Case-Based Decision Theory and Knowledge Representation

An Outline

by

Itzhak Gilboa

and

David Schmeidler

1. What Is There to Know?

When we address the question of how should we represent knowledge, we should first ask ourselves what is there to know. It seems unarguable that "knowledge representation" should only deal with the representation of that which can be known.

In this note we take the somewhat extreme view that one can never *know* a rule. At least since Hume (1748), it has been recognized that the process of induction fails to be logically justified. Instances of rules – that is, *cases* – may be known, but the rules themselves can at best be conjectured or believed.

It appears that many of the theoretical difficulties associated with knowledge representation are due to the fact that we often purport to "know" the rules obtained by induction. Thus it makes sense to consider a theory of knowledge representation which will obviate the pitfalls of induction.

(This suggestion is related to, and partly inspired by the theory of Case-Based Reasoning proposed by Schank (1986). See Gilboa and Schmeidler (1993b) for additional discussion.)

One may argue that even cases may not be "known", and that the observation of cases is not independent of the rules ("theory") one has in mind.

(See Hanson (1958).) However, a model according to which cases can be "known" in a sense that rules cannot still seems to be a useful "first approximation."

Before we proceed to describe such a model, a comment about beliefs is in order. The Bayesian approach holds that rules, which cannot be *known*, can at least be ascribed some prior probability, to be updated given new evidence. In order to make such probabilities consistent, and in order to be able to update them by Bayes' rule, one actually needs a prior probability measure over a large space of states of the world, which are truth functions defined on all possible observations (and thus also on all possible rules). The Bayesian approach, when applied to such a context, is fraught with conceptual and practical difficulties. We discuss this issue in Gilboa and Schmeidler (1992). As a matter of fact, Case-Based Decision Theory (CBDT), which we propose there and briefly describe below, was developed as an alternative to the Bayesian paradigm. However, we do not dwell on this aspect here, and rather discuss CBDT from the viewpoint of knowledge representation alone.

2. What Is Knowledge Representation Good For?

The notion of a "case" as an object of knowledge seems to provide us with very little structure. Cases are somewhat vague entities. What are they made of? How should they be represented?

The answers to these questions depend on one's goals. We are thus led to pose a more fundamental question, namely, why do we want to represent knowledge in the first place?

We suggest that the main reason to have and represent knowledge is the need to *act*. That is, knowledge is needed for decision making, and it follows that its representation should reflect this use.

Focusing on decision making as our main goal, it is natural to divide a "case" into three parts: (i) the decision problem; (ii) the act that was chosen; and (iii) the result that obtained. These are the "past," "present," and "future" when we set our clocks to the time of decision. We therefore define a "case" to be a triple of problem-act-result.

3. Similarity

Having an idea of what cases are and what they are good for, we now address the question, how are cases to be used in decision making? 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 offers the view that cases are used via analogies. In our context, it is natural to define the similarity on the decision problem, since this is all that is available to the decision maker while making a decision. Thus, having some notion of similarity, we suggest that one may solve decision problems by choosing acts which were successful in *similar* decision problems.

The precise way in which similarity is used is still to be determined. For instance, one natural choice which comes to mind is to employ "nearest neighbor" techniques, that is, to look for the "most similar" case recalled and act accordingly. However, upon closer inspection this technique seems unsatisfactory: suppose that the "most similar" case ended in a catastrophic outcome. What should one do next? Surely one is not tempted to choose the same act again, but there may be many other alternatives to choose from, and the "most similar" case does not provide any help in ranking them.

Furthermore, assume that the most similar case yielded a reasonable outcome, but that in many other cases, potentially "less similar," a different act was chosen, and that in all of them it yielded a very desirable outcome. Wouldn't one like to take these cases into account as well?

It therefore appears that one needs some measure of desirability of outcomes – a *utility* function – which may be aggregated over many cases to obtain an overall ranking of an alternative. This is the main idea of Case-Based Decision Theory (CBDT) as proposed in Gilboa and Schmeidler (1992). We devote the next section to an outline of CBDT.

4. An Overview of CBDT

The primitives of CBDT are:

P – a set of decision ***problems***

A – a set of available ***acts***

R – a set of possible ***results*** (or outcomes)

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 the problem encountered, a is the act chosen by the decision maker, and r is the result that was obtained in this case. We will assume that at any given point in time, a decision maker is equipped with some memory M, which is simply some subset of cases, and which will be interpreted as the set of problems the decision maker can remember.

CBDT postulates two main theoretical terms – "utility" and "similarity." As in classical decision theory, the utility measures the desirability of the results, and is thus a function

$$u: R \rightarrow \Re .$$

The notion of "similarity" corresponds in many ways to that of "subjective probability" in expected utility theory. Similarity measures the extent to which one decision problem is similar to another; that is, it is a function

$$s: P \times P \rightarrow [0,1] .$$

Finally, we may describe the decision rule which is the heart of CBDT: Suppose that a decision maker, characterized by the utility u and the similarity

s is faced a decision problem p, while his/her memory is $M \subseteq C$. Then every possible act $a \in A$ is evaluated by the functional

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

– and the decision maker will, according to CBDT, choose a maximizer of U.

A few comments are in order. First, notice that for two distinct acts $a,b \in A$, $U(a)$ and $U(b)$ are summations over *disjoint* sets of cases. Furthermore, for some acts this summation may be over an empty set, in which case its value is defined to be zero. This value plays a major role in the theory: one may think of it as the decision maker's "aspiration level." To be precise, this is the "default" (utility) value the decision maker seems to be attaching to an act that was never tried in the past (i.e., for which there are no cases in memory). If certain acts obtain higher U-value than zero, the decision maker is "satisficed" (in the sense of Simon (1957) and March and Simon (1958)), and will continue to choose from among them without trying new acts and *without trying to maximize* u. Once all the acts that were tried in the past turned out to be unsatisfactory – that is, to have negative U values – then the decision maker will choose a new act (assuming such exists. If there are several new acts, the choice among them will be arbitrary.)

One of the main features of CBDT is that it does not require the decision maker (DM) to "engage" in hypothetical reasoning: as opposed to expected utility theory, where the very definition of an "act" involves hypothetical statements such as "If state ω occurs then I get r", in CBDT all the DM is required to "know" is the history of cases which *actually happened* and the utility he/she *actually experienced.* (The terms "engage" and "know" above are within quotation marks since one may choose a purely behavioral interpretation of the theory, according to which the DM does not have to "know" or to reason about anything.)

Without details we mention here that the decision rule of CBDT, together with the theoretical terms "utility" and "similarity" may be axiomatically derived from preferences, in a way which parallels the axiomatic derivations of "utility" and "probability", combined with the expected utility formula, in models such as Savage (1954). (See Gilboa and Schmeidler (1992) for one such axiom system, as well as additional discussions.)

The notion of a "case" will sometimes be interpreted in a broader fashion. For instance, a case in a decision maker's memory need not necessarily have been experienced by the same DM. It may well be a "story" told by someone else. Furthermore, it need not be a real case – it may be a hypothetical one, reflecting the DM's knowledge (or belief) about what would have occurred as a result of a possible choice.

Finally, let us briefly mention a few variants of the basic CBDT model:

– *Averaged similarity* in which one uses a functional similar to U above, with the sole difference that for each act a, the similarity coefficients $s(p,q)$ are normalized to sum up to 1;

–*Act Similarity* according to which acts may also be similar to each other, and the evaluation of an act a depends not only on its own performance in the past, but also on that of similar acts. Furthermore, the similarity judgments of acts and of problems may depend on each other. Thus, the similarity function should be defined on problem-act pairs, and an act a is evaluated by

$$U'(a) = U'_{p,M}(a) = \sum_{(q,b,r)\in M} s((p,a),(q,b))u(r)$$

– *Memory-Dependent Similarity* which allows the similarity function to depend on past cases. In particular, the decision maker may "learn" that certain features of problems are more or less similar, or that certain attributes of a "problem" are more or less important, based on past experience. In this case the similarity function is likely to evolve with memory.

5. Is This Knowledge?

Having surveyed the theory, one is tempted to ask, to what extent is this a theory of knowledge representation? Do CBDM's (Case-Based Decision Makers) "know" anything? Do they "know" rules?

Our answer is partly in the affirmative. When rules seem to work, CBDM's behave *as if* they knew them. Following CBDT, a decision maker takes into account all past cases, and if these cases warrant the induction of a rule, such a decision maker would also act in accordance with this rule.

However, the strength of CBDT is demonstrated precisely when rules seem too crude to represent the knowledge base. When induction leads to contradictions between rules or between rules on the one hand and observations on the other – a CBDM may keep using all past cases and make decisions in a coherent way. (See also Gilboa and Schmeidler (1993a) for optimality results.) As opposed to a rule-based decision maker, a case-based one need not undergo any revision of beliefs or major restructuring of his/her knowledge base.

Differently put, CBDT faces no logical problems since cases, as opposed to rules, cannot be contradicted. Avoiding induction in the first place saves the need to retract its unwarranted conclusions.

References

Gilboa, I. and D. Schmeidler (1992). Case-Based Decision Theory. Discussion Paper, Northwestern University. Revised, November 1993.

Gilboa, I. and D. Schmeidler (1993a). Case-Based Optimization. mimeo., Northwestern University.

Gilboa, I. and D. Schmeidler (1993b). Case-Based Knowledge Representation: An Expository Note. mimeo., 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.

March, J. G. and H. A. Simon (1958). Organizations. New York, Wiley.

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.

Simon, H. A. (1957). Models of Man. New York, Wiley.