Discussion Paper No. 1053

Case-Based Knowledge Representation*

by

Itzhak Gilboa+

and

David Schmeidler++

May, 1993

Revised Version, January 1994

* We are grateful to Eva Gilboa, Kimberly Katz, Akihiko Matsui, and Ariel Rubinstein for discussions and references.

+ KGSM-MEDS, Northwestern University, Leverone Hall, Evanston, IL 60208.

++ Department of Statistics and Recanai School of Business, Tel Aviv University, Tel Aviv 69978, Israel; Department of Economics, Ohio State University, Columbus, OH 43210-1172. Currently visiting the Department of Economics and KGSM, Northwestern University, Evanston, IL 60208.

Abstract

The representation of knowledge in terms of rules is fraught with theoretical problems, such as the justification of induction, the "right" way to do it, and the revision of knowledge in face of contradictions.

In this paper we argue that these problems, and especially the inconsistency of "knowledge," are partly due to the fact that we pretend to know what in fact cannot be known. Rather than coping with the problems that explicit induction raises, we suggest to avoid it. Instead of formulating rules which we supposedly "know," we may make do with the knowledge of actual cases from our experience.

Starting from this viewpoint, we continue to derive Case-Based Decision Theory (CBDT), and propose it as a less ambitious, yet less problematic theory of knowledge representation. CBDT deals with decision making under uncertainty, and can be viewed as performing implicit induction, that is, as using past experience to make decisions, without resorting to the explicit formulation of rules.

We discuss two levels on which implicit induction takes place, and the corresponding two roles that "rules" may have in case-based decision making. We also discuss the process of learning and the concept of "expertise" as they are reflected in our model.

1. Introduction

The literature in philosophy and artificial intelligence (AI) typically assumes that one type of objects of knowledge are "rules," namely general propositions of the form "For all x, $P(x)$." While some of these rules may be considered, as a first approximation at least, "analytic propositions," a vast part of our "knowledge" consists of "synthetic propositions."[1] These are obtained by induction, that is, by generalizing particular instances of them.

The process of induction is very natural. Asked, "What do you know about... ?", people tend to formulate rules as answers. Yet, as was already pointed out by Hume (1748), induction has no logical justification, and may well lead to erroneous conclusions. Furthermore, it is not clear how does one generalize instances – or "cases" – into rules. Wittgenstein (1922, 6.363), for instance, argued that "The procedure of induction consists in accepting as true the *simplest* law that can be reconciled with our experiences." But the notion of "simplicity" may be very vague, subjective and language-dependent. (See, for instance, Sober (1975) and Gärdenfors (1990).) Thus it is far from obvious what the "right" way to perform induction is (from a normative point of view), or how people actually do it (taking a descriptive approach).

The subjective nature of induction, coupled with its dubious justification, lead to the problem of knowledge revision and update. Much attention has been devoted to this problem in the recent literature in philosophy and AI. (See Levi (1980), McDermott and Doyle (1980), Reiter (1980) and others.) While we do not attempt to provide here even the briefest survey of any of the above-mentioned fields, we would like to point out that induction raises three main questions:

(i) How and to what extent can induction be justified?

(ii) How do people do induction, and how should it be done?

(iii) How do people revise their knowledge when they discover that their induction process has yielded unwarranted conclusions, and how should they revise their knowledge?

In view of these problems, we suggest to consider knowledge representation models which avoid induction to begin with. Rather than

---

1 While this distinction was already rejected by Quine (1953), we still find it useful for the purposes of the present discussion.

dealing with rules, one may restrict the knowledge represented to cases. This proposal was first made by Schank (1986) and Riesbeck and Schank (1989). In Section 2 we motivate and discuss our approach, and compare it with theirs.

One may ask, in what sense do cases represent knowledge, and are they good enough? This leads us to a basic question, namely, why do we want to represent knowledge in the first place? In Section 3 we present the view that knowledge, and perforce knowledge representation, are needed for decision making. This leads us to a model of a "case" which emphasizes the role of one's decision. The structure which follows is identical to that we assumed as primitive in Gilboa and Schmeidler (1992).

In Section 4 we deal with the question, what does one do with this knowledge? We consider the notion of "similarity" between cases, and some ways in which it can be used in decision-making, making a case for the particular way we use it in case-based decision theory (CBDT). This theory suggests that, when confronted with a decision problem under uncertainty, people tend to choose acts which performed well in similar cases in the past. It may thus be viewed as modeling "implicit induction:" according to CBDT, decision makers behave as if they believed in generalizations of specific instances. However, the theory does not explicitly mention "rules."

Section 5 is devoted to the relationship between CBDT and the notion of "rules." We argue in sub-section 5.1 that a decision-maker who follows CBDT may be described as if (s)he "knows" certain rules at certain points of time. However, when these rules seem to be contradicted by observations, or to contradict each other, a case-based decision maker continues to make decisions according to CBDT in a coherent manner, which may or may not be also represented as following a new set of rules. In sub-section 5.2 we discuss two ways in which rules may be incorporated into the CBDT model. Rather than treating rules as objects of knowledge, we think of them as offering an efficient language to transfer information. The distinction between the two roles of rules is based on the subject matter of the information transferred: cases vs. similarity judgments.

In Section 6 we discuss some limitations and possible extensions of CBDT. Specifically, we present an intuitive example in which our original, somewhat naïve version of CBDT fails. Analyzing this example, we note that, as opposed to the model presented in Section 4, the similarity function itself may change as a result of past experience. This leads us to distinguish between two levels of (implicit) induction: first-order induction, which is captured by CBDT, and second-order induction, that is, learning how to perform first-order induction, which calls for a generalization of CBDT in its present form.

Section 7 summarizes the implications of sections 4-6 with respect to the process of learning and the definition of "expertise." Finally, section 8 concludes with some comments.

2. What Can Be Known?

When we consider the issue of knowledge representation, we should first ask ourselves, what is there to represent? What is it that we know? And, more specifically, can we know general rules?

To this last question one can hardly find a more well-known answer than Hume's. In Hume (1748, Section IV) we find,

> "... The contrary of every matter of fact is still possible; because it can never imply a contradiction, and is conceived by the mind with the same facility and distinctness, as if ever so conformable to reality. *That the sun will not rise to-morrow* is no less intelligible a proposition, and implies no more contradiction than the affirmation, *that it will rise*. We should in vain, therefore, attempt to demonstrate its falsehood."

That is, no synthetic proposition whose truth has not yet been observed is to be deemed necessary. In particular, to the extent that rules are useful – that is, to the extent that they generalize our experience and have implications regarding the future – they cannot be *known*.

We suspect that many of the theoretical problems relating to knowledge representation are due to the unwarranted induction process. Hume argues (and seems to convince most of us[2]) that we cannot know the generalized rules induction offers. If we nevertheless insist to treat these rules as "known," there is little reason to wonder at the theoretical difficulties that ensue.

---

2 Quine (1969) writes, "... I do not see that we are further along today than where Hume left us. The Humean predicament is the human predicament."

It is therefore natural to consider an alternative approach, which, instead of dealing with the problems induction poses, will attempt to avoid induction in the first place. According to this approach, knowledge representation should confine itself to those things which can indeed be known. And these include only facts which were observed, not "laws;" cases can be known, while rules can at best be conjectured. One of the theoretical advantages of this approach is that, while rules tend to give rise to inconsistencies, cases cannot contradict each other.

Our approach is closely related to (and partly inspired by) the theory of Case-Based Reasoning (CBR) proposed by Schank (1986) and Riesbeck and Schank (1989). (See also Kolodner and Riesbeck (1986) and Kolodner (1988).) In this literature, CBR is proposed as a better AI technology, and a more realistic descriptive theory of human reasoning than rule-based models (or systems). However, our approach differs from theirs in motivation, emphasis and the analysis that follows. We suggest the case-based approach as a solution to, or rather a way to avoid the theoretical problems entailed by explicit induction. Our focus is decision-theoretic, and therefore our definition of a "case" highlights the aspect of decision making. Finally, our emphasis is on a formal model of case-based decisions, and the extent to which such a model captures basic intuition.

As we will see in the sequel, our model does involve some implicit process of "induction;" indeed, whenever one learns from experience one may be claimed to engage in "implicit induction." Yet our model does not resort to explicit formulation of general "rules" which are supposedly known. "Rules" may still play certain roles in possible extensions of the model, but they are not essential to it, and, more importantly, they are not the object of knowledge. We remain faithful to Hume's claim that it is only a case that can be known.

One has to admit, however, that even cases may not be "objectively known." The meaning of "empirical knowledge" and "knowledge of a fact" are also a matter of heated debate. (For a recent anthology on this subject, see Moser (1986).) Furthermore, as has been argued by Hanson (1958), observations tend to be theory-laden; hence the very formulation of the "cases" that we allegedly observe may depend on the "rules" that we believe to apply. It therefore appears that one cannot actually distinguish cases from rules, and even if one could, such a distinction would be to no avail, since cases cannot be known with certainty any more than rules can.

While we are sympathetic to both claims, we tend to view them as somewhat peripheral issues. We still believe that the theoretical literature on epistemology and knowledge representation may benefit from drawing a distinction between "theory" and "observations," and between the knowledge of a case and that of a rule. Philosophy, like other social sciences, often has to make do with models which are "first approximations," "metaphors," or "images," – in short, models that should not be taken too literally, nor be expected to be completely accurate. We will therefore allow ourselves the idealizations according to which cases can be "known," and can be observed independently of rules or "theories."

The discussion so far excludes induction, and leaves us with the knowledge of rather amorphous "cases." What are they made of? How do we use them? In what sense do they represent knowledge, apart from merely existing? – In the next sections we make some preliminary steps towards answering these questions.

3. What Is Knowledge Representation Needed For?

When the object of knowledge is a rule, or a "theory," logic provides a natural (even if not necessarily unique) language to describe their structure. Thus logic allows us to reason about rules in a general and abstract way, while capturing their essence. At this point of our discussion, the notion of a "case" is too vague to allow any analysis of interest. It seems that we need to endow "cases" with some structure to make them more tangible and amenable to analysis.

What is a case made of, then? Or, what type of creature will it be represented by in a formal model? The answer, we believe, lies in understanding what we expect of the model. That is, we should first ask ourselves, why would we represent knowledge at all?

We take the view that knowledge representation is required to facilitate the use we make of the knowledge. And we use knowledge when we act, that is, when we have to make decisions. This implies that the structure of a "case" should reflect the decision-making aspect of it.

We make the somewhat-idealized assumption that all cases involve acts, or decisions. (See comment 8.2 below for a discussion of this assumption and possible relaxations of it.) Focusing on the decision made in a given case, it is natural to divide a "case" to three components: (i) the conditions at which the decision was made; (ii) the decision; and (iii) the results. If we were to set our clocks to the time of decision, these components would correspond to the past, the present and the future, respectively, as perceived by the decision maker. In other words, the "past" contains all that was known at the time of decision, the "present" – the decision itself, while the "future" – whatever followed from the "past" and the "present." In Gilboa and Schmeidler (1992) we dub the three components "problem," "act," and "result."

Formally, in CBDT we postulate three abstract sets as primitive:

P – the set of (decision) *problems*
A – the set of possible *acts*
R – the set of conceivable *results*.

Correspondingly, the set of *cases* is simply the set of ordered triples, or the product of the above:

$$C \equiv P \times A \times R$$

where, at any given point of time, the decision maker has a *memory*, which is a finite subset of cases $M \subseteq C$.

We thus assume that one's knowledge is given by one's memory, and it is simply a collection of cases. Each case describes a certain decision situation which one may recall. While one cannot presume to know more than bare cases, one's subjective judgment will still play a role in the way one uses one's knowledge.

4. The Introduction of Similarity

How are the cases used in decision making, then? Again, we resort to Hume (1748) who writes, "In reality, all arguments from experience are founded on the similarity which we discover among natural objects, and by which we are induced to expect effects similar to those which we have found to follow from such objects. ... From causes which appear *similar* we expect similar effects. This is the sum of all our experimental conclusions."

Thus Hume suggests the notion of similarity – in his formulation, between "causes" as well as between "effects" – as key to the procedure by which we use cases. Since this procedure is to be applied to a given decision problem $p \in P$, it makes sense to define the similarity on problems as well (as opposed, say, to whole cases). Thus we postulate a *similarity function* which may be normalized to take values not exceeding unity:

$$s: P^2 \to [0,1].$$

It is important to note that the similarity function is not assumed to be "known" by the decision maker in the same sense cases are. While cases are claimed to be "objectively known," the similarity is a matter of subjective judgment. Where does it come from, then? Or, to be precise, what determines our similarity judgments (from a descriptive viewpoint), and what should determine it (taking a normative view) ?

Unfortunately, we cannot offer any general answers to these questions. However, it is clear that the questions will be better defined, and the answers easier to obtain, if we know how the similarity is used in the decision making process. Differently put, the similarity function, being a theoretical construct, will gain its meaning only through the procedure in which it is used.

One relatively obvious candidate for a decision procedure is a "nearest neighbor" approach. It suggests that, confronted with a decision problem, we should look for the most similar problem which has been encountered in the past. Unfortunately, this approach does not seem to be very satisfactory: if one is happy with the result of the "nearest" case, one may indeed choose the same act which was chosen in that case. But which act is to be chosen if the "nearest" case ended up yielding a disastrous outcome? Furthermore, assume that the nearest case resulted in a reasonable outcome, but in many other cases, which are somewhat less similar to the problem at hand, a different act was chosen, and it yielded particularly good outcomes in all of them. Is it still reasonable, from either descriptive or normative point of view, to choose the act which was chosen in the *most* similar case?

It therefore appears that a more sensible procedure would not rely solely on the similarity function, nor would it depend only on the most similar case. First, one needs to have a notion of a *utility*, i.e., a measure of desirability of outcomes, and to take into account both the similarity of the problem and the utility of the result when using a past case in decision making. Second, one should probably use as many cases as deemed relevant to the decision at hand, rather than the "nearest neighbor" alone.

We thus postulate a *utility function*

$$u: R \to \Re$$

which assigns numerical values to outcomes, and which is to be interpreted as a measure of desirability. CBDT prescribes that acts be evaluated by a similarity-weighted sum of the utility they yielded in past cases. That is, given memory $M \subseteq C$ and a problem $p \in P$, every act $a \in A$ is evaluated by the functional

$$U(a) = U_{p,M}(a) = \sum\nolimits_{(q,a,r)\in M} s(p,q)u(r)$$

where a maximizer of this functional is to be chosen. (In case the summation is over an empty set, the act is assigned a "default value" of zero. This value may be interpreted as the "aspiration level" of the decision maker. For details, see Gilboa and Schmeidler (1992).)

This decision rule is perhaps the simplest one which takes into account both similarity and utility, and does so for all the cases in which a given act was chosen. Yet one may consider many other decision criteria which will make use of all the above information in a sensible way. The particular rule we suggest here should be taken as a rough, preliminary model; like most models in the social sciences, it should not be taken too literally. For instance, it should not come as a great surprise to us if in certain circumstances the additive separability of the formula above will prove to be too restrictive. (In particular, see the discussion in Section 6 below.)

At any rate, in Gilboa and Schmeidler (1992) we provide an axiomatic derivation of this rule. That is, we assume as given datum a binary relation over (actual and hypothetical) acts, to be interpreted as a preference relation, and provide necessary and sufficient conditions on it for there to exist a similarity function such that for every pair of acts, the one which has the higher U-value is the one preferred by the given relation. Thus the notion of "similarity" in our model may be derived from preferences, in a way that parallels de Finetti's derivation of "subjective probability." (See de Finetti (1937).)

While de Finetti assumed a concept of "utility" as primitive, Savage (1954) provides an axiomatic derivation of both "subjective probability" and "utility" simultaneously. Similarly, one may extend our model so that "(subjective) similarity" and "utility" would be derived from preferences together.

The axiomatic derivation serves several purposes. First, and perhaps most importantly, it relates theoretical constructs (such as "similarity") to observable terms (such as "preference"), thus endowing the model with "cognitive significance." Second, it may help us judge the reasonability of a particular theory. In the case of the functional above, the axioms are quite simple and intuitive (as the additively-separable functional form itself). Yet they are not unique in being so; indeed, we also offer an equally-attractive axiomatization of another rule, in which the similarity function is "normalized" so as to sum up to unity for each act separately. Finally, and correspondingly, the axioms may be of help in testing the theory and in measuring the "similarity" function.

The importance of axiomatic derivations notwithstanding, we should emphasize that an axiomatization does not explain whence the similarity stems. It cannot replace a psychological theory in accounting for similarity judgments, nor is it of great help to someone who would like to have a "reasonable" similarity function. Thus, while we can argue for employing a certain decision procedure rather than another, and can attempt to rank procedures in terms of descriptive and normative appeal, we can still say very little about the origin of the similarity function.

5. Where Do Rules Fit In?

5.1 Do CBDM's Know Rules?

Equipped with knowledge of past cases, similarity judgments and utility valuations, case-based decision makers (CBDM's) may go about their business, taking decisions as the need arises, without being bothered by induction and general rules, and without ever having to deal with inconsistencies. Yet, if they are asked why they made a certain decision, they are likely to give answers in the form of rules. For instance, if you turn a door's handle in order to open it, and you are asked why you chose to turn the handle, it is unlikely that your answer would be a list of particular cases in which this trick happened to work. You are most likely to say, "Because the door opens when one turns the handle," i.e., to formulate your knowledge as a general rule.

In a sense, then, case-based decision makers often behave *as if* they knew certain rules. Furthermore, whenever the two descriptions are equivalent, the language of "rules" is much more efficient and parsimonious than that of "cases." However, the advantage of a case-based description is in its flexibility. In the same example given above, suppose that a person (or, say, a robot we are now programming) finds out that the door refuses to open despite the turning of the handle. A rule-based decision maker, while struggling with the door, also undergoes internal, mental commotion. Not only is the door still shut, one's rule base has been contradicted. Some rules will have to be retracted, or suspended, before our decision maker will be able to use the second rule "If turning the handle doesn't work, call the janitor." By contrast, a CBDM has only the door to struggle with. The failure of the turn-the-handle act will make it less attractive, and will make another act the new U-maximizer. Perhaps the very failure of the first act will make the whole situation look more similar to other cases, in which only a janitor could open the door. At any rate, a CBDM does not have to apply any special procedures to re-organize his/her knowledge base. Like the rule-based one, the case-based decision maker also has to wait for the janitor; but (s)he does so with peace of mind.

In other words, the rules-vs.-cases choice faces us with a familiar tradeoff between parsimony and accuracy. Rules are simply described, and they are therefore rather efficient as a means to represent knowledge; but they tend to fail, namely to be contradicted by evidence and by their fellow rules. Cases tend to be numerous and repetitive, but they are never a source of inconsistency. In view of the theoretical problems associated with rules, it appears that case-based models are a viable alternative.

In a slightly different metaphor, "cases" may be considered a low-level description of human knowledge, while "rules" are the higher-level one. As long as the two are equivalent, one would typically prefer the high-level description. When the latter fails, one has to settle for the low-level one.

In the sciences we find examples of both cases. For instance, all natural sciences (and perhaps the social ones as well) are commonly believed to be in-principle reducible to physics. Yet this reduction is used only as a last resort. In the case of chemistry and biology, for instance, most of the discourse is done in a high-level language, without explicit reference to the underlying physical principles. In the example of meteorology, on the other hand, high-level theories such as "Clouds imply rain" or "Whenever Zeus is angry, a thunderstorm occurs" do not seem to be very useful, and are reluctantly discarded in favor of dreary and detail-laden flow equations.

Similarly, it is possible, in principle, that rule-based models will be as successful as, say, biological ones, and that the level of cases will be completely redundant for knowledge representation. On the other hand, it may also be the case that rule-based models will turn out to be hopelessly naïve. While it is probably too early to decide one way or the other, it appears that rule-based models are theoretically problematic enough to warrant a competing theory of knowledge representation.

5.2 Two Roles of Rules

While CBDT rejects the notion of "rules" as objects of knowledge, it may still find them a useful tool. Even if one is convinced that they are too crude to be "correct," rules are still a convenient "first-approximation" of cases. Furthermore, they provide a language for efficient information transmission. As such, rules can have two roles:

(i) A rule may summarize many cases. If we think of a rule as an "ossified case," (Riesbeck and Schank (1989)) it is natural to imagine one individual (system) telling another about many cases by conveying a single rule which applies in all of them;

(ii) A rule may point to similarity among cases. That is, even if two people (systems) have the same cases in their memory, one may be unaware of certain common denominators among them. Especially when the amount of information is vast, an abstract rule may help in finding analogies. For instance, claims such as "Peace-keeping forces can succeed only if the belligerent forces want them to" or "The stock market always plunges after presidential elections" serve mainly to draw the reader's attention to known cases, rather than to tell him/her about unknown ones. Furthermore, many "laws" in the social sciences, though formulated as rules, should be thought of as "proverbs:" they do not purport to be literally true. Rather, their main goal is to affect similarity judgments. In this capacity, the fact that rules tend to contradict each other poses no theoretical difficulty. Indeed, it is well-known that proverbs are often contradictory.[3] Once they are incorporated into the similarity function, the latter will determine which prevails in each given decision situation.

To sum, CBDT may incorporate rules, and experts' knowledge formulated as rules, either as a summary of cases or as a similarity-defining feature of cases. Yet, within the framework of CBDT, rules are not taken literally, they are not assumed "known," and their contradictions are blithely ignored.

---

3 The notion of a "rule" as a "proverb" also appears in Riesbeck & Schank (1989). They distinguish among "stories," "paradigmatic cases," and "ossified cases," where the latter "look a lot like rules, because they have been abstracted from cases." Thus, CBR systems would also have "rules" of sorts, or "proverbs," which may, indeed, lead to contradictions.

6. Memory-Dependent Similarity and Two Levels of Induction

In the model of Section 4 the similarity function is assumed to be memory-independent. However, the similarity function may also depend on the problems that were encountered in the past, as well as on the results obtained in them.

We start with the following example. Consider a hypothetical decision maker named Agent. For grammatical simplicity let us endow Agent with gender, say, female. Suppose that Agent has two coffee shops in her neighborhood, 1 and 2. Once in a shop (which determines a "problem"), she has to decide what to order (which act to choose). In the past, she has visited both of them once in the morning (M) and once in the evening (E), ordering "cafe latte" in each of these four problems. The four problems she recalls are: $(M1, M2, E1, E2)$. (Notice that which shop to go to is *not* a decision variable in this story.) Now assume that the quality of the coffee she had was either 1 (high) or -1 (low). Let us compare two possible memories: in the first, the result sequence is $(1,1,-1,-1)$, while in the second it is $(1,-1,1,-1)$. In the first case Agent would be tempted to assume that what determines the quality of cafe latte is the time of the day. Correspondingly, she is likely to put more weight on this attribute in future similarity judgments. On the other hand, in the second case, the implicit induction leads to the conclusion that coffee shop 1 simply serves a better latte than 2, and more generally, that the coffee shop is a more important attribute than the time of the day. In both cases, the way similarity will be evaluated in the future depends on memory as a whole, including the results that were obtained.

Generally, one may distinguish between two levels of inductive reasoning in the context of CBDT. First, there is "first order" induction, by which similar past cases are implicitly generalized to bear upon future cases, and, in particular, to affect Agent's decision in a new problem. The version of CBDT presented here does attempt to model this process, if only in a rudimentary way. However, there is also "second order" induction, by which Agent learns not only what to "expect" in a given case, but also *how to conduct first-order induction*, namely, how to judge similarity of problems. The current version of CBDT does not model this process. Moreover, it implicitly assumes that it does not take place at all.

Specifically, one would expect that when some process of "second-order induction" affects similarity judgments, there would be some plausible counterexamples to U-maximization. Indeed, consider the following example: coffee shops 1, 2, 3, and 4 serve both cafe latte and cappuccino. Shops 1 and 2 are in Agent's neighborhood, and she had the opportunity to try both orders in each of them, both in the morning and in the evening. Shops 3 and 4 are in a different town, and probably bear little resemblance to either 1 or 2. So far Agent has only tried the latte in 3 in the afternoon (A), and the cappuccino in 4 at night (N). Both resulted in high-quality coffee. The next afternoon she is in shop 4, trying to decide what to order. Based on her experience, both acts are likely to have a positive U-value. Yet, Agent may still distinguish between them depending on her similarity function. If she puts more weight on the time of the day, the latte, which was successfully tried in the afternoon, is a more promising choice; if, however, she tends to "believe" that similarity is mostly determined by the shop, she should perhaps order cappuccino, as she did yesterday night in the same shop. Let us now consider the following vectors (where empty entries denote zeroes):

act	Problems									
profiles	$M1$	$M2$	$E1$	$E2$	$M1$	$M2$	$E1$	$E2$	$A3$	$N4$
x	1	1					−1	−1	1	
y			−1	−1	1	1				1
z	1		1			−1		−1	1	
w		−1		−1	1		1			1
d		−1	1			−1	1			

In this example memory contains ten cases. It is convenient to think of the ten problems as formally distinct; for instance, each may be uniquely identified by a time parameter. However, in the table we suppress this parameter and specify only the features of the problem which are deemed relevant, namely the coffee shop and the time of the day.

We further assume that in each of the problems only two acts – "cafe latte" and "cappuccino" – were available. The vectors x, y, z, w, and d are "act profiles;" that is, they designate a conceivable history of an act. If the act was not chosen in a particular case, it is assigned a default "utility" value of zero. If it was, it is assigned the actual utility that resulted from it in this case.

We would now like to consider the preferences between cafe latte and cappuccino under two separate scenarios. In the first, the preference question would reduce to comparing x and y, while in the second – to comparing z and w. Suppose first that x is the act profile of "cafe latte" and y – of "cappuccino." Focusing on the first two rows in the table, the results obtained in the first eight problems clearly indicate that it is the time of the day that matters: all morning coffees were high quality, all evening ones were low quality. Based on this "general observation," Agent has learnt to appreciate the crucial role of the time of the day, and she is unlikely to put too much weight on the "night problem" $N4$ when making a decision in the afternoon. Thus, she expresses a preference for x over y when faced with the problem $p = A4$.

By a similar token, when comparing z and w, Agent concludes that the shop is very important, but the time of the day does not really matter. Hence she puts more weight on the experience in the same shop – problem $N4$ – and decides to order cappuccino, i.e., she prefers w over z (for the same decision problem $p = A4$).

It is easily verified that this preference pattern is inconsistent with U-maximization for a fixed similarity function s. Indeed, for any such function s, since $z - x = w - y = d$, we have

$$U(z) - U(x) = U(w) - U(y) = U(d)$$

from which we derive

$$U(x) - U(y) = U(z) - U(w).$$

That is, x is preferred to y *if and only if* z is preferred to w, in contradiction to the preference pattern we motivated above. Thus second-order induction may result in violations of the version of CBDT presented in Section 4.

Similar examples may be constructed, in which the violation of U-maximization stems from learning that certain *values* of an attribute are similar, rather than that the attribute itself is of importance. That is, instead of learning that the coffee shop is an important factor, Agent may simply learn that coffee shop 1 is similar to coffee shop 2.

One obvious drawback of the functional U which is highlighted here is the fact that it is additively separable across cases. Specifically, second-order induction renders the "weight" of a set of cases a non-additive set function. Since several cases *in conjunction* may implicitly suggest a "rule," the effect of all of them together may exceed the sum of their separate effects. Differently put, the "marginal contribution" of a case to overall preferences depends not only on the case itself, but also on the other cases it is lumped together with. For instance, a utility value of 1 in problem $M1$ has a different effect when coupled with the value 1 in problem $E1$ (as in vector z) than it has when coupled with the same value in $M2$ (as in x).

A possible generalization of additive functionals which may account for this "non-additivity" involves the use of non-additive measures, where aggregation of utility is done by the Choquet integral. (Choquet (1953-4). See also Schmeidler (1989), who introduced this technique to decision making under uncertainty.) However, it should be noted that when second-order induction takes place, it is not only the case-additivity assumption which is being challenged. With similar examples one may convince oneself that the very assumption that preference between acts is determined solely by their "act profiles" may fail in the presence of inductive learning of the similarity function. For instance, consider a matrix as above, where the acts chosen in the first eight problems were neither "cafe latte" nor "cappuccino," but rather two different ones, which yielded the results given by the table. Agent would still draw the same general conclusions about the relative importance of the two attributes of a "problem," and her preference between the latte and the cappuccino would thus depend on all of her memory, including the act profiles of other acts. In particular, second-order inductive reasoning is one plausible example in which case-based preferences do not necessarily satisfy "independence of irrelevant alternatives." That is, the preference between two acts may change when other acts are introduced into the choice set, even if the latter are considered "worse choices" than both of the former.

To sum, CBDT as presented in Section 4 is a rather sketchy "first approximation;" for many problems of interest the formula U should be interpreted so as to allow the similarity function to depend on all of the decision maker's memory.

7. Learning and Expertise

In this section we survey the implications of our model to the process of learning and the definition of the notion of "expertise."

A case-based decision maker learns in two ways: first, by introducing more cases into memory; second, by refining the similarity function based on past experiences. By learning more cases, our decision maker ("Agent") obtains a wider "data base" for future decisions. This process should generally improve her decision making. Of course, the cases learnt may be biased or otherwise misleading; yet, one may expect that, as a rule, and barring computational costs, the knowledge of mcre cases leads to a "better" first-order induction as embodied in case-based decision making.

This improvement of the first-order induction may be viewed as "quantitative." That is, to the extent that CBDT performs implicit first-order induction, it does so even with a meager memory. Thus the introduction of more cases does not change first-order induction in a fundamental or even qualitative way; it does so only quantitatively. (The term "quantitative" may be misleading, since memories are only partially ordered by inclusion; but the addition of cases has a flavor of "more of the same.") On the other hand, second-order induction may be viewed as a *qualitative* improvement of first-order induction. That is, refining the similarity judgment introduces a new dimension to the process of learning. Rather than simply knowing more, it suggests that a better use may be made of the knowledge of the same set of cases.

Knowledge of cases, which we may dub "type I knowledge," is relatively "objective." Though cases may be construed in different ways, there seems to be relatively little room for dispute about them, since they purport to be "facts." By contrast, knowledge of the similarity function, which we refer to as "type II knowledge," is inherently subjective. Correspondingly, it is easier to compare people's type I knowledge than it is to compare type II knowledge. While even knowledge of type I cannot be easily quantified, it does suggest a clear definition of "knowing more," namely, having a memory which is larger (as defined by set inclusion). On the other hand, it is much more difficult to provide a formal definition of "knowing more" in the sense of "having a better similarity function." It seems that what is meant by that is a similarity function which resembles that of an expert, or one which in hindsight can be shown to have performed better in decision making.

The two roles that rules may play in a case-based knowledge representation system correspond to the two types of knowledge, and to the two levels of induction. Specifically, the first role, namely to summarize many cases, may be thought of as succinctly providing knowledge of type I. Correspondingly, only first-order induction is required to formulate such rules: given a similarity function, one simply lumps similar cases together and generates a "rule."[4] By contrast, the second role – drawing attention to similarity among cases – may be viewed as expressing type II knowledge. Indeed, one needs to engage in second-order induction to formulate these rules: it is required that the similarity be *learnt* in order to be able to observe the regularity the rule should express.

These distinctions also allow us to define the concept of "expertise." What is expertise? What makes one an expert in a given field? How can expertise be learnt? Our answer is that expertise also has two aspects. First, being an "expert" in any given field typically involves a rich memory, the acquaintance with many cases, or, in short – knowledge of type I. However, an expert can also do more with the same information. That is, (s)he has a more "accurate" and/or more "sophisticated" similarity function, and in our terminology – possesses "more" (or "better") knowledge of type II.

These distinctions may also have implications for the implementation of computerized systems. A case-based expert system would typically involve both types of knowledge. The discussion above suggests that it makes sense to distinguish between them. For instance, one would like to separate the "hard," "objective" type I knowledge that may be learnt from an expert from the "soft" and "subjective" type II knowledge provided by the same expert. The first is less likely to change than the second. Furthermore, one may wish to use one expert's knowledge of cases with another expert's similarity judgments.

As a final remark, we would like to draw the reader's attention to the fact that even in the presence of second-order induction, case-based knowledge representation incorporates modifications in a "smooth" way. That is, one may sometimes wish to update the similarity values; this may

---

4 Notice, however, that this is first-order *explicit* induction, i.e., a process that generates a general rule, as opposed to the *implicit* induction performed by CBDT.

lead to different decisions based on the same set of cases. But this process does not pose any theoretical difficulties such as those entailed by explicit induction.

8. Concluding Remarks

8.1 We started out by enumerating three major theoretical problems that induction raises. Out of the three, two are solved, or rather obviated, by CBDT: (i) since explicit induction is not done, there is no need to justify it; and (iii) in the absence of unwarranted generalizations, there are no inconsistencies to deal with.

However, problem (ii) disappeared only to re-appear in a different guise: when we avoid induction we need not ask how it is done; but the specification of the similarity function basically poses the same question in the language of CBDT. Indeed, if a similarity function over cases is given, the problem of "How to induct?" or, "What aspects of a given case should be generalized?" becomes much easier to solve.

As mentioned above, we do not purport to provide any general insights into the question of "Similarity – Whence?". From a descriptive point of view, this problem is studied in the psychological literature. (See Tversky (1977), Gick and Holyoak (1980), and Gick and Holyoak (1983), among others.) Taking a normative approach, answers are sometimes given in specific domains in which "cases" are an essential teaching technique (such as law, medicine, and business).

At any rate, we believe that the language of CBDT may also be helpful in dealing with the similarity problem. In particular, second-order induction – as defined in the context of CBDT – may provide some hints regarding the evolution of similarity judgments.

8.2 One may ask, is it indeed the case that all decision-relevant knowledge involves decisions? For instance, suppose that you want to buy a new car, and you know a general rule, by which last year models are offered at discounted prices when new models arrive. This seems extremely relevant knowledge, yet it does not involve any decision.

As we argue in Gilboa and Schmeidler (1992), however, one need not take too narrow a definition of a "case." First, the cases in one's memory need not necessarily be one's own personal experience. Thus, it suffices to have a friend who happened to have bought a new car after the next year models have arrived, to conclude that "wait" may be a smart decision. Secondly, some of the cases one uses may be hypothetical. In fact we argue that any knowledge, in as much as it pertains to the decision to be made, may be translated to a hypothetical case. While this representation raises again the issue of efficiency and parsimony, it still avoids the pitfalls of induction.

8.3 We divide a case into three, the "past," "present," and "future" of the decision. Since the "past" should reflect what is known to the decision maker at the time of decision, it is actually a "subjective" past; that is, the description of a decision problem does not contain facts which are not known to the decision maker, even if they are known to, say, an outside observer.

However, when the decision maker learns of such facts, one faces a choice: to incorporate them into the description of the case, and to refine the similarity judgment accordingly, or to refrain from doing so, in the hope that truly similar cases will be encountered in the future with a similar level of (lack of) information.

While the second approach appears to be more consistent, it may sometimes be "rational" for one to change the description of past problems if one has become aware of certain features one had ignored before. Specifically, if it is likely that in the future these features will be observed in new problems, it makes sense to change one's memory by noting the same features in the previously-encountered ones.

8.4 In the discussion of knowledge representation, our main focus is on that "knowledge" used by people (or machines) in everyday situations. However, one may ask to what extent the case-based model applies to the representation of scientific knowledge, or even mathematical knowledge.

Starting with the latter, we agree with Riesbeck and Schank (1989) that a mathematician's knowledge and reasoning technique is most accurately represented by case-based reasoning. Ideas and solutions are considered "creative" precisely when they cannot be generated algorithmically, that is when the knowledge base they rely on does not contain a wide enough array of obviously-similar cases to induct rules from. Indeed, an idea is "creative" if it does not resemble any known case, or if it relies on original analogies. In other words, creative thinking requires originality either in the (hypothetical) cases considered (corresponding to type I knowledge) or in the similarity function (i.e., type II knowledge).

On the other hand, the *product* of a mathematician's work is almost by definition in the form of rules. Mathematicians are, by and large, interested in producing theorems, or, at most, counterexamples to conjectured ones. Thus the "mathematical knowledge," the accumulation of which is, supposedly, the "aim" of mathematicians, is closer to rules than it is to cases.

When it comes to the sciences, an expert's knowledge is probably best described, as in mathematics, by cases. The product of the scientific work, it would seem, is generally expected to be in the form of rules. However, the degree to which this partly-implicit goal is achieved varies. Generally it appears that in a simple enough environment, which allows for many almost-identical cases to be observed, rules are indeed formulated. But when the environment tends to be unique, scientific knowledge may also take the form of a collection of cases.

CBDT being a scientific theory, one can hardly fail to ask how it applies to itself. Indeed, it does attempt to be a general, rule-style theory, describing decision making and knowledge representation at large. Thus, to the extent that it fails, its failure should be taken as proof of its virtue. (On the other hand, if it happens to be a valid description of reality, it is not refuted; rules which happen to be "true" can be translated back to the collection of cases from which they were derived in the first place.)

8.5 While our main interest is in the theoretical aspects of knowledge representation, we would like to note that case-based models need not be impractical. Indeed, rules appear to be much more efficient than the cases from which they were originally derived. Yet one need not actually program each and every case into memory. For instance, repeated cases may be represented by higher similarity values, thus saving both memory and computation time.

8.6 Throughout the discussion above we confronted cases with "rules." One may argue that, while rules cannot be *known*, they can be *believed* to a certain degree. That is, we may adopt a Bayesian approach, according to which rules are assigned probability values.

Formulating rules to be believed (rather than known) still raises problem (ii), namely, how do induct. Yet it alleviates problems (i) and (iii): since rules are not taken to be certainties, the Bayesian approach is not subject to Hume's criticism. Furthermore, this approach copes with belief-revision quite elegantly as long as no zero-probability event occurs. However, it suffers from other weaknesses, which were the original motivation for the development of CBDT. As these issues are discussed at length in Gilboa and Schmeidler (1992), we ignore them here.

References

Choquet, G. (1953-4), "Theory of Capacities," Annales de l'Institute Fourier, 5, 131-295.

de Finetti, B. (1937), "La Prevision: Ses Lois Logiques, Ses Sources Subjectives." Annales de l'Institute Henri Poincare, 7, 1-68.

Gärdenfors, P. (1990), "Induction, Conceptual Spaces and AI," Philosophy of Science, 57, 78-95.

Gick, M. L. and K. J. Holyoak (1980), "Analogical Problem Solving," Cognitive Psychology 12, 306-355.

Gick, M. L. and K. J. Holyoak (1983), "Schema Induction and Analogical Transfer," Cognitive Psychology 15, 1-38.

Gilboa, I. and D. Schmeidler (1992), "Case-Based Decision Theory," Discussion Paper, Northwestern University.

Hanson, N. R. (1958), Patterns of Discovery. Cambridge, England, Cambridge University Press.

Hume, D. (1748), Enquiry into the Human Understanding. Oxford, Clarendon Press.

Kolodner, J. L., Ed. (1988), Proceedings of the First Case-Based Reasoning Workshop. Los Altos, CA, Morgan Kaufmann Publishers.

Kolodner, J. L. and C. K. Riesbeck (1986), Experience, Memory and Reasoning. Hillsdale, NJ, Lawrence Erlbaum Associates.

Levi, I. (1980), The Enterprise of Knowledge. Cambridge, MA, MIT Press.

McDermott, D. and J. Doyle (1980). "Non-Monotonic Logic I," Artificial Intelligence 25, 41-72.

Moser, P. K., Ed. (1986), Empirical Knowledge. Rowman and Littlefield Publishers.

Quine, W. V. (1953), "Two Dogmas of Empiricism," in From a Logical Point of View. Cambridge, MA, Harvard University Press.

Quine, W. V. (1969), "Epistemology Naturalized," in Ontological Relativity and Other Essays. New-York, Columbia University Press.

Reiter, R. (1980), "A Logic for Default Reasoning," Artificial Intelligence 13, 81-132.

Riesbeck, C. K. and R. C. Schank (1989), Inside Case-Based Reasoning. Hillsdale, NJ, Lawrence Erlbaum Associates, Inc.

Savage, L. J. (1954), The Foundations of Statistics. New York, John Wiley and Sons.

Schank, R. C. (1986), Explanation Patterns: Understanding Mechanically and Creatively. Hillsdale, NJ, Lawrence Erlbaum Associates.

Schmeidler, D. (1989), "Subjective Probability and Expected Utility without Additivity," Econometrica, 57, 571-587.

Sober, E. (1975), Simplicity. Oxford, Clarendon Press.

Tversky, A. (1977), "Features of Similarity." Psychological Review 84, 327-352.

Wittgenstein, L. (1922), Tractatus Logico Philosophicus. London, Routledge and Kegan Paul.