_0, such that $u(r_0) = 0$.[6] The set of cases is defined to be $C \equiv P \times A \times R$. That is, a case is a triple (p, a, r) where p is a problem, a is an action, and r is its result. The decision rule of CBDT is described as follows. Suppose that a decision maker, who is characterized by u and s, is faced with a decision problem p, while the history at time

[4] In this case, $V(a|\bar{h})$ is expressed as

$$V(a|\bar{h}) = \sum_{r \in R} \mu_{\bar{h}}(f_T(a) = r) \left[u(r) + \max_{a' \in A} V(a'|\bar{h} \circ (a, r)) \right]$$

where $\bar{h} \circ (a, r)$ is a concatenated history. Expand the space of results to $R^* = H \times A \times R$. Then define $u^*(r)$ $(r \in R^*)$ as

$$u^*(r^*) = u^*(\bar{h}, a, r) = u(r) + \max_{a' \in A} V(a'|\bar{h} \circ (a, r)).$$

If the horizon is finite, construct them by backward induction. If the horizon is infinite, some additional assumptions are needed to ensure the existence of maximum.

[5] See Gilboa and Schmeidler (1995) for other modified forms.

[6] The subsequent argument would go through but become more complicated without this assumption.

T ($T = 1, 2, \cdots$) is $h_T = (c^1, \cdots, c^{T-1})$. In particular, the null history is denoted h_1. Let H_T be the set of all histories at time T. Write $H = \cup_{T=1}^{\infty} H_T$. Given a problem p and a history $h_T = ((p^1, a^1, r^1), \cdots, (p^{T-1}, a^{T-1}, r^{T-1})) \in H_T$ ($T = 1, 2, \cdots$), every action $a \in A$ is evaluated by

$$U(a|p, h_T) = \sum_{t=1}^{T-1} s((p, a), (p^t, a^t); r^t) u(r^t). \tag{2}$$

The decision maker chooses a maximand of U. A past case which is more similar to the possible present case is assigned more weight in evaluating an action. Write $s((p, a), c') = s((p, a), (p', a'); r')$.

Note that the present formulation is a generalization of the original model of Gilboa-Schmeidler, in which the value of an action, say, a, is given by

$$\sum_{(q,a,r)\in M} s(p, q) u(r), \tag{3}$$

where M is the set of past cases, and p is the present problem.[7]

Notice that in (3) the summation is only over the set of past cases in which action a was taken. In (2), the similarity measure is defined over the set of pairs of problems and actions as opposed to the set of problems alone as in Gilboa-Schmeidler. Their version also does not rely on the past result. The original theory is obtained if it is assumed that $s((p, a), (q, b); r) = 0$ for all $a \neq b$, and $s((p, a), (q, a); r) = s((p, b), (q, b); r')$ for all $p, q \in P$, all $a, b \in A$, and all $r, r' \in R$. These differences are necessary for the equivalence result. For example, if two acts have never been taken in the past, they are assigned the same utility values in (3), which are typically not the case in EU.

The present framework resembles the satisficing theory to the extent that the

[7] Gilboa and Schmeidler (1995) discuss several types of generalization, too. Also, see Gilboa and Schmeidler (1997a), proposing act-similarity functions which are "similar" to the present formulation.

original CBDT resembles it. The interpretation of the present formulation is that the decision maker evaluates each action by comparing the past scenarios with possible scenarios instead of comparing problems only. In fact, there is no reason to preclude this possibility *a priori.* Throwing a stone at someone is often more similar to kicking him than to shaking hands with him.

3 Embedding

This section shows that CBDT models can be embedded into EU models, and EU models can in turn be embedded into CBDT models. Throughout the section, we assume that P and R (or $\bar{R}$), the set of results, are all countable (can be finite).

Given a CBDT model, $< P, A, R, s, u >$, this model is said to be *embedded* in an EU model if there exists a model of the form $< (\Omega, \mathcal{F}), A, \bar{R}, f, \bar{u}, \mu >$ such that for all $h \in H$, $p \in P$, and all $a, b \in A$, $V(a|\bar{h}) \geq V(b|\bar{h})$ holds if and only if $U(a|p,h) \geq U(b|p,h)$ holds where (p, h) and $\bar{h}$ correspond to each other as described later. Similarly, given an EU model, $< (\Omega, \mathcal{F}), A, \bar{R}, f, \bar{u}, \mu >$, it is said to be *embedded* in a CBDT model if there exists a model of the form $< P, A, R, s, u >$ such that for any history $\bar{h} \in \bar{H} \backslash \{\bar{h}_1\}$ except the null history, any problem p, and for all $a, b \in A$, $U(a|p,h) \geq U(b|p,h)$ if and only if $V(a|\bar{h}) \geq V(b|\bar{h})$ holds. In the second definition, the null history should be excluded since CBDT can do little without a case.

3.1 Embedding CBDT into EU

Consider a model of CBDT, $< P, A, R, s, u >$ with H being the set of all histories. Construct its embedded model of EU, $< (\Omega, \mathcal{F}), A, \bar{R}, f, \bar{u}, \mu >$, in the following manner. First, let

$$\bar{R} = R \times P,$$

and

$$\bar{u}(r,p) = u(r)$$

for all $(r,p) \in \bar{R}$. A result specifies the problem of the next period as well. Then let

$$\Omega = P \times \prod_{t=1}^{\infty}(\bar{R}^{P\times A\times H_t}).$$

Interpretation is as follows. In each state $\omega = (p^1, \omega^1, \omega^2, \cdots) \in \Omega$, p^1 is a possible problem encountered by the decision maker in the first period, and $\omega^t = \{\bar{r}(p,a,h)\}_{(p,a,h)\in P\times A\times H_t}$ $(t = 1, 2, \cdots)$ is a possible realization in the tth period, i.e., p^t is a possible problem encountered by the decision maker in question, and $\bar{r}(p,a,h)$ is a possible result including the problem of the next period, which typically depends on the current problem p and choice a, and history h. Let $\mathcal{F}_t$ be the σ-algebra generated by the first $t+1$ elements of $(p^1, \omega^1, \cdots, \omega^t)$, i.e., $\mathcal{F}_t$ is the smallest σ-algebra which contains any set (cylinder) F_ω such that $\omega' = (p'^1, \omega'^1, \cdots)$ is in F_ω if and only if $p^1 = p'^1$ and $\omega^\tau = \omega'^\tau$ for all $\tau = 1, \cdots, t$. Let $\mathcal{F} = \sigma(\cup_{t=1}^{\infty}\mathcal{F}_t)$, the σ-field generated by $\mathcal{F}_t$'s.

With this preparation, the main task of embedding is to construct a probability measure on $(\Omega, \mathcal{F})$ such that for all $T = 1, 2, \cdots$, for any history $h_T = ((p^1, a^1, r^1), \cdots, (p^{T-1}, a^{T-1}, r^{T-1})) \in H_T$, and any current problem p^T, and any two actions a, $b \in A$, $V(a|\bar{h}_T) \geq V(b|\bar{h}_T)$ holds if and only if $U(a|p^T, h_T) \geq U(b|p^T, h_T)$ holds where $\bar{h}_T$ is the history (in EU) corresponding to (p^T, h_T), i.e., $\bar{h}_T = (p^1, (a^1, (r^1, p^2)), \cdots, (a^{T-1}, (r^{T-1}, p^T))) \in \bar{H}$. The following does this task inductively.

In the beginning of the Tth period, the decision maker knows $\bar{h}_T$, or equivalently, a problem p^T and the history $h_T = (c^1, \cdots, c^{T-1})$ with $c^t = (p^t, a^t, r^t)$. An action a

is evaluated by

$$U(a|p^T, h_T) = \sum_{t=1}^{T-1} s((p^T, a), c^t)u(r^t). \tag{4}$$

Let $\mu_{\bar{h}_T}$ satisfy

$$\mu_{\bar{h}_T}(f_{\bar{h}_T}(a, \cdot) = r) = \frac{1}{\mu^*} \sum_{t \in \{\tau | r^\tau = r\}} s((p^T, a), c^t), \tag{5}$$

for all $a \in A$ and all $r \in R \backslash \{r_0\}$, where μ^* is given by

$$\mu^* = \max_{a \in A} \sum_{r \in R} \sum_{t \in \{\tau | r^\tau = r\}} s((p^T, a), c^t), \tag{6}$$

which is always finite since only finitely many $s(\cdot)$'s are added. We need this normalization to make the probability of the entire set be one. For r_0, define

$$\mu_{\bar{h}_T}(f_{\bar{h}_T}(a, \cdot) = r_0) = 1 - \sum_{r \neq r_0} \mu_{\bar{h}_T}(f_{\bar{h}_T}(a, \cdot) = r). \tag{7}$$

Now, calculate $V(a|\bar{h}_T) = \sum_{r \in R} \mu_{\bar{h}_T}(f_{\bar{h}_T}(a, \cdot) = r)u(r)$. Substituting (5) into the right hand side of this expression, we obtain

$$\frac{1}{\mu^*} \sum_{t=1}^{T-1} s((p^T, a), c^t)u(r^t). \tag{8}$$

From (4), (8) is equal to

$$\frac{1}{\mu^*} U(a|p^T, h_T),$$

for all $a \in A$ where we also make use of $u(r_0) = 0$. Therefore, $V(a|\bar{h}_T) \geq V(b|\bar{h}_T)$ holds if and only if $U(a|p^T, h_T) \geq U(b|p^T, h_T)$ holds. Repeat this exercise for all $p \in P$ and all $h \in H$. The state space is so large that the above construction is consistent across problems and histories. A CBDT model is embedded into an EU model.

Note that in this embedding, an increase in $s((p^T, a), c^t)$ results in an increase in $\mu_{\bar{h}_T}(f_{\bar{h}_T}(a, \cdot) = r)$, i.e., the more similar the current situation is to the past case where action a led to result r, the higher the conditional probability of r given a. In other words, high similarity corresponds, at least in the above sense, to high correlation.

3.2 Embedding EU into CBDT

Consider a model of EU, $< (\Omega, \mathcal{F}), A, \bar{R}, f, \bar{u}, \mu >$ with $\bar{H}$ being the set of histories as defined in the previous section. Construct its embedded CBDT model, $< P, A, R, s, u >$, in the following manner. Let P, the set of problems, be defined as

$$P = \bar{H},$$

let $R = \bar{R}$, and let

$$u(r) = \bar{u}(r) + \epsilon$$

for some ϵ so that $\bar{u}(r) + \epsilon \neq 0$ for all $r \in R$. This linear transformation is needed since some ratios of utilities are needed later. In this formulation, a problem is considered as a history, i.e., totality of what is known at the time the problem arrives.

With this preparation, a similarity function $s : (P \times A)^2 \times R \to \mathbb{R}$ is inductively constructed. This task is started with the second period since case-based decision theory is meaningless without a case.[8] In the second period, $p^2 = \bar{h}_2$ is the problem encountered by the decision maker. Let s satisfy

$$s((\bar{h}_2, a), c^1) = \alpha \sum_{r \in R} \mu_{\bar{h}_2}(f_2(a) = r) \frac{u(r)}{u(r^1)} + \frac{1}{2} \tag{9}$$

where $\alpha = \alpha(\bar{h}_2, c^1)$ is constant across actions, but may depend on $\bar{h}_2$ and c^1 so that (9) is always between zero and one. Then it is verified that in the second period,

[8] One can cope with the first period problem if one modifies CBDT so that aspiration level is different across actions.

$U(a|\bar{h}_2, c^1) \geq U(b|\bar{h}_2, c^1)$ if and only if $V(a|\bar{h}_2) \geq V(b|\bar{h}_2)$ since

$$U(a|\bar{h}_2, c^1) = s((\bar{h}_2, a), c^1)u(r^1) = \alpha \sum_{r \in R} \mu_{\bar{h}_2}(f_2(a) = r)u(r) + \frac{1}{2} = \alpha V(a|\bar{h}_2) + \frac{1}{2},$$

holds.

Suppose that $s(\cdot, \cdot)$ has been defined up to period $T-1$ ($T = 3, 4, \cdots$). That is, $s((p, a), (p', a'); r')$has been defined for all $a, a' \in A$, all $r' \in R$, and all $p = \bar{h}_t$ and $p' = \bar{h}'_{t'}$ with $h_t = (c^1, \cdots, c^t)$ and $h'_{t'} = (c'^1, \cdots, c'^{t'})$ such that $t' < t \leq T-1$ ($T = 2, 3, \cdots$) and $c^\tau = c'^\tau$ for $\tau = 1, \cdots, t'$. Now define $s(\cdot, \cdot)$ for the problems possibly encountered by the decision maker in the Tth period. Let

$$s((p, a), (p', a'); r') = \frac{\alpha}{T-1} \sum_{r \in R} \mu_{\bar{h}_T}(f_T(a) = r) \frac{u(r)}{u(r')} + \frac{1}{2} \tag{10}$$

for all $a, a' \in A$, $r' \in R$, and all $p = \bar{h}_T$ and $p' = \bar{h}'_t$ with $h_T = (c^1, \cdots, c^{T-1})$ and $h'_t = (c'^1, \cdots, c'^t)$ such that $t < T$ and $c^\tau = c'^\tau$ for $\tau = 1, \cdots, t$. In (10), α is, as before, constant across actions and may depend on the past history so that the expression is between zero and one. In this manner, the similarity function s is defined for a relevant domain. To other pairs of cases, simply assign any arbitrary numbers between zero and one.

It is now verified that for any $\bar{h} \in \bar{H}$, $U(a|\bar{h}, h) \geq U(b|\bar{h}, h)$ if and only if $V(a|\bar{h}) \geq V(b|\bar{h})$. Indeed, for all $T = 1, 2, \cdots$, all $h_T = ((p^1, a^1, r^1), \cdots, (p^{T-1}, a^{T-1}, r^{T-1})) \in H_T$, and all $a \in A$,

$$\begin{aligned}
U(a|\bar{h}_T, h_T) &= \sum_{t=1}^{T-1} s((\bar{h}, a), (p^t, a^t); r^t)u(r^t) \\
&= \sum_{t=1}^{T-1} \left[\sum_{r \in R} \frac{\alpha}{T-1} \mu_{\bar{h}}(f_T(a) = r)u(r) + \frac{1}{2} \right] \\
&= \alpha \sum_{r \in R} \mu_{\bar{h}}(f_T(a) = r)u(r) + \frac{1}{2}(T-1) \\
&= \alpha V(a|\bar{h}_T) + \frac{1}{2}(T-1),
\end{aligned}$$

holds where the second equality is derived from (10).

Like in the previous embedding from CBDT to EU, there is a loose correspondence between correlation and similarity: if the previously encountered case gives the decision maker positive (resp. negative) utility, then the higher the probability of obtaining a result with positive (resp. negative) utility under a certain action, the more similar a pair of the present problem and the action is to the past case.

4 Conclusion

The present paper proved an equivalence result between expected utility theory and case-based decision theory. Two implications of this equivalence result are:

(i) for an outside modeler, these two theories are observationally equivalent so that it is impossible to distinguish the two unless he can access to the cognitive process of the decision maker;

(ii) neither theory is rejected in favor of the other on the ground of the lack of generality.

Several remarks are in order. First, although we regard embedding as the criterion for equivalence, this is certainly only one way to make comparison between the two theories. In particular, introducing some measure for complexity might help differentiating EU and CBDT. Indeed, in the present formulation, given an original model, its embedded model often becomes complicated. For example, for a CBDT model with a finite set of problems, one typically needs an infinite state space for embedding. On the other hand, for an EU model with a finite information set, one typically needs an infinite set of problems in order to embed the CBDT model. Moreover, while one can embed a model based on one theory into a model based on the other theory, intuition which motivates the original model may be lost. It is intuition on phenomena rather than formal capability of description of certain situations that differentiates the two theories.

Another way to differentiate EU and CBDT is to impose appropriate restrictions on these two theories to see their respective predictive power. There are many empirical studies based on EU, which are normally conducted by imposing certain restrictions on the model. Sometimes, constructing a specific model itself becomes a restriction. Building a descriptive CBDT model is the next step if one wants to compare the two theories.

Finally, this paper does not intend to dethrone expected utility theory, nor does it try to diminish the significance of case-based decision theory. If one understands the relation between the two processes of decision makings better than before, the goal of the present paper would be achieved.

Appendix

This appendix provides the set of axioms that leads to the decision rule in CBDT. It essentially follows Gilboa and Schmeidler (1995) with minor modification. A preference order over actions is to be defined. Unlike Gilboa-Schmeidler, for every p and history h, there is a separate order $\succeq_{p,h}$ over the finite set of actions, A. It is assumed that $R = \mathbb{R}$, and results are already measured in utility term. It is also assumed that problems are always different after different histories. In evaluating an action, take into account the results that other actions led to. Formally, given h, let $X = \mathbb{R}^{|h|}$, where $|h|$ is the number of cases in h, and the preference order is on $Z = X^{A\times A}$. Given p and h, each act profile, a history related to an action, is represented by a point in Z. If a case $c = (p, a, r)$ is the tth case in h, then an entry for an action b and case c is given by

$$z_t(a_1, a_2) = \begin{cases} u(r) & \text{if } a_1 = b, a_2 = a, \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

For example, suppose that there are only two actions, a and a', and two results, r and r'. Suppose further that $h = ((p, a, r), (p', a', r'))$. Then the act profile for a is represented by

$$((u(r), 0; 0, 0), (0, u(r'); 0, 0)). \tag{12}$$

In this expression, the first four entries correspond to the first case, the second of the final four entries is $u(r')$ since this entry corresponds to action a, which is now evaluated, and the action chosen, a', and so forth.

It is not assumed that $\succeq_{p,h}$ is a complete order on Z. As one may see from the above expression, it is meaningless to compare two vectors in Z unless they satisfy (11). Let Z^* be the set of all the vectors in Z which satisfy (11). Two act profiles may be compared if and only if both are in Z^*. Now state four axioms on $\succeq_{p,h}$.

A1 For all $p \in P$ and all h, $\succeq_{p,h}$ is reflexive and transitive on Z. For all $x, y \in Z^*$, either $x \succeq_{p,h} y$ or $y \succeq_{p,h} x$ holds.

A2 (continuity) For all p, all h, and all $x \in Z^*$, $\{y | y \succeq_{p,h} x\}$ and $\{y | x \succeq_{p,h} y\}$ are closed (in the standard topology in R^n).

A3 (separability) For all p, h, and all $x, y, z, w \in Z^*$ with $x + z$ and $y + w$ in Z^* as well, if $x \succeq_{p,h} y$, and $z \succeq_{p,h} w$, then $(x + z) \succeq_{p,h} (y + w)$.

A4 (strict separability) For all p, h, and all $x, y, z, w \in Z^*$, if $x \succeq_{p,h} y$, and $z \succ_{p,h} w$, then $(x + z) \succ_{p,h} (y + w)$.

Proposition. If A1-A4 hold, then for all $p \in P$ and all $h = (c^1, \cdots, c^{T-1}) \in H_T$ $(T = 2, 3, \cdots)$ with $c^t = (p^t, a^t, r^t)$ $(t = 1, \cdots, T-1)$, there exists a function $s_{p,h} : \{c^1, \cdots, c^{T-1}\} \to \mathbb{R}$ such that $x \succeq_{p,h} y$ iff

$$\sum_{t=1}^{T-1} \sum_{a \in A} s_{p,h}(c^t) x(c^t, a, a^t) \geq \sum_{t=1}^{T-1} \sum_{a \in A} s_{p,h}(c^t) y(c^t, a, a^t).$$

This proposition is proven step by step. In the following, write

"$\succeq$" instead of $\succeq_{p,h}$ whenever there is no confusion.

Observation.

1. For all $x, y \in Z^*$,

$$x \succeq y \Leftrightarrow -y \succeq -x.$$

2. For all $x, y, z, w \in Z^*$ such that $x + z$, $y + w \in Z^*$, and $z \sim w$,

$$x \succeq y \Leftrightarrow (x + z) \succeq (y + w),$$

holds

Proof of the observation.

1. It is sufficient to prove one direction. Suppose $x \succeq y$, and $-x \succ -y$. Then by A4, $0 = x + (-x) \succ y + (-y) = 0$. A contradiction.

2. Assume the presumptions of the claim. A3 and the assumptions imply that

$$x \succeq y \Rightarrow (x + z) \succeq (y + w).$$

As for the converse, by the first observation, $-z \succeq -w$ holds. Then A3 implies

$$(x + z) \succeq (y + w) \Rightarrow x = (x + z + z') \succeq (y + w + w') = y.$$

One problem is that Z^* is not a linear vector space. Given $p \in P$ and h, another binary relation $\succeq_{p,h}$ is defined as follows. First, it is complete and satisfies

$$x \succeq_{p,h} y \quad \text{iff} \quad x \succeq_{p,h} y,$$

for all $x, y \in Z^*$. Omit the subscript from $\succeq'$, too.

A1' $\succeq'$ is complete (note that $\succeq$ is not), reflexive and transitive on Z.

A2' (continuity) $\forall x \in Z$, $\{y | y \succeq' x\}$ and $\{y | x \succeq' y\}$ are closed.

A3' (separability) $\forall x, y, z, w \in Z$, $x \succeq' y$ and $z \succeq' w \Rightarrow (x + z) \succeq' (y + w)$.

A4'(strict separability) $\forall x, y, z, w \in Z$, $[x \succeq' y$ and $z \succeq' w] \Rightarrow (x+z) \succeq' (y+w)$.

It is verified that A1-A4 imply A1'-A4' for those vectors in Z^*. We need some other facts.

F1 $\forall x, y \in Z$, $x - y \succeq' 0 \Leftrightarrow x \succeq' y$.

F2 (convexity) $\forall x \in Z$, sets $\{y|y \succeq' x\}$, $\{y|x \succeq' y\}$, $\{y|y \succ' x\}$, and $\{y|x \succ' y\}$ are convex.

F3 In Z, define

$$A = \{x|x \succeq' 0\},$$

and

$$B = \{x|0 \succ' x\}.$$

Then A is closed and convex, while B is open and convex. $A \cap B = \emptyset$ and $A \cup B = Z$ hold, too.

F4 If $B = \emptyset$, then $s(\cdot) \equiv 0$ satisfies the representation condition. If $B \neq \emptyset$, then by the separating hyperplane theorem there exists a linear functional $S : Z \to \mathrm{R}$ such that

$$S(x) \geq 0, \forall x \in A,$$

and

$$S(x) < 0, \forall x \in B.$$

Thus, there exists $s_{p,h} : (P \times A \times R) \times A^2 \to \mathrm{R}$ such that

$$x \succeq_{p,h} y \text{ iff } \sum_{t=1}^{T-1} \sum_{a \in A} s_{p,h}(c^t, a, a^t) x(c^t, a, a^t) \geq \sum_{t=1}^{T-1} \sum_{a \in A} s_{p,h}(c^t, a, a^t) y(c^t, a, a^t), \tag{13}$$

for all x and y in Z. By F1, for all $x, y \in Z^*$, $x \succeq_{p,h} y$ holds if and only if (13) holds.

This proves the proposition. Since any two problems with different histories are different, CBDT constructed in the text is obtained by letting $s((p, a), (p', a'); r') = \alpha s_{p,h}((p', a', r'), a, a') + \beta$ where α and β are chosen in such a way that the similarity value $s(\cdot)$ fits between zero and one.

References

[1] Allais, M. & O. Hagen (1979) (eds.), *Expected Utility Hypotheses and the Allais Paradox,* Dordrecht, Holland: Reidel.

[2] Ellsberg, D. (1961), "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics,* 643-669.

[3] Gilboa, I. & D. Schmeidler (1995), "Case-Based Decision Theory," *Quarterly Journal of Economics*, 605-639.

[4] Gilboa, I. & D. Schmeidler (1996), "Case-Based Optimization," *Games and Economic Behavior*, 1-26.

[5] Gilboa, I. & D. Schmeidler (1997a), "Act Similarity in Case-Based Decision Theory," *Economic Theory*, 47-61.

[6] Gilboa, I. & D. Schmeidler (1997b), "Cumulative Utility and Consumer Theory," *International Economic Review*, 737-761.

[7] Machina, M. J., (1987), "Choice under Uncertainty: Problems Solved and Unsolved." *Journal of Economic Perspectives*, 121-154.

[8] March, J.G. & H.A. Simon (1958), *Organizations*, New York: Wiley.

[9] Savage, L.S. (1954), *The Foundations of Statistics*, New York: Wiley.

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