An Axiomatic Treatment of Three Qualitative Decision Criteria

RONEN I. BRAFMAN

Ben-Gurion University, Beer-Sheva, Israel

AND

MOSHE TENNENHOLTZ

Technion–Israel Institute of Technology, Haifa, Israel

Abstract. The need for computationally efficient decision-making techniques together with the desire to simplify the processes of knowledge acquisition and agent specification have led various researchers in artificial intelligence to examine qualitative decision tools. However, the adequacy of such tools is not clear. This paper investigates the foundations of *maximin*, *minmax regret*, and *competitive ratio*, three central qualitative decision criteria, by characterizing those behaviors that could result from their use. This characterization provides two important insights: (1) under what conditions can we employ an agent model based on these basic qualitative decision criteria, and (2) how "rational" are these decision procedures. For the *competitive ratio* criterion in particular, this latter issue is of central importance to our understanding of current work on on-line algorithms. Our main result is a constructive representation theorem that uses two choice axioms to characterize *maximin*, *minmax regret*, and *competitive ratio*.

Categories and Subject Descriptors: I.2 [**Artificial Intelligence**]

General Terms: Theory

Additional Key Words and Phrases: Agent models, choice theory, foundations of online algorithms, qualitative decision theory

1. *Introduction*

Decision theory plays an important role in fields such as statistics, economics, game-theory, and industrial engineering. More recently, the realization that decision making is one of the central tasks of artificial agents has led to much interest in this area within the artificial intelligence (AI) research community. Some of the more recent work on decision theory in AI concentrates on qualitative decision-making tools. For example, Boutilier [1994] and Tan and

Authors' addresses: R. I. Brafman, Computer Science Department, Ben-Gurion University, P.O.B. 653, Beer-Sheva, 84105 Israel, e-mail: brafman@cs.bgu.ac.il; M. Tennenholtz, Faculty of Industrial Engineering and Management, Technion–Israel Institute of Technology, Haifa 32000, Israel.

Pearl [1994] examine semantics and specification tools for qualitative decision makers, while Darwiche and Goldszmidt [1994] experiment with qualitative probabilistic reasoning in diagnostics; many other contributions to this area appear in Doyle and Thomason [1997]. Such work has two central motivations in mind. First, one hopes that qualitative tools, because of their simplicity, will lead to faster algorithms. Second, qualitative representations may be easier to specify and easier to obtain from experts, leading to a simpler knowledge acquisition process. Indeed, for these reasons AI has often been concerned with qualitative tools and representation techniques.[1]

Research into the foundations of decision theory is motivated by two major applications: agent modeling and decision making. Agent modeling is often the main concern of economists and game-theorists; they ask questions such as: under what assumptions can we model an agent as an expected utility maximizer? In artificial intelligence, we share this concern in various areas, most notably in multi-agent systems, where agents must represent and reason about other agents. Decision making is often the main concern of statisticians, decision analysts, and engineers; they ask, how should we model our state of information, and how should we choose an appropriate action based on this model? The relevance of this question to the design of artificial agents is obvious. The foundational approach helps answer these questions by describing the basic principles that underlie various decision procedures.

One of the most important foundational results in the area of classical decision theory is Savage's theorem [Savage, 1972], described by Kreps [1988] as the "crowning achievement" of choice theory. Savage provides a number of conditions on an agent's preference among actions. Under these conditions the agent's choices can be described as stemming from the use of probabilities to describe her state of information, utilities to describe her preferences over outcomes, and the use of expected utility maximization to choose her actions. Economists use Savage's results to understand the assumptions under which they can use probabilities and utilities as the basis of agent models; decision theorists rely on the intuitiveness of Savage's postulates to justify the use of the expected utility maximization principle. There are numerous other axiomatizations of the principle of expected utility maximization,[2] and much effort has been made to relax the standard representation to deal with various variants of the probabilistic decision-making model, in particular, nonadditive, nontransitive and nonregular preferences over acts [Fishburn 1988].

While the emphasis on the quantitative probabilistic framework is natural given the prominence of probability theory in various scientific disciplines, more effort is needed to understand the basic principles underlying qualitative decision making. The use of qualitative decision tools is likely to come with a price in terms of decision quality, and it is important to develop a better understanding of the basic properties of different qualitative decision making approaches. Our aim in this paper is to initiate work on the foundations of qualitative decision making

[1] For example, work on qualitative notions of knowledge and belief [Halpern 1988; Levesque 1986], belief revision [Gardenfors 1992], nonmonotonic reasoning [Ginsberg 1987], planning [Allen et al. 1990], and qualitative physical models [Forbus 1985].

[2] See, for example, Savage [1972], Anscombe and Aumann [1963], Blum et al. [1991], Kreps [1988], and Hart et al. [1994].

that can help clarify these properties. Our main contribution consists of a number of representation theorems for three qualitative decision criteria: *maximin* [Wald 1950], *minmax regret* and *competitive ratio*. The first two criteria are well known in the decision theory literature [Luce and Raiffa 1957; Milnor 1954] while *competitive ratio* is used in the theoretical computer science literature as the primary optimization measure for on-line algorithms (see, e.g., Borodin and El-Yaniv [1998] and Papadimitriou and Yannakakis [1989]). A central property in all these results is that of *closure under union*: if an agent prefers action a over b given that the possible worlds are s_1 and s_2 and she prefers a over b when the possible worlds are s_3 and s_4, then she still prefers a over b when the possible worlds are s_1, s_2, s_3 and s_4. This condition is strictly stronger than a similar version of Savage's *sure-thing principle* which would require that $\{s_1, s_2\} \cap \{s_3, s_4\} = \emptyset$. The other conditions are more technical, and we defer their presentation to Sections 3 and 4.

The term *qualitative* decision theory is somewhat fuzzy, and within AI one can identify two central formal approaches. The first approach attempts to provide qualitative tools for making decisions that are consistent with classical decision theory. Here, the central idea is to replace probability and utility functions with non-quantitative descriptions of beliefs and preferences that are, nevertheless, consistent with classical decision theory. Special decision making algorithms can be applied to these representations yielding behavior consistent with classical decision theory. For example, Doyle and Wellman [1994] considered formal languages for making ceteris paribus statements. Similarly, Boutilier et al. [1997] use conditional statements of preference to construct a preference relation consistent with the Von Neumann–Morgenstern axioms. However, a complete framework supporting this type of inference has yet to be developed.

A second approach aims to supply qualitative tools for representing and reasoning with beliefs and preferences that do not necessarily conform to the axioms of classical decision theory. Examples of work along these lines include Boutilier's work on the representation of and reasoning with qualitative statements of preference and normality [Boutilier 1994], similar work by Tan and Pearl [1994] on specifying and querying statements of conditional preference, the work of Dubois and Prade [1995] and Dubois et al. [2000] on possibilistic analogous of classical decision theoretic tools, and Lehmann's work on more qualitative versions of Savage's theorem [Lehman 1996]. Our work, too, fits within this approach since it focuses on decision criteria that are not consistent with the principle of expected utility maximization. The three criteria discussed in this paper do not require a quantitative measure of likelihood and their representation of preferences requires integer valued value functions or even a simple preorder relation. Indeed, in their basic form, these criteria completely ignore degree or cumulation of belief, although these can be introduced, to some extent, as we discuss in Section 7.

The main motivation for our study is the foundations of qualitative decision theory, but our results have two additional interesting ramifications. First, the *competitive ratio* decision criterion plays a major role in the analysis of on-line algorithms [Borodin and El-Yaniv 1998; Papadimitriou and Yannakakis 1989]. Hence, our representation theorems for this criterion should be relevant to the foundations of research in that area. Second, our results show how certain agents' behaviors can be represented compactly: There are different ways in which we can encode an agent's behavior (or program). One simple, but space-consuming manner for representing an agent's behavior is as an explicit mapping from the agent's local state to actions. Alternatively, probability and utility functions (or their qualitative counterparts) can be used to implicitly represent certain behaviors if we wish to cut down program storage and program transmission costs. The (constructive) existence theorems in this paper characterize behaviors that can be represented in $O(nm \log(nm))$ space, where n is the number of states of the environment, and m is the number of possible actions. This is to be contrasted with a possibly exponential explicit representation.

In Section 2, we define a model of a situated agent and two alternative representations for its program or behavior. One is a simple policy that maps an agent's state of information to actions, while the other represents the agent's program (or behavior) implicitly using the *maximin* decision criterion. Our aim is to present conditions under which simple policies can be represented implicitly using qualitative decision criteria such as *maximin*. This will be carried out in three steps: In Section 3, we provide an axiomatization for *maximin* in the case of agents that must decide between two actions in various states. In Section 4, we consider agents that choose among an arbitrary, finite set of actions. We show an axiomatization for *maximin* in this more elaborate setting. In Section 5, we show that a policy has a *maximin* representation if and only if it has a *minmax regret* or a *competitive ratio* representation. This immediately implies that the axiomatization of *maximin* applies to *minmax regret* and *competitive ratio* as well. In Section 6, we consider some of the implications of our axiomatization on the adequacy of qualitative decision procedures. As the reader will surely note, our formulation differs from many other standard models in the form given to the utility function and in the role of beliefs. These differences are examined more closely in Section 7. We conclude with a brief summary and discussion of related work in Section 8. Proofs appear in the appendix.

2. *The Basic Model*

In this section, we define a fairly standard agent model. Then, we define the concept of a policy, which describes the agent's behavior, and finally, we suggest a qualitative manner for implicitly representing (some) policies using the concept of value and the *maximin* decision criterion.

Definition 1. Let States be the (finite) set of possible states of the world. An *agent* is a pair of sets, (*LocalStates*, *Actions*), which are called, respectively, the agent's set of *local states* and *actions*.

Let PW: *LocalStates* $\rightarrow 2^{States} \setminus \emptyset$ be a function on *LocalStates* such that (1) $PW(l) = PW(l')$ iff $l = l'$, and (2) For each subset S of *States*, there exists some $l \in$ *LocalStates* such that $PW(l) = S$. We refer to $PW(l)$ as the set of worlds possible in (or consistent with) local state l.

Each state in the set *States* describes one possible state of the *external* world, that is, the agent's environment. This description does not tell us about the internal state of the agent (e.g., the content of its registers) that is described by an element of the set of local states. We view this set as describing the agent's possible states of information, or its *knowledge* (see e.g., Fagin et al. [1995] and Rosenschein [1985]). Hence, the sets *LocalStates* and *States* are disjoint, though closely related through the function $PW(l)$. In addition to a set of possible local states, the agent has a set *Actions* of actions. One can view these actions as the basic control signals the agent can send to its actuators. This concept of action differs from the notion of *act* found in most other work in this area, and we find it more natural for our purpose. We return to this issue in Section 7.

With every local state $l \in LocalState$ we associate a subset $PW(l)$ of *States*, understood as the possible states of the world consistent with the agent's information at l. That is, $s \in PW(l)$ iff the agent can be in local state l when the current state of the world is s. In this paper, we identify l with $PW(l)$ and use both interchangeably. We require that $l = l'$ iff $PW(l) = PW(l')$ and that for every $S \subseteq States$ there exists some $l \in LocalStates$ such that $PW(l) = S$.[3] Notice that the set *LocalStates* does not constitute a partition of the set *States*. Moreover, for any two different local states l, l' there are no constraints on the relationship between $PW(l)$ and $PW(l')$ (except that they must be different).[4]

Like other popular models of decision making (e.g., Savage [1972] and Anscombe and Aumann [1963]), our model considers one-shot decision making. The agent starts at some initial state of information (i.e., some local state) and chooses one of its possible actions. This choice of action is a function of the agent's state of information, as described by the agent's *policy* (also called *protocol* in Fagin et al. [1995] and a *strategy* in the game-theoretic literature [Luce and Raiffa 1957]), which maps each information state to an action.

Definition 2. A *policy* for an agent (*LocalStates*, *Actions*) is a function $\mathcal{P}$: *LocalStates* $\rightarrow$ *Actions*.

A naive description of the policy as an explicit mapping between local states and actions is exponentially large in the number of possible worlds because $|LocalStates| = 2^{|States|}$. Moreover, requiring a designer to supply this mapping explicitly is unrealistic. Hence, a method for implicitly specifying policies is desirable. In particular, we would like a specification method that helps us judge the quality of a policy. Classical decision theory provides one such manner: the policy is implicitly specified using a probability assignment *pr* over the set *States* and a real valued utility function u over a set O of action outcomes. The action to be performed at local state l is obtained using the principle of expected utility maximization:

$$argmax_{a \in Actions} \left\{ \sum_{s \in PW(l)} pr(s) \cdot u(a(s)) \right\},$$

where $a(s)$ is the outcome of action a when the state of the world is s. We wish to present a different, more qualitative representation. We will not use a probability function and our value function $u(\cdot,\cdot)$ takes both the action and the state of the world as its arguments and returns some value in a totally preordered set. (Notice the use of qualitative values rather than quantitative utilities.) For

[3] This assumption is made to simplify presentation. With small modifications, our discussion and results still hold when the set of local states is smaller.

[4] We can define the set of global states of the world, each of which describes a possible joint state of the agent and the external world as the set $\mathcal{G} = \{(l, s): s \in PW(l)\}$. One can identify local states with partitions of $\mathcal{G}$.

convenience, we will use integers to denote the relative positions of elements within this set.

Our choice of a value function dependent on both states and actions differs from the more typical definition of utility functions which do not depend on an action component. However, in this respect, it resembles much of the earlier work on the decision criteria discussed here (e.g., Milnor [1954] and Luce and Raiffa [1957]), which used the state-action formulation. In Section 7.2, we discuss this choice in more detail and sketch an alternative approach based on the action independent formulation.

In our representation, the agent's action in a local state l is defined as:

$$argmax_{a \in Actions}\left\{ \min_{s \in PW(l)} (u(a, s)) \right\}.$$

That is, the agent takes the action whose worst-case value is maximal. This decision criterion is referred to in the literature as *maximin* and it embodies a cautious, risk-averse attitude to decision making.

Definition 3. A policy $\mathcal{P}$ has a *maximin representation* if there exists a value function on *Actions* × *States* such that for every $l \in LocalStates$

$$\mathcal{P}(l) = argmax_{a \in Actions}\left\{ \min_{s \in PW(l)} (u(a, s)) \right\}.$$

That is, $\mathcal{P}$ has a *maximin* representation if there is a value function such that in every local state l, if the agent were to make a decision based on this value function and its state of information using the *maximin* criterion, it would come up with the action $\mathcal{P}(l)$.

Given an arbitrary agent and a policy $\mathcal{P}$ adopted by the agent, it is unclear whether this policy has a *maximin* representation. It is the goal of this paper to characterize the class of policies that have such representations. From this result, we hope to learn about the conditions for modeling agents using the *maximin* decision criterion and to understand the rationality of using this criterion. As it turns out, this will also teach us about the conditions for modeling agents using the *minmax regret* and *competitive ratio* decision criteria and the rationality of using these decision criteria. Besides the primary relevance of these results to the foundations of qualitative decision theory, they can be interpreted as characterizing policies with compact representation. As we remarked, policies can be represented naively using space exponential in *States*. The implicit *maximin* representation provides a much more succinct representation; its size is $O(log(|States| \cdot |Actions|) \cdot |States| \cdot |Actions|)$.

3. *Existence Theorems for Binary Decisions*

This section and the following one present three representation theorems for *maximin*. Here, we examine policies that choose between two actions, and later, we provide more general representation theorems for *maximin*.

A property of protocols that plays a central role in obtaining *maximin* representations is that of *closure under unions*. Policy $\mathcal{P}$ is closed under unions if whenever the same action is chosen in two different local states l, l' in which the agent considers possible the sets of states U and V, respectively, this same action is chosen when the states the agent considers possible are $U \cup V$.

Definition 4. We say that a policy $\mathcal{P}$ is *closed under union* if $\mathcal{P}(U) = \mathcal{P}(W)$ implies $\mathcal{P}(U \cup W) = \mathcal{P}(U)$, where $U, W \subseteq$ *States*.

Example 1. As an example, suppose that our agent is instructed to bring coffee when it knows that the temperature is either cold or ok and when it knows that the temperature is either ok or hot. Hence, if the agent has little information about the temperature, and all it knows is that the weather is either of cold, ok, or hot, it should still bring coffee if its policy is closed under unions. This sounds perfectly reasonable. Consider another example: Alex likes Swiss chocolate, but dislikes all other chocolates. He finds an unmarked chocolate bar and must decide whether or not he should eat it. His policy is such that, if he knows that this chocolate is Swiss or American, he will eat it; if he knows that this bar is Swiss or French, he will eat it as well. If Alex's policy is closed under unions, he will eat this bar even if he knows it must be Swiss, French, or American. If fact, suppose that Alex would eat this bar whenever he knows that it is either Swiss or x, for any other manufacturing country x. If his policy is closed under unions, he will still eat the bar if all he knows is that it must be Swiss, German, French, American, Canadian, or Israeli. Hence, we see that Alex dreads the possibility of not eating Swiss chocolate, and would risk the inferior taste of all other brands rather than miss out on this opportunity.

Consider attempting to find a *maximin* representation for Alex's policy by attempting to fill in the values for the following table:

	Swiss	x	y
eat	v_1	v_3	v_5
dont-eat	v_2	v_4	v_6

Assume that $v_2 < v_1$, $v_3 < v_4$, and $v_5 < v_6$. What determines which action is recommended by the *maximin* criterion when more than one world is possible is which of the lower values is worst, for example, is $v_3 < v_2$ or $v_2 < v_3$? In our case, we know that $v_2 < v_3$ and that $v_2 < v_5$ because Alex would prefer to eat any chocolate as long as it is possible that it is Swiss. Thus, it is easy to see how the closure under union property must follow: v_2 will remain the lowest value in the combined context where we have v_3 and v_5 as well. In general, if some value is worst in one particular context and it is worst in another particular context, it will be worst when we "combine" these contexts. Because, in *maximin*, it is these worst-case values that really matter, we get the closure under union property.

We can show that, in the restricted binary action case, closure under unions is also sufficient to obtain a *maximin* representation because it enforces a preorder relation over the values of the worst-case actions for all states. Hence, our first representation theorem for *maximin* shows that policies containing two possible actions that are closed under unions are maximin representable using a value function defined on *Actions* $\times$ *States*.

THEOREM 1. *Let $\mathcal{P}$ be a policy assigning only one of two possible actions at each local state, and assume that $\mathcal{P}$ is closed under unions and States is finite. Then, $\mathcal{P}$ is maximin representable.*

It is easy to see that every *maximin* representable policy would be closed under union. Hence, we cannot expect a weaker characterization of this decision criterion.

The following example illustrates our result.

Example 2. Consider the following policy (or precondition for wearing a sweater) in which Y stands for "wear a sweater" and N stands for "do not wear a sweater."

	{Cold}	{Ok}	{Hot}	{C,O}	{C,H}	{O,H}	{C,O,H}
sweater	Y	N	N	Y	N	N	N

It is easy to verify that this policy is closed under unions. For example, the sweater is not worn when the weather is ok or when the weather is either hot or cold; hence, it is not worn when there is no information at all (i.e., the weather is either cold, ok, or hot).

Using the proof of Theorem 1, we construct the following value function that represents the policy above:

	cold	ok	hot
Y	3	2	0
N	1	3	3

It is easy to verify that using this value function we obtain the original policy.

The proofs of Theorem 1 and all other theorems appear in the Appendix.

A slight generalization of this theorem allows for policies in which the agent is indifferent between the two available choices. In the two action case discussed here, we capture such indifference by assigning both actions at a local state, for example, $\mathcal{P}(l) = \{a, a'\}$. Hence, we treat the policy as assigning sets of actions rather than actions. We refer to such policies as set-valued policies, or *s-policies*. We redefine the condition of closure under union in this context as follows:

Definition 5. We say that an s-policy $\mathcal{P}$ is closed under unions if for every pair of local states U, $W \subseteq$ *States*, $\mathcal{P}(U \cup W)$ is either $\mathcal{P}(W)$, $\mathcal{P}(U)$, or $\mathcal{P}(U) \cup \mathcal{P}(W)$.

Notice that when $\mathcal{P}(U) = \mathcal{P}(W)$ closure under unions implies, as before, that $\mathcal{P}(U) = \mathcal{P}(U \cup W)$.

We require a number of additional definitions before we can proceed with the representation theorem for s-policies. First, we define the following two binary relationships on subsets of *States*.

Definition 6. Let U, $W \subseteq$ *States*. $U >_{\mathcal{P}} W$ whenever $\mathcal{P}(U \cup W) = \mathcal{P}(U)$ and $\mathcal{P}(U) \neq \mathcal{P}(W)$. $U =_{\mathcal{P}} W$ whenever $\mathcal{P}(U)$, $\mathcal{P}(W)$ and $\mathcal{P}(U \cup W)$ are all different or if $U = W$.

$U >_{\mathcal{P}} W$ tells us that the preferred action in U is preferred in $U \cup W$. $U =_{\mathcal{P}} W$ is basically equivalent to $U \not>_{\mathcal{P}} W$ and $W \not>_{\mathcal{P}} U$.

Again, motivating this definition is the fact that what really matters in *maximin* is worst case values. Suppose that there is a *maximin* representation for $\mathcal{P}$ and that $\mathcal{P}(U) \neq \mathcal{P}(W)$ and $\mathcal{P}(U \cup W) = \mathcal{P}(U)$. It follows that the worst-case value assigned to any state in W (corresponding to the worst action on W) must be lower than the worst-case value assigned to any state in U. Similarly, if $\mathcal{P}(U)$, $\mathcal{P}(W)$ and $\mathcal{P}(U \cup W)$ are all different then in any *maximin* representation of $\mathcal{P}$ we must have that the worst-case values for U and W must be identical. These intuitions give rise to the use of the $<$ and $=$ notations. They also make it clear that some form of relationship resembling transitivity must exist with respect to the $>_{\mathcal{P}}$ and $=_{\mathcal{P}}$ relations if $\mathcal{P}$ is *maximin*-representable.

Definition 7. We say that $>_{\mathcal{P}}$ is *closed under chains* if whenever $U_1 *_1 \cdots *_{k-1} U_k$, where $*_j \in \{>_{\mathcal{P}}, =_{\mathcal{P}}\}$, and $\mathcal{P}(U_1) \neq \mathcal{P}(U_k)$, we have that $U_1 * U_k$. Here, $*$ is $>_{\mathcal{P}}$ if any of the $*_i$ is $>_{\mathcal{P}}$, and $*$ is $=_{\mathcal{P}}$ otherwise.

The condition of being closed under chains helps us ensure that values themselves can be mapped into a transitive binary relation. This will become clearer when we present an extension of this property to arbitrary sets of actions.

Example 3. Suppose there are four possible states of the world s_1, s_2, s_3, s_4. The following (partially specified) policy $\mathcal{P}$ is closed under chains.

	$\{s_1\}$	$\{s_2\}$	$\{s_3\}$	$\{s_4\}$	$\{s_1, s_2\}$	$\{s_2, s_3\}$	$\{s_3, s_4\}$	$\{s_1, s_4\}$
	Y	N	Y	N	Y	{Y,N}	{Y,N}	{Y}

We see that $s_1 >_{\mathcal{P}} s_2$, $s_2 =_{\mathcal{P}} s_3$, $s_3 =_{\mathcal{P}} s_4$, and $s_1 >_{\mathcal{P}} s_4$. Here is one value function that could give rise to such preferences. Notice, in particular, the worst-case values for each state.

	s_1	s_2	s_3	s_4
Y	5	1	7	1
N	0	7	1	8

Finally, we say that $\mathcal{P}$ *respects domination* if the action assigned to the union of a number of sets does not depend on the action assigned to those sets in this union that are dominated by other sets with respect to $>_{\mathcal{P}}$.

Definition 8. We say that $\mathcal{P}$ *respects domination* if, for all W, U, V, $X \subseteq$ *States*, we have that $W >_{\mathcal{P}} U$ and $V >_{\mathcal{P}} X$ implies that $\mathcal{P}(W \cup U \cup V \cup X) = \mathcal{P}(W \cup V)$.

Again, we can see how domination respect arises due to *maximin*'s emphasis on worst-case values. If $W >_{\mathcal{P}} U$, then the worst-case value obtained on W is worse than the worst-case value on U. Similarly, if $V >_{\mathcal{P}} X$, then the worst-case value on V is less than the worst-case value on X. Clearly, for the case of two actions, the worst-case value on $W \cup U \cup V \cup X$ will be the lesser value between the worst case values of W and V. The action on which this value is obtained will determine the action assigned to $W \cup V$. Thus, while closure under unions enforces the needed constraints on the union of pairs of state sets, domination respect enforces the needed constraints on larger unions. Together with the closure under chains property, these conditions enable us to define a preorder on the worst-case values for each state. Consequently, we can state the following representation theorem for s-policies:

THEOREM 2. *Let $\mathcal{P}$ be an s-policy for an agent* (*LocalStates*, *Actions*) *such that* (1) $|Actions| = 2$, (2) *$\mathcal{P}$ is closed under unions*, (3) *$\mathcal{P}$ respects domination*, (4) $>_{\mathcal{P}}$ *is closed under chains*. *Then*, *$\mathcal{P}$ is maximin representable*.

It is readily seen that properties (2)–(4) hold for any policy that is *maximin* representable whenever $|Actions| = 2$.

4. *General Existence Theorems*

The representation theorems we have seen so far have dealt with agents that choose between two actions. Now, we wish to generalize these results to arbitrary (finite) sets of actions. We will assume that, rather than a single most preferred action, the agent has a total order over the set of actions associated with each local state. This total order can be understood as telling us what the agent would do should its first choice became unavailable. The corresponding representation using *maximin* will tell us not only which action is most preferred, but also, which action is preferred to which. That is, the agent will prefer action a to a' in local state l if and only if the worst-case outcome of performing a is better than the worst-case outcome of performing a' when the possible states of the world are those in $PW(l)$.

The intuitions we developed in dealing with the binary choice case will serve us well here. As we will see, the properties we will concentrate on in the general case are natural extensions of the properties developed in the binary case, both of which stem from the central role that worst-case values play in the *maximin* criterion. Our proofs, too, will rely on these similar intuitions, where our aim will be to guarantee conditions that enforce a total preorder on the worst-case values for each possible state of the world.

Definition 9. A *generalized policy* for an agent with local states *LocalStates* and actions *Actions* is a function $\mathcal{P}$: *LocalStates* $\rightarrow$ *TO*(*Actions*), where *TO*(*Actions*) is the set of total orders on *Actions*.

Generalized policy $\mathcal{P}$ is *maximin representable* if there exists a value function $u(\cdot, \cdot)$ on *Actions* $\times$ *States* such that for every pair of actions $a, a' \in$ *Actions* and for every local state $l \in$ *LocalStates*, a is preferred to a' in l iff

$$\min_{s \in PW(l)} (u(a, s)) > \min_{s \in PW(l)} (u(a', s)).$$

When we allow an arbitrary set of actions, the generalization of closure under unions to generalized policies is not a sufficient condition on policies for

s	s'	$s \cup s'$
a'	a	a'
a	a'	a

	s	s'
a	1	3
a'	3	2

FIG. 1. $(a, s) <_{\mathcal{P}} (a', s')$.

obtaining a *maximin* representation. The following definition introduces an additional property needed:

Definition 10. Let $\mathcal{P} = \{>_W | W \subseteq \mathit{States}\}$ be a set of total orders over *Actions*. Given $s, s' \in \mathit{States}$ and $a, a' \in \mathit{Actions}$, we write $(a, s) <_{\mathcal{P}} (a', s')$ if (1) $a' >_s a$, $a >_{s'} a'$, and $a' >_{\{s,s'\}} a$; or (2) $s = s'$ and $a' >_s a$.

We say that $<_{\mathcal{P}}$ is *closed under chains* if whenever $(a_1, s_1) <_{\mathcal{P}} (a_2, s_2) <_{\mathcal{P}} \cdots <_{\mathcal{P}} (a_k, s_k)$ and either (1) $a_k >_{s_1} a_1$ and $a_1 >_{s_k} a_k$ or (2) $s_1 = s_k$, then $(a_1, s_1) <_{\mathcal{P}} (a_k, s_k)$.

As in Theorem 2, the property of closure under chains is required to ensure transitivity of preference among outcomes and hence the possibility of mapping outcomes to an ordered domain.

The left table in Figure 1 helps us clarify this definition. In it, we depict the conditions under which $(a, s) <_{\mathcal{P}} (a', s')$ holds. There are three columns in this table, each showing the agent's preference relation over actions in different local states. The possible worlds in these local states are s, s', and $\{s, s'\}$. In s the agent prefers a' over a, in s' it prefers a over a', but when all the agent knows is that the world is either in state s or s', it prefers a' over a. Roughly, we can say that $(a, s) <_{\mathcal{P}} (a', s')$ if the agent dislikes taking action a in state s more than it dislikes taking action a' in state s'.

We can see how these ideas generalize those presented in the previous section. The $s >_{\mathcal{P}} s'$ relation used in the binary case is really $(a, s) <_{\mathcal{P}} (a', s')$, where a is the less preferred action in s and a' is the less preferred action in s'.

The following example illustrates the closure under chains property:

Example 4. Suppose that there are three possible states of the world: snowing and cold, raining and cold, or warm and neither snowing nor raining. I prefer skiing to walking when it is snowing, but prefer walking to skiing when it is raining. However, when I am uncertain about whether it will rain or snow, I'd choose to walk. In this case (*ski*, *rain*) $<_{\mathcal{P}}$ (*walk*, *snow*). I prefer skiing to jogging when it is warm, and I prefer jogging to skiing when it is raining. However, I really dislike jogging when it is not cold, so I prefer skiing to jogging if I am uncertain whether it is warm or raining. Hence, (*jog*, *warm*) $<_{\mathcal{P}}$ (*ski*, *rain*). Suppose that, in addition, I prefer walking to jogging when it is warm, and I prefer jogging to walking when it snows. The closure under chains condition implies that (*jog*, *warm*) $<_{\mathcal{P}}$ (*walk*, *snow*), and hence I'd prefer walking to jogging if I am uncertain whether it will be warm or it will snow.

What we can see is that closure under chains is closely related to a transitivity of dislikes property: If I dislike skiing in the rain more than I dislike walking in the snow and I dislike jogging in warm weather more than I dislike skiing in the rain, I clearly dislike jogging in the rain more than I dislike walking in the snow. Although this sounds quite plausible, the problematic part of the *maximin* criteria is its exclusive focus on these "dislikes" in making a decision (from which its risk averse flavor stems). This worst-case perspective is enforced by the closure under chains property.

As the following theorem shows, closure under chains together with a natural generalization of closure under unions suffice to guarantee a *maximin* representation when $|Actions| \geq 2$. Indeed, our preference relation over arbitrary sets of actions induces many preferences over binary action choices. We can understand condition (1) below as requiring each of these preference relations to satisfy the closure under unions property.

THEOREM 3. *Let Actions be an arbitrary set of actions, and let* $>_W$, *for every* $W \subseteq States$, *be a total order such that*

(1) *if* $a >_W a'$ *and* $a >_V a'$, *then* $a >_{W \cup V} a'$, *and*
(2) $<_{\mathcal{P}}$ *is closed under chains.*

Then, the generalized policy $\mathcal{P}$ *described by* $\{>_W \mid W \subseteq States\}$ *is maximin representable.*

Finally, in the most general (finite) case, one has an arbitrary set of actions and preferences that are described using a total preorder. That is, the agent can express indifference among a set of actions in a particular local state.

We introduce the following definitions:

Definition 11. A generalized s-policy for agent (*LocalStates*, *Actions*) is a set $\{\preceq_W | W \subseteq States\}$ of total preorders over *Actions*. For each $W \subseteq States$, the preorder $\preceq_U$ describes the preferences of the agent over actions in the local state U.

We write $a \prec_U a'$ when $a \preceq_U a'$ and $a' \npreceq_U a$.

The treatment of generalized s-policies is motivated by the same approach used for s-policies: We extend the ideas used to handle s-policies in the case of binary choice by concentrating on the binary relations that are naturally induced in the more general case; namely, the relation between the value of state-action pairs. Again, the aim is to sufficiently constrain these relations so that a total preorder on state-action pairs can be defined. The value function will simply reflect this total preorder.

The notion of closure under unions for generalized s-policies is closely related to the definition for the binary case, and is defined as follows:

Definition 12. A generalized s-policy $\{\preceq_U | U \subseteq States\}$ is closed under unions if

- $a \preceq_U a'$ and $a \preceq_W a'$ implies $a \preceq_{U \cup W} a'$; and
- $a \prec_U a'$ and $a \prec_W a'$ implies $a \prec_{U \cup W} a'$.

We introduce the following two binary relations on $O = Actions \times States$.

Definition 13. $(a, s) <_{\mathcal{P}} (a', s')$ if

- $a \neq a'$, $s \neq s'$, $a \preceq_s a'$, $a' \preceq_{s'} a$, and $a \prec_{\{s,s'\}} a'$.
- $s = s'$ and $a \prec_s a'$.

$(a, s) =_{\mathcal{P}} (a', s')$ if

- $a \neq a'$, $s \neq s'$, $a \preceq_s a'$, $a' \preceq_{s'} a$, $a' \preceq_{\{s,s'\}} a$, and $a \preceq_{\{s,s'\}} a'$.

- $s = s'$, $a' \preceq_s a$, and $a \preceq_s a'$.

The definition of closure under chains is much as before:

Definition 14. $<_{\mathcal{P}}$ and $=_{\mathcal{P}}$ are *closed under chains* if whenever $(a_1, s_1) *_1 (a_2, s_2) *_2 \cdots *_{k-1} (a_k, s_k)$, where each $*_i$ is either $=_{\mathcal{P}}$ or $<_{\mathcal{P}}$, and either (1) $a_1 \neq a_k$, $s_1 \neq s_k$, $a_1 \preceq_{s_1} a_k$, and $a_k \preceq_{s_k} a_1$; or (2) $s_1 = s_k$; then $(a_1, s_1) * (a_k, s_k)$, where $*$ is $=_{\mathcal{P}}$ if all the $*_i$ are $=_{\mathcal{P}}$, and it is $<_{\mathcal{P}}$ otherwise.

THEOREM 4. *Let Actions be an arbitrary set of actions, and let* $\{\preceq_U \mid U \subseteq$ *States*$\}$, *be a set of total preorders over Actions such that*

(1) $\{\preceq_U | U \subseteq$ *States*$\}$ *is closed under unions, and*
(2) $<_{\mathcal{P}}$ *and* $=_{\mathcal{P}}$ *are closed under chains.*

Then, the generalized policy $\{\preceq_U | U \subseteq$ *States*$\}$ *is maximin representable.*

Again, it is easy to see that a preference order based on *maximin* will have the properties described in this theorem.

As far as offering compact representation to policies, we can understand this theorem as characterizing a class of policies that can be represented using polynomial space. If we measure the size of the representation as a function of $|States| \cdot |Actions| = n$, then we get a *maximin* representation of size $O(n \cdot log(n))$, while the naive representation of a policy is of size $O(2^n)$. In that sense, a similar theorem in the case of classical decision theory would characterize the class of policies that are representable using the maximal expected utility criteria. However, as far as we know, no such characterization exists when the state space is finite. In the infinite case (as in Savage's treatment), space conservation is of little interest.

Unfortunately, in many cases of interest the state space is prohibitively large. Hence any representation linear in $|States|$ is still too large. Today, many AI researchers are interested in structured representation that center on features of states (i.e., state variables) and their values. Here, *influence diagrams* [Howard 1961] are of particular interest because they can be viewed as a compact decision-theoretic representation of probability and utility functions, and consequently, of policies. In many interesting cases, they provide a representation that is logarithmic in the size of $|States| \cdot |Actions|$. A similar approach can be taken in the context of *maximin* representation when for "structured" value functions, for example, when the value is simple function (such as sum) of state features. We are not aware of work along the axiomatic approach that has attempted to give characterization of policies in this manner (either using the maximal expected utility criterion or other criteria) although there is well-known work on structured utility functions (see Luce and Raiffa [1957] and reference therein). It should be noted, though, that as implicit representation of policies, influence diagrams are somewhat problematic because the process of showing that an actual policy is optimal (and hence represented by the diagram) is NP-hard [Cooper 1990; Shimony 1994]. However, we believe that structured representations based on *maximin* will not present this computational difficulty.

5. *Minmax Regret and Competitive Ratio*

Maximin is a classical qualitative decision criterion. In this paper, we also consider two other basic decision criteria, which are defined below.

Definition 15. Given a value function U on *Actions* $\times$ *States*, a state $s \in$ *States*, and an action $a \in$ *Actions*, define $R(a, s) = max_{a' \in A}(U(a', s) - U(a, s))$. In local state l, the *minmax regret* decision criterion selects an action

$$a = \underset{a' \in A}{arg\,min}\left\{ \underset{s \in PW(l)}{max} R(a', s) \right\}.$$

$R(a, s)$ measures the agent's "regret" at having chosen a when the actual state of the world is s. That is, it is the difference in value between the optimal outcome on state s and the outcome resulting from the performance of a in s. *Minmax regret* attempts to minimize the maximal regret value over all possible worlds.

Definition 16. Given a value function U on *Actions* $\times$ *States*, $s \in$ *States*, and $a \in$ *Actions*, define $R(a, s) = \max_{a' \in A}(U(a', s)/U(a, s))$.[5] In local state l, the *competitive ratio* decision criterion selects an action

$$a = \underset{a' \in A}{\arg\min}\left\{ \max_{s \in PW(l)} R(a', s) \right\}.$$

Much like *minmax regret* the *competitive ratio* criterion attempts to optimize behavior relative to the optimal outcome. The only difference is that here we are interested in ratio, rather than difference.

To illustrate these rules, consider the following decision matrix, each action of which would be chosen by a different decision criterion:

	s_1	s_2	chosen by:
a_1	60	10	*minmax regret*
a_2	40	20	*competitive ratio*
a_3	30	21	*maximin*

Notice that *minmax regret* and *competitive ratio* are somewhat more quantitative than *maximin*, since they care about the actual numbers, their difference, or ratio. However, unlike the expected utility criterion, the three criteria discussed in this paper do not require a quantitative measure of likelihood. In addition, it will be evident from the previous representation theorems and the result that follows that we can restrict our attention to integer valued value functions when we use these decision criteria. Finally, notice that all three decision criteria use space polynomial in the number of states and actions to represent the agent's preferences.

[5] For ease of exposition we assume that values are greater than 0; in particular, the division is well defined.

Minmax regret is well known in the decision theory literature [Luce and Raiffa 1957; Milnor 1954]. The *competitive ratio* decision rule is popular in the theoretical computer science literature (e.g., Papadimitriou and Yannakakis [1989]), where it is used as the primary optimization measure for on-line algorithms. Hence, representation theorems that teach us about the conditions under which an agent can be viewed as using each of these decision criteria are interesting both from the AI and theoretical computer science perspectives.

Definition 17. A policy $\mathcal{P}$ is *minmax regret/competitive ratio representable* if there exists a value function $u(\cdot, \cdot)$ on *Actions* $\times$ *States* such that for every pair $a, a' \in$ *Actions* and for every local state $l \in L$ it is the case that a is preferred to a' in l iff

$$\max_{s \in PW(l)} R(a, s) < \max_{s \in PW(l)} R(a', s).$$

Notice that the definitions for the *minmax regret* and the *competitive ratio* representations are similar. The difference stems from the way $R(a, s)$ is defined in these cases. Notice that the value function assigns natural numbers to the elements of *Actions* $\times$ *States*. Given these values the agent applies the *min* and *max* operators to select its favorite actions.

We can prove the following:

THEOREM 5. *A policy is maximin representable iff it is minmax regret/competitive ratio representable*.

Given this result, it follows that the representation theorems and axiomatizations for *maximin* will hold for the two other criteria as well. Interestingly, although this result is not hard to show, it has not been observed before.

6. *Interpreting the Results*

What is the significance of our results? First, they imply that from a modeling perspective, all three decision criteria are similar: Any agent whose choice behavior can be modeled using *maximin*, *minmax regret*, or *competitive ratio* can be modeled using any of the other decision criteria. However, these models will differ in the value function they use. Second and most importantly, our results expose the two fundamental properties of a number of basic qualitative choice criteria: *closure under chains* and *closure under unions*, each of which captures one essential aspect of the criteria we looked at.

Closure under chains is seemingly a restricted form of transitivity requirement on the $<_{\mathcal{P}}$ relation. In fact, it is responsible for the emphasis on worst case performance placed by the three decision criteria in question. This, in turn, leads to their perception as supporting a risk-averse form of reasoning emphasizing safety levels. This emphasis of worst case outcomes follows from closure under chains because, as defined, $(s, a) <_{\mathcal{P}} (s', a')$ assumes that a is worse than a' on s and a' is worse than a on s'. Had we defined the $<_{\mathcal{P}}$ relation by concentrating on the top performers (i.e., assuming a was better on s and a' was better on s') we would have obtained a representation theorem for the *maximax* criterion. This greedy risk-seeking criterion prefers the action that has the best outcome on some of the states.

Closure under unions stipulated that if given a set V of possible worlds the agent prefers action a over a', and given another set W of possible worlds the agent prefers a over a' as well, then it prefers a over a' given $V \cup W$. When V and W are disjoint, we obtain a property analogous to Savage's *sure-thing principle* [Savage 1972]. In this restricted form, this property seems essential when we assume that actions are deterministic and all uncertainty about their effects is modeled as uncertainty about the state of the world (as we do here).[6] Closure under disjoint unions is a basic property of another decision criterion, Laplace's *principle of indifference* in which the action maximizing the sum of values is preferred.

When the sets V and W are not disjoint, closure under unions is a somewhat less natural property of a rational decision maker. Intuitively, it is responsible for the weak, qualitative notion of beliefs that these three decision criteria support. To understand this, we recall the *column duplication* property [Milnor 1954].

Definition 18. A decision criterion has the *column duplication* property if whenever it prefers an action a over an action a' given a set V of possible worlds, it prefers a over a' given any set of possible worlds $V \cup \{s\}$, where $s \notin V$ is such that there exists some $s' \in V$ such that $U(\hat{a}, s) = U(\hat{a}, s')$ for all actions $\hat{a} \in Actions$.

Intuitively, the column duplication property asserts that the agent's preferences do not change if it considers another state possible which is identical, in terms of its effects, to some existing state. Note that column duplication is just the flip side of an equivalent property allowing for removal of one of two states with identical values assigned to outcomes. Column duplication/removal allows us to ignore the cardinality (or weight) of identical columns, and thus, cannot occur where a stronger notion of state plausibility exists.

The following informal observation may be useful in understanding the closure under unions property: *A decision criterion is closed under unions iff it is closed under disjoint unions and has the column duplication property*. A precise formulation of this result appears in the appendix.[7]

It has been observed that column duplication is a basic property of all three decision criteria [Milnor 1954]. For instance, here is the table presented in Section 5 with the column corresponding to s_2 duplicated. As can be seen, the three decision criteria we have discussed choose the same action as before.

	s_1	s_2	s_3	chosen by:
a_1	60	10	10	*minmax regret*
a_2	40	20	20	*competitive ratio*
a_3	30	21	21	*maximin*

[6] However, when "actions" represent multi-step conditional plans, during whose execution the agent's state of information can change, this is no longer true.

[7] The column duplication property is associated with decision tables, for example, as discussed by Milnor [1954] and we have to relate it to our formulation. In particular, we need to define a general notion of a decision criterion and allow more flexible manipulation of states.

Whether or not column duplication is reasonable depends on the state of information of the agent and the conceptualization of the domain. It has been suggested that this property is characteristic of states of complete ignorance [Luce and Raiffa 1957].

In the literature (e.g., Luce and Raiffa [1957]), one finds various examples of counterintuitive choices made by various qualitative criteria in various settings. For instance, one can argue against *maximin* using the following matrix:

	s_1	s_2	$\ldots$	s_{99}	s_{100}
a_1	1	1	1	1	1
a_2	1000	1000	1000	1000	0

Under *maximin*, the first action will be preferred, and this seems counterintuitive. While it is *not* our goal to advocate *maximin*, it is worthwhile to point out a certain known problem with such examples; a problem which lies with the meaning of the numbers used within the decision matrix. If the numbers in the matrix above correspond to dollar amounts, or some objective criteria, such as execution time, then *maximin* may not make much sense.[8] However, in many AI contexts, we are not concerned with monetary payoffs. In that case, one may suppose that the numbers used signify utilities. Yet, the concept of *utility* is meaningless unless it is specified in the context of a decision criterion. For example, the standard notion of utility is *tailored* for expected utility maximizers, and it is somewhat awkward to use it in the context of a *maximin* agent. Of course, once we interpret these values as "utilities" ascribed to a *maximin* agent, this example is no longer counterintuitive.

Finally, we observe that another well-known qualitative decision criterion, Hurwicz's criterion, does not satisfy the property of closure under disjoint unions (although it has the column duplication property). Hurwicz's criterion is the following generalization of *maximin* and *maximax* (*maximax* chooses the action that maximizes the best payoff).

Definition 19. Given a value function U on *Actions* $\times$ *States* and a local state l, the *Hurwicz decision criterion* selects an action a such that

$$a = \arg\max_{a' \in A} \left\{ \left(\alpha \cdot \min_{s \in PW(l)} U(a', s) \right) + \left((1 - \alpha) \cdot \max_{s \in PW(l)} U(a', s) \right) \right\}.$$

[8] Despite such examples, people often make decisions based on worst-case judgments or seek performance guarantees, at least with respect to a set of plausible states (see Section 7.2 for a discussion of this notion). In many domains, it is difficult to make reasonable likelihood estimates and people opt for the more qualitative worst-case analysis. Indeed, the routine use of the competitive ratio criterion in theoretical computer science stems from the difficulty of defining appropriate notions of average case analysis.

When $\alpha = 1$, we obtain the *maximin* criterion, and when $\alpha = 0$, we obtain *maximax*. The following matrix shows that Hurwicz's criterion is not, in general, closed under disjoint unions.

	s_1	s_2	s_3	s_4
a_1	50	50	4	-1
a_2	50	10	1	1

Suppose that $\alpha = 0.5$. Under Hurwicz's criterion, a_1 is preferred over a_2 given either $\{s_1, s_2\}$ or $\{s_3, s_4\}$. However, given $\{s_1, s_2, s_3, s_4\}$, a_2 is preferred over a_1.

7. *Extensions*

Classical decision theory prescribes an agent's behavior using both a utility function and a probability function, while our representation theorems have emphasized value functions. The closest analogue to probabilities, or beliefs, in our model is the agent's state of information. This state of information is described by the agent's local state with which a set of consistent states of the world is associated by means of the function $PW(\cdot)$. We do not distinguish between these states, that is, we do not ascribe different degrees of plausibility to different consistent states. In addition, the value function we use in our representation theorems is defined on state/action pairs, rather than on a separate space of outcomes, as is generally the case in similar representation theorems. Here, we take a closer look at our definition of values, and we show one manner in which qualitative degrees of plausibility can be incorporated into a qualitative representation.

7.1. Defining Values. Most standard treatments of decision theory, such as Savage's, reserve the term *utility function* to real valued functions on a space of outcomes, consisting of a set of possible post-action states of the world. In particular, in Savage's formulation, the notion of an action (or rather, an act) is reserved to a function that maps possible states of the world to outcomes. It is possible for two different acts to result in the same outcome on a given state and on two different states; and it is possible for a single action to have the same outcome on different states. In contrast, we employed the term value to describe a qualitative, for example, integer valued, function on state/action pairs. Moreover, our framework views actions as primitive entities. We believe there is merit in both views. In particular, when we consider an artificial agent, the only physically observable aspect of this agent's behavior is its choice of action, usually in the form of a control signal sent to the agent's actuators. Our representation is identical to the standard representation if it is assumed that the outcomes of different actions on different states are different. Moreover, using value functions that depend on both the state and the action makes practical sense in our qualitative context: it is reasonable when the manner in which the outcome was received is important, for example, the cost of an action, and it allows us to use the value function to encode both the desirability of the action's outcomes and the likelihood of the state in which it is obtained. Finally, our use of this representation has enabled us to require an ordering relation over real actions as opposed to the formulation of Savage and others, which requires an ordering relation over acts, most of which are fictitious, unintuitive entities.

Nevertheless, it is desirable to understand how, for example, a *maximin* representation would be obtained using a framework that employs the concept of an outcome. In what follows, we outline such a result. We assume that there exists a set *Outcomes* of outcomes and that *acts* are functions from *States* to *Outcomes*. We shall use $\mathcal{A}$ to denote the set of all acts. We assume that $\leq$ is a binary relation over $\mathcal{A}$, and we say that f is at least as good as g whenever $g \leq f$ holds. We define $f < g$ and $f = g$ in the standard fashion. Such a preference over the set of acts is *maximin* representable if an ordinal value function u on *Outcomes* exists (in essence, a total preorder on *Outcomes*) such that act f is as preferred as g, written $g \leq f$, iff the worst-case value in f is at least as preferred as the worst-case value in g.

Let us denote by $f[^s_o]$ the act that is identical to f except possibly on $s \in States$, where it assigns o, and let $f(s)$ stand for the act in which $f(s)$ is obtained on all states in *States*. All *maximin* representable preference relations satisfy the following:

Definition 20. $\leq$ *respects domination* if $f(s) \leq f(s')$ implies $f[^{s'}_{f(s)}] = f$.

That is, $\leq$ respects domination if it is indifferent between any act f and a similar act in which s' in assigned $f(s)$ instead of $f(s')$ for some outcome $f(s)$ that is less preferred than $f(s')$.

THEOREM 6. *If $\leq$ is a total preorder on $\mathcal{A}$ that respects domination then $\leq$ is maximin representable.*

The proof of this theorem is left to the interested reader. It is worthwhile to point out that, given a preference relation $\leq$, we cannot define the concept of f being preferable to g given $A \subseteq States$ in a manner analogous to that of Savage [1972]. This stems from the fact that in a preference relation induced by *maximin* it is possible to have acts f, f', g, g' such that (1) f, f' agree on A, (2) g, g' agree on A, (3) f, g agree on $States \backslash A$, (4) f', g' agree on $States \backslash A$, (5) $f < g$, and (6) $f' = g'$. In order to see this, one may consider the values of f, g, f', g' when we have two states s and s'. Let $u(g(s)) = u(g'(s)) = 1$, $u(g(s')) = u(f(s')) = 3$, $u(f(s)) = u(f'(s)) = 2$, and $u(g'(s')) = u(f'(s')) = 1$, then we have that conditions (1)–(6) are satisfied by *maximin*, where $A = \{s\}$. This contradicts one of Savage's nontrivial axioms.

7.2. DEFINING BELIEFS. Quantitative representations of preference usually employ a nontrivial representation of belief, usually in the form of a probability assignment over the set of possible states of the world. Closure under union is the manifestation of what most people would consider unintuitive in the qualitative criteria we discuss—their complete indifference to the likelihood of each of the states. In our current representation, one distinguishes between two types of state: those that are consistent with the agent's local state and those that are inconsistent. However, much work in artificial intelligence employs a richer, though still qualitative, notion of belief, which partitions the set of states that are consistent with the agent's state of information into two sets: the set of *plausible* states and the set of *implausible* states. Finer distinctions are possible, whereby the set of implausible sets is partitioned again, and so on. It is commonly envisioned that the agent makes its decision by considering only the most plausible among its possible worlds. This concept of belief, or plausibility, is easily incorporated into our current model by imposing additional structure on the set *States* in the form of a ranking function. This model has been suggested by, for example, Brafman and Tennenholtz [1994; 1997], Friedman and Halpern [1994], and Lamarre and Shoham [1994]. Given a ranking function r: *States* $\rightarrow N$, we define the agent's beliefs at local state l as:

$$B(l) = \{s \in PW(l) \mid \text{if } s' \in PW(l) \text{ then } r(s) \leq r(s')\}.$$

$B(l)$ are often called the agent's *plausible* states at the local state l. It is natural to modify the decision criterion so that *maximin* is applied only to the plausible states, instead of the possible states (see, e.g., Brafman and Tennenholtz [1994; 1997]). Therefore, at state l the agent chooses the following action:

$$argman_{a \in Actions}\left\{ \min_{s \in B(l)} u(a, s) \right\}.$$

It is natural to ask whether our results can shed some light on this alternative approach to qualitative decision making. Clearly, any behavior that is *maximin* representable can be represented using the ranked *maximin* representation suggested above. (We use a ranking function that maps all states to the same integer.) However, the converse is true as well. That is, if an agent can be represented as using ranked *maximin*, it can also be represented as using the standard *maximin* approach discussed in this paper. The proof of this result is based on the following observations. Given a ranked maximin representation of a policy, where there are m ranks and with maximal value n, we can replace the value $u(a, s)$ of an outcome (a, s) which corresponds to a state s in the mth rank, by $u'(a, s) = m \cdot n + u(a, s)$. It is easy to see that by applying the maximin criterion to u', we obtain the same behavior as with the application of the ranked *maximin* criterion to u. The implications of this observation is that the ranked *maximin* representation is no more expressive than the standard *maximin* representation, that is, it can capture the same set of behaviors. Hence, ranked *maximin* is not, a priori a more rational decision criterion. Similarly we can define ranked *minmax regret* and ranked *competitive ratio* and show that they can represent the same set of policies as *minmax regret* and *competitive ratio* do (respectively).

If we wish to represent preference relations among acts using values assigned to outcomes, as discussed in Section 7.1, the ranked *maximin* representation does provide additional expressivity, allowing for a representation of a larger class of preference relation. One way to see this difference is to observe that, given any permutation π of *States*, we have that $f = \pi(f)$ for any act f whenever $\leq$ is induced by *maximin*. However, when $\leq$ is induced by ranked *maximin* $f = \pi(f)$ only when π is a product of permutations of equally ranked states, that is, permutations π such that for every $s \in$ *States* $r(\pi(s)) = r(s)$. It is possible to obtain postulates that capture this class of preference relations. Two steps must be taken to achieve such postulates: First, we must provide a suitable definition of conditional preference, that is, preference given A for every $A \subseteq$ *States*. As we remarked above, we cannot imitate Savage's definition of conditional preferences when the preference relation is induced by *maximin*. Next, we must uncover the ranking of the set of *States*. That is, we must define a partial preorder on *States* based on the preference relation $\leq$, representing the agent's ranking of the possible states. For example, in order to find out the relationship between s and s', we will use the agent's conditional preferences given $\{s, s'\}$. For example, suppose $u(f(s)) = 0$, $u(f(s')) = 2$, $u(g(s)) = 0$, $u(g(s')) = 1$ and $g < f$, then clearly $r(s) > r(s')$. The interested reader will be able to fill in the details.

8. *Summary and Related Work*

Decision theory is clearly relevant to AI and there is little doubt about the need for decision making techniques that are more designer friendly and have nice computational properties. Qualitative decision theory could offer such an alternative, but the question is: how rational is this approach in different domains? One method of addressing this question is experimentation, as in Darwiche and Goldszmidt [1994]. However, the prominent approach for understanding and justifying the rationality of decision criteria has been the axiomatic approach. This approach yields results that are more general than the experimental approach because they are not domain specific. Using these properties, a designer can assess the desirability of using a particular decision criterion in her domain of application. Our work provides one of a few results within the axiomatic approach that deal with qualitative decision criteria and helps us understand the inherent properties of *maximin*, *minmax regret*, and *competitive ratio*. Our representation theorems are constructive and suggest algorithms for compiling an agent's plan into concise mental-level representation. They show that if we are willing to choose actions that are consistent with the conditions of closure under unions and closure under chains, we can use a simple ordinal value function to represent the agent's preferences.

The reader should be careful not to confuse the work carried out in this paper with the axiomatization of *maximin* and *minmax regret* by Milnor [1954]. Milnor supplies axioms under which an agent would use *maximin* (and axioms under which an agent would use *minmax regret*) to make its move in a game against nature. In Milnor's setting, the values are given and not ascribed to the agent based on his choice among actions. In addition, Milnor does not consider the behavior of the agent in various information states. Our axiomatization provides conditions on the agent's choice of action in various information states under which the agent can be ascribed qualitative values and can be viewed *as if* it adopts the *maximin* criterion to make its decision based on these values. We then show that our axiomatization can be extended to the case where the agent uses *minmax regret* or *competitive ratio* to choose its actions. Therefore, our work is much closer in spirit to the approach taken by Savage.

Following Milnor, there has not been much interest among decision-theorists in the use of qualitative rules such as *maximin* and *minmax regret*, and most work in this area centered around the principle of expected utility maximization. Nevertheless, in recent years, there is an increasing effort by game-theorists to understand various more qualitative settings in which *maximin* plays a prominent role. Of particular note are work by Gilboa and Schmeidler [1989], which attempts to integrate the maximin perspective with an expected utility maximization perspective which has led to similar related efforts. Another related axiomatization, by Hart et al. [1994] axiomatizes *maximin* in the context of 2-person zero-sum games. However, this axiomatization is probabilistic and does not fit the framework of qualitative decision theory as considered in this paper; in their setting the agent's choice among probabilistic lotteries is given, and their *maximin* decision criterion refers to probabilistic strategies.[9] It is worth noting that the notion of safety level, as captured by *maximin* is the basis of standard solution concepts in game-theory, and in particular, that of Nash equilibrium [Fudenberg and Tirole 1991].

An area where the use of a quantitative decision criterion, namely, *competitive ratio* is of paramount important is that of on-line algorithms in Computer Science. There, the need for an adequate axiomatization for this decision criterion has been recognized [Bowdin and El-Yaniv 1998]. Indeed, a Milnor's style axiomatization is mentioned in the literature on on-line algorithms [Borodin and El-Yaniv 1998] and is presented in El-Yaniv [2000]. Our representation theorem is the first to introduce a Savage-style axiomatization (i.e., where utilities are ascribed to the agents) for this basic and widely used decision criterion.

Although researchers in artificial intelligence (and to some extent statisticians) have examined and used various qualitative notions of knowledge, belief (e.g., Doyle [1989], Halpern and Moses [1990], and Brafman and Friedman [1995]) and preference (e.g., Doyle and Wellman [1994]) most effort has been on defining these concepts and understanding the process of (qualitative) belief revision following new information. More recently, ideas on how qualitative notions of belief and preference can be combined have been examined [Tan and Pearl 1994; Boutilier 1994]. However, investigations of the foundations of such decision criteria have not been conducted. A few related works exist: Lehmann [1996] presents an axiomatization of qualitative probabilities, similar to Savage's [1976] where preferences among acts are only partially ordered. Dubois and Prade [1995] present a qualitative axiomatization of utilities, similar to von Neumann and Morgenstern's [1944] axiomatization of the principle of expected utility maximization. Rather than use probabilities, Dubois and Prade use possibility measures [Zadeh 1978] to define a qualitative notion of mixtures. Roughly speaking, in this representation, sum over utilities is replaced by the *min* operator (due to a risk averse attitude embodied in the axioms) while multiplication is replaced by a max operator. Besides the technical differences with our work, much like von Neumann and Morgenstern use of probabilities, Dubois and Prade assume a predefined concept of belief in the form of a possibility distribution over the states of the world. A recent extension of this line of work [Dubois et al. 2000] presents a Savage-style axiomatization that is closer in spirit to the approach presented here.

A number of natural extensions remain for future work, most prominently, the generalization of our results to infinite state spaces. At this point, it is not clear to us whether a straightforward generalization exists. In addition, an analysis of additional qualitative representations of preference is needed.

[9] Recent work by Gilboa and Schmeidler [1995] examines case-based decision making, which should be of interest to researchers in qualitative decision theory.

Finally, the state space representation used in this paper are still too large in most applications of interest. An important question for future work is characterizing the class of behaviors that can be encoded compactly using a more structured representation of states (e.g., as truth assignments to some set of propositions).

Appendix A. Proofs

PROOF OF THEOREM 1. We will use the following definition in the proof:

Definition 21. $U >_{\mathcal{P}} W$, where $U, W \subseteq$ *States*, if $\mathcal{P}(U \cup W) = \mathcal{P}(U)$ and $\mathcal{P}(U) \neq \mathcal{P}(W)$.

CLAIM 1. *If* $U_1 >_{\mathcal{P}} \cdots >_{\mathcal{P}} U_k$ *and* $\mathcal{P}(U_1) \neq \mathcal{P}(U_k)$, *then* $U_1 >_{\mathcal{P}} U_k$. *We will refer to this property as closure under chains.*

PROOF. First, recall that we are assuming only two possible actions. Suppose that $U >_{\mathcal{P}} V$, $V >_{\mathcal{P}} W$, $W >_{\mathcal{P}} X$, and that $\mathcal{P}(U) = a$. We must show that $\mathcal{P}(U \cup X) = a$ as well. Suppose not. In that case, we have that $\mathcal{P}(U \cup X) = a'$ and that $\mathcal{P}(V \cup W) = a'$ (since $V >_{\mathcal{P}} W$ and $\mathcal{P}(V)$ must be a' because $U >_{\mathcal{P}} V$). Hence, since $\mathcal{P}$ is closed under unions, we have that $\mathcal{P}(U \cup V \cup W \cup X) = \mathcal{P}((U \cup X) \cup (V \cup W)) = a'$. However, $\mathcal{P}(U \cup V) = a$ and $\mathcal{P}(W \cup X) = a$ (because $U >_{\mathcal{P}} V$ and $W >_{\mathcal{P}} X$), which implies, using closure under unions, that $\mathcal{P}(U \cup V \cup W \cup X) = a$, a contradiction. The same proof works for an arbitrary chain of the form $U_1 >_{\mathcal{P}} \cdots >_{\mathcal{P}} U_k$.

As we explain in the body of the paper, the main idea behind the various constructions we use is to impose a total preorder on the worst-case values for each state, from which we can derive a *maximin* representation. We now show how the value function is constructed and prove that this value function represents the policy.

(1) Set i to 0, and S to *States*.
(2) For each state $v \in$ *States*, let $u(\mathcal{P}(v), v) = n$. (We let the value of the preferred action in v be $n = |States|$).
(3) Let $U = \{s \in S \mid \text{there is no } v \in S \text{ such that } v >_{\mathcal{P}} s\}$. (Among the states to which the worst-case action has not been assigned a value, we now find those that are the worst.)
(4) Let $u(a, s) = i$ for every $s \in U$ and a such that $a \neq \mathcal{P}(s)$.
(5) Let $i := i + 1$ and $S := S \backslash U$.
(6) If $S = \emptyset$, then terminate; otherwise, goto (3).

First, notice that the algorithm terminates because each time Step (3) is taken, the set U is nonempty. If this were not the case, then, at Step (3), for every $v \in S$, there would be some v' in S such that $v' >_{\mathcal{P}} v$. But, because closure under chains and the fact that *States* is finite, this would imply the (impossible) existence of two distinct states v, v' such that $v' >_{\mathcal{P}} v$ and $v >_{\mathcal{P}} v'$.

CLAIM 2. *Given the value function generated above, the maximin criteria yields a unique most preferred action for every local state* W.

PROOF. Both actions can be equally preferred only if they have equal worst-case outcomes. This could be the case if the set U constructed in Step (3) at some particular iteration contains two states, s and v, on which $\mathcal{P}$ assigns a different action. However, if that was the case then either $s >_{\mathcal{P}} v$ or $v >_{\mathcal{P}} s$, depending on the action assigned on $\{s, v\}$, and so both cannot appear in U simultaneously.

CLAIM 3. *By using maximin, we obtain behavior identical to that of the original policy.*

First, notice that this claim holds when the decision is made under certainty, that is, when $|PW(l)| = 1$. This is immediate from our construction process. Let $Actions = \{a, a'\}$ and suppose that *maximin* chooses a on U. By definition of *maximin*, it must be the case that for some $w \in U$, $u(a', w)$ is the worst value possibly obtained on U. Let A_a be those U worlds on which a is preferred and let $A_{a'}$ be those U worlds on which a' is preferred. We claim that if $v \in A_{a'}$ then $w >_{\mathcal{P}} v$. To see this, consider our construction process. Since $u(a', w)$ is minimal, w must have been chosen in Step (3) before all other elements of U. Hence, we have that, for no $v \in A_{a'}$, is it the case that $v >_{\mathcal{P}} w$. However, since the action taken on elements of A'_a and on w is different, we must have that $v \in A_{a'}$ implies $w >_{\mathcal{P}} v$. (Because for every pair v, v' of states that are assigned a different action by $\mathcal{P}$ either $v >_{\mathcal{P}} v'$ or $v' >_{\mathcal{P}} v$.) This in turn implies that on each local state of the form $\{v, w\}$, where $v \in A_{a'}$, the action taken is a. Since $\mathcal{P}$ is closed under unions, we have that a is taken on $A_{a'} \cup w$. Since, by definition, a is taken on A_a then by the fact that $\mathcal{P}$ is closed under unions, we have that a is taken on $A_a \cup A_{a'} \cup \{w\} = U$.

This concludes the proof.

PROOF OF THEOREM 2. We construct the value function much like in Theorem 1.

(1) Set i to 0, and S to *States*.
(2) Let $U = \{s \in S \mid \text{there are no } v, w \in S \text{ such that } w =_{\mathcal{P}} s \text{ and } v >_{\mathcal{P}} w\}$.
(3) Let $u(a, s) = i$ for every $s \in U$ and a such that $a \notin \mathcal{P}(s)$.
(4) Let $u(a, s) = i$ for every $s \in U$ and a such that $\mathcal{P}(s) = \{a, a'\}$.
(5) Let $i := i + 1$ and $S := S \backslash U$.
(6) If $S = \emptyset$, then let u assign i in all other cases and terminate; otherwise, goto (2).

First, notice that the algorithm terminates because each time Step (2) is taken, the set U is nonempty. If this were not the case, then, at Step (2), for every $v \in S$ there would be some $v', v'' \in S$ such that $v =_{\mathcal{P}} v'$ and $v'' >_{\mathcal{P}} v'$. Because S is finite, this implies that there are some states $s_1, \ldots, s_j \in S$ such that $s_1 =_{\mathcal{P}} s_2 >_{\mathcal{P}} s_3 >_{\mathcal{P}} \cdots >_{\mathcal{P}} s_{j-1} =_{\mathcal{P}} s_j >_{\mathcal{P}} s_1$. By closure under chains, we deduce that $s_1 >_{\mathcal{P}} s_j$ and $s_j >_{\mathcal{P}} s_1$, which is impossible.

CLAIM 4. *Suppose that $\mathcal{P}(w) = a$, $a' \in \mathcal{P}(v)$ and $u(a', w) < u(a, v)$ for some $w, v \in States$. Then, $w >_{\mathcal{P}} v$.*

PROOF. Because of our construction process, $a' \in \mathcal{P}(v)$ implies that $u(a, v) \leq u(a', v)$.

Because $u(a', w) < u(a, v)$, we know that w was chosen before v during the construction of $u(\cdot)$. We know that $\mathcal{P}(w) \neq \mathcal{P}(v)$. There are three options: (1) $\mathcal{P}(w) = \mathcal{P}(\{v, w\})$ which implies, by definition, that $w >_{\mathcal{P}} v$. (2) $\mathcal{P}(v) = \mathcal{P}(\{v, w\})$ which implies, by definition, that $v >_{\mathcal{P}} w$. However, in that case, v would have been chosen before w in the above algorithm. In that case, we would have had $u(a, v) > u(a', w)$, which contradicts our assumptions. (3) $\mathcal{P}(\{v, w\})$ is different from $\mathcal{P}(w)$ and $\mathcal{P}(v)$. Since we know that $\mathcal{P}(v)$ and $\mathcal{P}(w)$ are different, this implies, by definition, that $v =_{\mathcal{P}} w$. Let us consider this possibility. We know that $u(a, v) > u(a', w)$. From the algorithm above, we see that this implies that when w was chosen as part of the set U, v was still unassigned and was not chosen as part of the set U. Hence, at that time, there was some unassigned $s, s' \in$ *States* such that $v =_{\mathcal{P}} s$ and $s' >_{\mathcal{P}} s$. Therefore, we have that $s' >_{\mathcal{P}} s =_{\mathcal{P}} v =_{\mathcal{P}} w$. If $s' >_{\mathcal{P}} v$ as well, then from the construction process, we know that w could not have been chosen at that stage. (Since s', v satisfy the conditions in Step (2) for not including w in U.) Otherwise, it follows from closure under chains that $\mathcal{P}(s') = \mathcal{P}(v)$. Because $\mathcal{P}(w) \neq \mathcal{P}(v)$ we get that $\mathcal{P}(s') \neq \mathcal{P}(w)$. Therefore, by closure under chains we have $\mathcal{P}(s') >_{\mathcal{P}} \mathcal{P}(w)$. Consequently, w could not have been chosen as part of U at this stage, contradicting the fact that it was chosen prior to v. In conclusion, we have seen that only option (1) is consistent with our assumptions, and thus, $w >_{\mathcal{P}} v$ must hold.

CLAIM 5. *Suppose that* $\mathcal{P}(w) = a$, $\mathcal{P}(v) = a'$ *and* $u(a', w) = u(a, v)$ *for some* $w \neq v \in$ *States. Then,* $w =_{\mathcal{P}} v$.

PROOF. Because $u(a', w) = u(a, v)$, we know that both w and v were chosen at the same stage during the construction of $u(\cdot)$. Hence, neither $w >_{\mathcal{P}} v$ nor $v >_{\mathcal{P}} w$ hold. This can only be the case if $\mathcal{P}(\{w, v\}) = \{a, a'\}$, which implies $w =_{\mathcal{P}} v$.

CLAIM 6. *By using maximin on the value function obtained above, we obtain behavior identical to that of the original s-policy.*

PROOF. We prove this claim by induction on the size of the set of possible worlds $|W|$. As before, we first notice that the construction process guarantees this claim will hold whenever $|PW(l)| = 1$. Next, let *Actions* $= \{a, a'\}$ and suppose that *maximin* chooses x on W. We will show that for every possible choice of x, this claim holds.

(1) Suppose that $x = a$. By definition of *maximin*, there must be some state $w \in W$ such that for every $w' \in W$, we have that $u(a', w) < u(a, w')$. Because the claim is valid for singletons, we know that $\mathcal{P}(w) = a$. Let $A_a = \{s \in W \mid \mathcal{P}(s) = a\}$ and let $A_{a'} = \{s \in W \mid a' \in \mathcal{P}(s)\}$. By Claim 4, we know that $w >_{\mathcal{P}} v$ for all $v \in A_{a'}$. Consequently, $\mathcal{P}(\{w, v\}) = \mathcal{P}(w) = a$. Using closure under unions, we conclude that $\mathcal{P}(w \cup A_{a'}) = a$. Another sequence of applications of closure under unions implies that $\mathcal{P}(w \cup A_{a'} \cup A_a) = a$. However, $W = w \cup A_{a'} \cup A_a$.

(2) If $x = a'$, the same argument works.

(3) Suppose $x = \{a, a'\}$. There are two subcases to consider:

(a) There is some $w \in W$ such that $u(a, w)$ and $u(a', w)$ are the minimal values assigned to an action and a state in W. We notice that because of closure under unions, when $\mathcal{P}(U) = \{a, a'\}$, then for every $V \subseteq States$, $\mathcal{P}(U \cup V) = \mathcal{P}(U)$ or $\mathcal{P}(U \cup V) = \mathcal{P}(V)$. Hence, for every $v \in W$, either $w >_{\mathcal{P}} v$, $v >_{\mathcal{P}} w$ or $\mathcal{P}(w) = \mathcal{P}(v)$; but because of our construction process, $v \not>_{\mathcal{P}} w$. We conclude that $w >_{\mathcal{P}} v$ for all $v \in W$ such that $\mathcal{P}(w) \neq \mathcal{P}(v)$. In turn, this means that in all local states $\{\{w, v\} \mid v \in W, \mathcal{P}(v) \neq \{a, a'\}\}$ the s-policy $\mathcal{P}$ assigns $\{a, a'\}$. We conclude using closure under unions.

(b) There are a number of distinct states $w_1, \ldots, w_m \in W$ and actions $a_1, \ldots, a_m \in Actions$ such that $u(a_1, w_1) = \cdots = u(a_m, w_m)$ and for any other pair $(b, s) \in Actions \times W$ it is the case that $u(a_1, w_1) < u(b, s)$. Without loss of generality, we will assume no state in W is assigned $\{a, a'\}$ by $\mathcal{P}$. We can always use closure under unions to add these states later on.

We consider two cases: (i) $w_1 \cup \cdots \cup w_m = W$ in which case we conclude by invoking Claim 5 (which implies that for any two states w_i, w_j such that $1 \leq i, j \leq m$ $\mathcal{P}(\{w_i, w_j\}) = \{a, a'\}$) and closure under unions. (ii) $W \backslash (w_1 \cup \cdots \cup w_m) = W' \neq \emptyset$. Let us denote by W_a (respectively, $W_{a'}$) the set of state in W' in which $\mathcal{P}$ assigns a (respectively, a'). Consider any w_i in which $\mathcal{P}$ assigns a and $w \in W_{a'}$. Using Claim 4, we obtain that $\mathcal{P}(\{w, w_i\}) = \mathcal{P}(w_i)$, and using closure under union, we obtain that $\mathcal{P}(w \cup W_{a'}) = \mathcal{P}(w_i)$. Hence, $w_i >_{\mathcal{P}} W_{a'}$. Similarly, if $\mathcal{P}(w_j) = a'$, then we have $w_j >_{\mathcal{P}} W_a$. Because $\mathcal{P}$ respects domination, we have that $\mathcal{P}(w_i \cup w_j \cup W_a \cup W_{a'}) = \mathcal{P}(w_i \cup w_j)$, which, by Claim 5, must equal $\{a, a'\}$. Using a number of applications of closure under union, we obtain that $\mathcal{P}(W) = \{a, a'\}$ as desired.

This concludes the proof.

PROOF OF THEOREM 3. Let $O = Actions \times States$. Hence, u is a function of O. We construct the value function u over O using the following algorithm, with the aid of a Boolean function over O, DONE, and an integer-valued function over O, LIM. Notice that $<_{\mathcal{P}}$ is defined on O.

(1) On all elements of O, initialize u and LIM to 0 and DONE to `false`.
(2) While $\exists o' \in O$ such that DONE(o') = `false` do
   (a) Find some $o \in O$ such that
      (i) DONE(o) = `false`;
      (ii) For no other outcome $o' \in O$ such that DONE(o') = `false` is it the case that $o' <_{\mathcal{P}} o$.
   (b) Let $u(o)$ = LIM(o) and DONE(o) = `true`.
   (c) If $\exists o''$ is such that DONE(o'') = `false` and $o'' > o$, then let LIM(o'') = $u(o) + 1$.

Notice that it could be the case that neither $o <_{\mathcal{P}} o'$ nor $o' <_{\mathcal{P}} o$.

First, we must show that the above procedure terminates, that is, that we will always find an outcome o satisfying all conditions within the WHILE loop. This follows from our transitivity requirement. Suppose that for every o for which DONE(o) = `false` there is some o' for which DONE(o') = `false` such that $o' <_{\mathcal{P}} o$. Because O is finite, it must be the case that there are some $o_1, \ldots, o_m \in O$ for all of which Done($\cdot$) returns `false` such that $o_1 <_{\mathcal{P}} o_2 <_{\mathcal{P}} \cdots <_{\mathcal{P}} o_m <_{\mathcal{P}} o_1$. Hence, both $o_1 <_{\mathcal{P}} o_m$ and $o_m <_{\mathcal{P}} o_1$. However, an examination of the definition of $<_{\mathcal{P}}$ reveals that such a case is impossible.

Next, denote the preference order obtained by using *maximin* on the values obtained through the above construction by $>^m$. In order to complete the proof, it suffices to prove the following claim.

CLAIM 7. *$>$ and $>^m$ are identical.*

PROOF. First, notice that when restricted to single states, this claim is true. That is, under certainty, both preference relations are identical.

Next, suppose that $a >^m_U a'$. This implies that there is some state $w \in U$ such that $u(a', w)$ is the minimal value obtained by u in $\{a, a'\} \times U$. We claim that for every $s \in A_{a'} = \{v \in U \mid a' >_v a\}$ it is the case that $a >_{\{s,w\}} a'$. Suppose not. Then $a' >_{\{s,w\}} a$ for some such s. In that case, we have $a >_w a'$, $a' >_s a$ and $a' >_{\{s,w\}} a$. Hence, $(a', w) >_{\mathcal{P}} (a, s)$, and by our construction process, $u(a', w) > u(a, s)$. But this contradicts our initial assumption. Therefore, we must conclude that $a >_{\{s,w\}} a'$. Using closure under unions, we get that $a >_{\{w \cup A_{a'}\}} a'$. Since on all states in $U \backslash \{w \cup A_{a'}\}$ a is preferred over a', we conclude that $a >_U a'$ with one more application of closure under unions.

For the other direction, suppose that $a \not>^m_U a'$. If $a' >^m_U a$, then we proceed as above to show that $a' >_U a$ and, consequently, $a \not>_U a'$. If $a' \not>^m_U a$ as well, then there are some $w, w' \in U$ (where possibly $w = w'$) such that $u(a, w) = u(a', w')$ and these constitute the minimal values u obtains on $\{a, a'\} \times U$. From our construction process, it is clear that neither $(a, w) >_{\mathcal{P}} (a', w')$ nor $(a', w') >_{\mathcal{P}} (a, w)$, for otherwise step (2)(c) would have ensured that $u(a, w) > u(a', w')$ or $u(a, w) < u(a', w')$, contrary to our assumption that $u(a, w) = u(a', w')$. Therefore, we conclude that w and w' must be different (or else, one of the above relations must hold). We shall assume that a and a' are different as well, since we are not interested in the case $a = a'$.

We also know that $a' >_w a$ (using our initial remark about the validity of our claim when restricted to single states and the fact that $u(a, w)$ is minimal) and that $a >_{w'} a'$ (for the same reasons). It must be the case that either $a >_{\{w,w'\}} a'$ or $a' >_{\{w,w'\}} a$, since these are total orders. Hence, we conclude that $u(a, w) > u(a', w')$ or $u(a, w) < u(a', w')$, contradicting our earlier conclusion that neither of these hold. This in turn, contradicts the initial assumption that $a' \not>^m_U a$.

This concludes our proof.

PROOF OF THEOREM 4. We define $u(a, s)$ as follows:

(1) Mark all $o \in O$ unassigned and initialize c to 0.
(2) Choose all unassigned $o \in O$ such that there are no other unassigned o', $o_1, \ldots, o_k \in O$ (for $k \geq 0$) such that $o =_{\mathcal{P}} o_1 =_{\mathcal{P}} \cdots =_{\mathcal{P}} o_k <_{\mathcal{P}} o'$.
(3) Let $u(o) = c$ for all o chosen above and mark o assigned.
(4) Increment c by one and goto (2) if there are still $o \in O$ that are unassigned.

CLAIM 8. *The above algorithm terminates.*

PROOF. The algorithm will not terminate only if at some stage there are some unassigned elements of O, such that for each such unassigned o, there exists some unassigned $o', o_1, \ldots, o_k \in O$ such that $o =_{\mathcal{P}} o_1 =_{\mathcal{P}} \cdots =_{\mathcal{P}} o_k <_{\mathcal{P}} o'$. Since O is finite, this implies that there are unassigned $o_1, \ldots, o_m \in O$ such that $o_1 \star_1 o_2 \star_2 o_3 \star_3 \cdots \star_{k-1} o_k \star_k o_1$, where $\star_i \in \{>_{\mathcal{P}}, =_{\mathcal{P}}\}$ and for some $1 \leq j \leq k$, it is the case that $\star_j = >_{\mathcal{P}}$. In particular, based on the closure under chains property, there must be some $1 \leq l, n \leq k$ such that both $o_l >_{\mathcal{P}} o_n$ and $o_n >_{\mathcal{P}} o_l$, which is impossible.

CLAIM 9. *If* $(a, s) =_{\mathcal{P}} (a', s')$, *then* $u(a, s) = u(a', s')$.

PROOF. Suppose (a, s) is not chosen in Step (2) at some iteration. Hence, there are some other unassigned $o', o_1, \ldots, o_k \in O$ such that $(a, s) =_{\mathcal{P}} o_1 =_{\mathcal{P}} \cdots =_{\mathcal{P}} o_k <_{\mathcal{P}} o'$. But $(a, s) =_{\mathcal{P}} (a', s')$ so $(a', s') =_{\mathcal{P}} (a, s) =_{\mathcal{P}} o_1 =_{\mathcal{P}} \cdots =_{\mathcal{P}} o_k <_{\mathcal{P}} o'$. Hence, (a', s') could not have been chosen at Step (2) at this iteration either.

CLAIM 10. *Suppose that* $a \neq a'$; *then*

(1) $u(a, s) \leq u(a', s')$ *and* $u(a, s) \leq u(a', s)$ *implies that* $a \preceq_{\{s,s'\}} a'$;
(2) $u(a, s) < u(a', s')$ *and* $u(a, s) < u(a', s)$ *implies that* $a \prec_{\{s,s'\}} a'$.

PROOF. First, suppose that $s = s'$. If the consequence of (1) is false, then $a' \prec_s a$ holds which would ensure $u(a', s) < u(a, s)$. For (2), first, notice that using (1) we conclude that $a \preceq_{\{s,s'\}} a'$. Next, if the consequence of (2) does not hold, we know that $a' \preceq_{\{s,s'\}} a$. Therefore, $(a, s) =_{\mathcal{P}} (a', s)$. Based on Claim 9 we have that $u(a, s) = u(a', s)$, which contradicts the fact that $u(a, s) < u(a', s)$.

Now we proceed to prove Claim 10 when $s \neq s'$. Notice that from the above case, we know that $u(a, s) \leq u(a', s)$ implies that $a \preceq_s a'$ and $u(a, s) < u(a', s)$ implies that $a \prec_s a'$. We now derive (1) as follows:

Suppose that $u(a, s) \leq u(a', s')$ and assume that $a \npreceq_{\{s,s'\}} a'$. This implies that $a' \prec_{\{s,s'\}} a$. If $a' \preceq_{s'} a$, then together with $a' \prec_{\{s,s'\}} a$ and $a \preceq_s a'$, we conclude that $(a', s') <_{\mathcal{P}} (a, s)$. Hence, (a', s') must have been assigned value earlier than (a, s), which would mean that $u(a, s) > u(a', s')$ contradicting our initial assumption that $u(a, s) \leq u(a', s')$. Hence, we must have that $a \preceq_{s'} a'$. However, $a \preceq_{s'} a'$ and $a \preceq_s a'$ implies (1) using closure under under unions.

We derive (2) as follows: from $u(a, s) < u(a', s)$ we get $a \prec_s a'$ (based on the $s = s'$ case above). If $a \prec_{s'} a'$ holds, we get the desired result using closure under unions. Otherwise, $a' \preceq_{s'} a$ holds. From (1), we know that $a \preceq_{\{s,s'\}} a'$ holds as well. If the desired conclusion is false, then $a' \preceq_{\{s,s'\}} a$ and consequently, $(a, s) =_{\mathcal{P}} (a', s')$. However, based on Claim 9, this implies $u(a, s) = u(a', s')$, which contradicts the fact that $u(a, s) < u(a', s')$.

We must now show that $a \preceq^m_W a'$ iff $a \preceq_W a'$.

For the first direction, suppose that $a \preceq^m_W a'$ and let us show that $a \preceq_W a'$. Since $a \preceq^m_W a'$, we know that there must exist some $w \in W$ such that $u(a, w) \leq u(a', v)$ for all $v \in W$. Using Claim 10, we know that $a \preceq_{\{v,w\}} a'$ for all $v \in V$ that differ from w. Using closure under unions we conclude that $a \preceq_W a'$.

For the other direction, suppose that $a \preceq_W a'$ and let us show that $a \preceq^m_W a'$. Let $w \in W$ be such that for all $v \in W$, it is the case that $u(a, w) \leq u(a, v)$. We must show that for all $v \in W$ it is also the case that $u(a, w) \leq u(a', v)$. Suppose not. Then, there exists some $v \in W$ such that $u(a', v) < u(a, w)$. Hence, for every $v' \in W$, it is the case that $u(a', v) < u(a, v')$. Therefore, by Claim 10, we know that for all $v' \in W$, $a' \prec_{v,v'} a$. Using closure under unions, we obtain that $a' \prec_W a$, contradicting our initial assumption.

This concludes our proof.

PROOF OF THEOREM 5. In order to prove this theorem, it is sufficient to show that if there exists a value function under which one criterion represents some policy $\mathcal{P}$, we can generate another value function under which the other criterion will generate $\mathcal{P}$ as well.

First, notice that *minmax regret* and *competitive ratio* are equivalent. One talks about ratios while the other about differences. Consequently, through the use of logarithms and exponentiation we can transform a value function for *minmax regret* to an equivalent value function for *competitive ratio* and vice versa. (Given that S is finite, obtaining an integer valued function is then easy.)

Next, we show that *maximin* and *minmax regret* are equivalent. Notice that *maximin* and a similar (not so rational) criterion *minimax* that attempts to minimize the maximal value of the action are equivalent. We simply need to multiply utilities by -1. Consequently, it is sufficient to show that *minmax regret* and *minimax* are equivalent. To do this, we have to show that given a regret matrix, we can generate a value function that has these regret values, and this is straightforward.

Finally, we wish to provide a formal statement of the relationship between the properties of closure under union and column duplication discussed in Section 6. The relationship discussed holds in the context of Milnor's [1954] discussion of decision criteria for reasoning under ignorance. Column duplication is a property that is discussed in the context of decision tables, and we shall attempt to relate this context to our presentation. We formalize it as follows: First, we assume that a decision criterion takes as input a finite, two dimensional table whose entries correspond to state-action pairs (i.e., a standard decision table) and returns one or more rows of this table, corresponding to the preferred actions. Further, we must assume that the decision criterion does not care about actions or states name, but only about the entries in the decision table. Finally, we must assume that there is an infinite number of states and that for any particular values assigned to a state with respect to the set of actions (i.e., a column in the decision table), we have an infinite number of states with such assigned values. Under these assumptions, we can show:

LEMMA 1. *A decision criterion is closed under unions iff it is closed under disjoint unions and has the column duplication property*.

PROOF OF LEMMA 1. Suppose that a decision criterion is closed under unions. Further, suppose it prefers a to a' given some state set V. Hence, it prefers a to a' given $(V\backslash\{s'\}) \cup \{s\}$, where s and s' are as in Definition 18 because the decision table corresponding to V and $(V\backslash\{s'\}) \cup \{s\}$ are identical. Closure under unions implies that a is preferred given $V \cup \{s\}$ as well.

Suppose that a decision criterion is closed under disjoint unions and has the column duplication property. Let U and V be two sets of states given which a is preferred to a'. Let $W = U \cap V$, and let W' be a set of states disjoint from U and V but identical in terms of their columns. Let $V' = (V\backslash U) \cup W'$. By closure under disjoint unions we know that a is preferred to a' given $V' \cup U$. Furthermore, $V' \cup U = V \cup U \cup W'$, and $V' \cup U$ contains W. It is immediate to see that the property of column duplication is identical to the property of identical column removal. Hence, we know that the decision criterion has the same preferences given $V' \cup U\backslash W' = V \cup U$.

REFERENCES

ALLEN, J., HENDLER, J., AND TATE, A., EDS. 1990. *Readings in Planning*. Morgan-Kaufmann, San Francisco, Calif.

ANSCOMBE, F. J., AND AUMANN, R. J. 1963. A definition of subjective probability. *Ann. Math. Stat. 34*, 199–205.

BLUM, L., BRANDENBURGER, A., AND DEKEL, E. 1991. Lexicographic probabilities and equilibrium refinements. *Econometrica 59*, 61–79.

BORODIN, A., AND EL-YANIV, R. 1998. *On-Line Computation and Competitive Analysis*. Cambridge University Press, Cambridge, Mass.

BOUTILIER, C. 1994. Toward a Logic for Qualitative Decision Theory. In *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning*. Morgan-Kaufmann, San Francisco, Calif., pp. 75–86.

BOUTILIER, C., BRAFMAN, R. I., GEIB, C., AND POOLE, D. 1997. Reasoning with ceteris paribus preference statements. In *Proceedings of the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning*. AAAI Press, Menlo Park, Calif.

BRAFMAN, R. I., AND FRIEDMAN, N. 1995. On decision-theoretic foundations for defaults. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*.

BRAFMAN, R. I., AND TENNENHOLTZ, M. 1994. Belief ascription and mental-level modelling. In *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning*, J. Doyle, E. Sandewall, and P. Torasso, Eds., pp. 87–98.

BRAFMAN, R., AND TENNENHOLTZ, M. 1997. Modeling agents as qualitative decision-makers. *Artif. Int. 94*, 217–268.

COOPER, G. F. 1990. The computational complexity of probabilistic inference using Bayesian belief networks. *Artif. Int. 42*, 2–3, 393–405.

DARWICHE, A., AND GOLDSZMIDT, M. 1994. On the relation between kappa calculus and probabilistic reasoning. In *Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence* (*UAI '94*). Morgan-Kaufmann, San Francisco, Calif., pp. 145–153.

DOYLE, J. 1989. Constructive belief and rational representation. *Comput. Int. 5*, 1–11.

DOYLE, J., AND THOMASON, R. 1997. Qualitative preferences in deliberation and practical reasoning. *Working notes of the AAAI spring symposium*.

DOYLE, J., AND WELLMAN, M. 1994. Representing preferences as *ceteris paribus* comparatives. In *Proceedings of the AAAI Spring Symposium on Decision-Theoretic Planning*. AAAI Press, Menlo Park, Calif., pp. 69–75.

DOYLE, J., AND WELLMAN, M. P. 1989. Impediments to universal preference-based default theories. In *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning*, (Toronto, Ont., Canada, May). Ron Brachman and Hector Levesque, Eds. Morgan-Kaufmann, San Francisco, Calif., pp. 94–102.

DUBOIS, D., AND PRADE, H. 1995. Possibility theory as a basis for qualitative decision theory. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*. pp. 1924–1930.

DUBOIS, D., PRADE, H., AND SABBADIN, R. 2000. Decision-theoretic foundations of qualitative possibility theory. *Europ. J. Oper. Res.*, to appear.

EL-YANIV, R. 2000. On the decision theoretic foundations of the competative ratio criterion. Unpublished manuscript.

FAGIN, R., HALPERN, J. Y., MOSES, Y., AND VARDI, M. Y. 1995. *Reasoning about Knowledge*. MIT Press, Cambridge, Mass.

FISHBURN, P. C. 1988. *Nonlinear Preference and Utility Theory*. Johns Hopkins University Press, Baltimore, Md.

FORBUS, K. D. 1985. Qualitative process theory. In *Qualitative Reasoning About Physical Systems*. D. Bobrow, Ed. The MIT Press, Cambridge, Mass.

FRIEDMAN, N., AND HALPERN, J. Y. 1994. A knowledge-based framework for belief change. Part I: Foundations. In *Proceedings of the 5th Conference on Theoretical Aspects of Reasoning About Knowledge*, Morgan-Kaufmann, San Francisco, Calif. pp. 44–64.

FUDENBERG, D., AND TIROLE, J. 1991. *Game Theory*. MIT Press, Cambridge, Mass.

GARDENFORS, P. 1992. *Belief Revision*. Cambridge University Press, Cambridge, Mass.

GILBOA, I., AND SCHMEIDLER, D. 1989. Maxmin expected utility with a non-unique prior. *J. Math. Econ. 18*, 141–153.

GILBOA, I., AND SCHMEIDLER, D. 1995. Case-based decision theory. *Quart. J. Econ. 110*, 605–639.

GINSBERG, M. L., ED. 1987. *Readings in Nonmonotonic Reasoning*. Morgan-Kaufmann, San Francisco, Calif.

HALPERN, J. Y. 1988. Reasoning about knowledge: An overview. In *Theoretical Aspects of Reasoning About Knowledge: Proceedings of the 1988 Conference*. Morgan-Kaufmann, San Francisco, Calif., pp. 1–17.

HALPERN, J. Y., AND MOSES, Y. 1990. Knowledge and common knowledge in a distributed environment. *J. ACM 37*, (3) (July), 549–587.

HART, S., MODICA, S., AND SCHMEIDLER, D. 1994. A neo-Bayesian foundation of the maxmin value for two-parson zero-sum games. *Int. J. Game Theory*, 23: 347–358

HOWARD, R. A. 1960. *Dynamic Programming and Markov Processes*. MIT Press, Cambridge, Mass.

KREPS, D. M. 1988. *Notes on the Theory of Choice*. Westview Press, Boulder, Colo.

LAMARRE, P., AND SHOHAM, Y. 1994. Knowledge, certainty, belief and conditionalization. In *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning*. Morgan-Kaufmann, San Francisco, Calif., pp. 415–424.

LEHMANN, D. 1996. Generalized qualitative probability: Savage revisited. In *Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence* (*UAI '96*). Morgan-Kaufmann, San Francisco, Calif., pp. 381–388.

LEVESQUE, H. 1986. Knowledge representation and reasoning. *Ann. Rev. Comput. Sci. 1*, 255–287.

LUCE, R. D., AND RAIFFA, H. 1957. *Games and Decisions*. Wiley, New York.

MILNOR, J. 1954. Games against nature. In *Decision Processes*, R. M. Thrall, C. H. Coombs, and R. L. Davis, Eds. Wiley, New York.

PAPADIMITRIOU, C. H., AND YANNAKAKIS, M. 1989. Shortest paths without a map. In *Automata, Languages and Programming*. *16th International Colloquium Proceedings*. pp. 610–620.

ROSENSCHEIN, S. J. 1985. Formal theories of knowledge in AI and robotics. *New Gen. Comput*. *3*, 3, 345–357.

SAVAGE, L. J. 1972. *The Foundations of Statistics*. Dover Publications, New York.

SHIMONY, S. E. 1994. Finding MAPs for belief networks is NP-hard. *Artif. Int. 68*, (2) (Aug.), pp. 399–410.

TAN, S., AND PEARL, J. 1994. Qualitative decision theory. In *Proceedings of the 12th National Conference on Artificial Intelligence*. Morgan-Kaufmann, San Francisco, Calif., pp. 928–932.

TAN, S. W., AND PEARL, J. 1994. Specification and evaluation of preferences under uncertainty. In *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning*. pp. 530–539.

VON NEUMANN, J., AND MORGENSTERN, O. 1944. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, N. J.

WALD, A. 1950. *Statistical Decision Functions*. Wiley, New York.

ZADEH, L. A. 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets syst. 1*, 3–28.

RECEIVED JULY 1997; REVISED SEPTEMBER 1999; ACCEPTED SEPTEMBER 1999