

# Preference Rankings in the Face of Uncertainty

Wim H. Hesselink, 16th December 2002

Dept. of Mathematics and Computing Science, Rijksuniversiteit Groningen

P.O.Box 800, 9700 AV Groningen, The Netherlands

Email: [wim@cs.rug.nl](mailto:wim@cs.rug.nl), Web: <http://www.cs.rug.nl/~wim>

## Abstract

An agent may have to choose between actions based on incomplete knowledge of its environment. The incomplete knowledge is modelled as the local state of the agent, which represents the set of states of the environment that the agent deems possible. A policy determines a ranking (as a total preorder) of the set of actions as a function of the local state. A policy is *maximin* representable when it is based on a utility function via the *maximin* principle. The theory of Brafman and Tennenholz on necessary and sufficient conditions for policies to be *maximin* representable is sharpened, extended, and related to *maximax* and Laplace representability.

**Keywords:** qualitative decision theory, maximin principle, total preorder, choice theory, coalition theory, compromise rule.

## 1 Introduction

This paper deals with decision making policies for intelligent agents. According to [8], an agent is a computer system that is situated in some environment and that is capable of autonomous action in this environment to meet its design objectives. Intelligent agents are able to perceive their environment and respond (in a timely fashion) to changes that occur in it. They exhibit goal-directed behaviour, and are capable of interacting with other agents (and possibly humans). In this note, however, we are concerned with the decision making of an individual agent in an environment that may be uncertain, but does not contain other agents.

A decision making policy is a method for the agent to choose an available action based on information about the environment. Slightly more general, a preference ranking policy is a ranking of all available actions rather than making a specific choice. A preference ranking policy can be based on estimates of the utilities of actions. The main problem solved in this paper is to characterize preference ranking policies based on utility functions via the so-called *maximin* principle.

In this effort, we build on the ideas and results of Brafman and Tennenholz. Their paper [1] is an outstanding case of concept formation in the sense of Lakatos [5], with many simple examples and an extensive overview of the related artificial intelligence literature. Yet, some technical definitions can be improved and some theorems can be sharpened and simplified. We also present some new results and many technical examples.

The theory of [1] is characterized by two modelling decisions. Firstly, the uncertainty of the agent about its environment is modelled by representing the agent's knowledge as a set of possible states of the environment. In order to compare the agent's preferences under various circumstances, the agent is supposed to assign priorities to all available actions. This is formalized in the concept of policy. The agent is supposed to be purely reactive in the sense that it does not use knowledge of the history, but only knowledge of the current state. Of course, the history or a relevant part of it may be encoded in the current state. So, a *policy* is a prescription of priorities to actions as a function of incomplete knowledge of the environment. An agent represents its incomplete knowledge of the environment in its local state, which is a nonempty set of states (those it regards as possible). Following [1], we thus distinguish between the *states* (of the environment) and the *local states* (of the agent); every *local state* is a nonempty set of *states*.

A utility function is a function that evaluates an action in a given *state*. The *maximin* principle extends this evaluation to *local states* by taking the minimal value for all states in the local state. So, it evaluates the action by considering a worst-case scenario. A policy is *maximin* represented by the utility function when it maximizes this extended utility function.

A preference ranking policy is deterministic when it always gives a complete ranking of the actions. Utility functions do not always force a choice. It is therefore natural to investigate nondeterministic policies as well.

An important special case is when there are only two actions, e.g., yes or no, good or bad, true or false. We then speak of *binary* decisions and policies. In some sense, the binary case is general enough, since we regard a policy as a way to choose between any pair of the available actions. In this paper, however, we emphasize the general case and treat the binary case only as a corollary of it. Indeed, special methods for the binary case tend to require proofs with ugly case distinctions.

The paper [1] presents four theorems that axiomatically characterize *maximin* representable policies. It does this by generalizing step by step from special cases. Theorem 1 characterizes deterministic binary policies. Theorem 2 treats nondeterministic binary policies. General deterministic policies are handled in Theorem 3, while Theorem 4 is about general nondeterministic policies. Theorems 2, 3, and 4 each use a different concept of closure under chains, a weak and complicated form of transitivity.

In all three cases, we replace it by a condition that a certain directed graph has no cycles. We first sharpen Theorem 3 by showing, in Theorem D, that a deterministic policy is *maximin* representable if and only if it is “closed under unions” and a certain graph has no cycles. We then prove Theorem 1 as a corollary. For nondeterministic policies we also prove, in Theorem G, that a policy is *maximin* representable if and only if it is “closed under unions” and a certain graph has no cycles. This sharpens Theorem 4 of [1]. As a corollary, we give a Theorem G2A for nondeterministic binary policies sharpening Theorem 2 of [1]. It follows that the condition of “respect of domination” required in Theorem 2 of [1] is superfluous.

We also consider the *maximax* principle where the state-wise evaluation is extended to local states by taking the maximal value for all states in the local state. It turns out that a policy is *maximin* representable if and only if its opposite is *maximax* representable, where the opposite of a policy is obtained by reversing all priorities. This result is especially useful to generate examples. By comparing *maximin* and *maximax* policies, we discover the “compromise rule” for *maximin* policies, which says that, if an action  $x$  is better than  $y$  for local state  $L$  and better than  $z$  for local state  $M$ , the combination  $L \cup M$  regards  $x$  as better than at least one of the alternatives  $y$  or  $z$ . This rule does not hold for *maximax* policies.

The term “compromise rule” corresponds to a more subjective interpretation of the theory, where the states represent parties rather than possible states of the environment and where local states represent coalitions of parties that have to reconcile their priorities. In this interpretation, the case with two states gets extra significance, since coalitions of two parties are fairly common (e.g., marriages). We show that the compromise rule is sufficient for *maximin* representability for deterministic policies when there are only two states.

In Laplace’s principle, the state-wise evaluation is extended to local states by taking the sum of the values for the states in the local state and preferring the action for which the sum is maximal. This principle mainly serves here to generate examples, but we also show that every deterministic policy that is *maximin* or *maximax* representable is also Laplace representable.

**Overview.** Section 2 formalizes policies and utility functions and provides two simple examples. We indicate the duality between *maximin* and *maximax* policies, introduce the compromise rule and closure under unions, and we briefly discuss the Laplace’s principle of indifference which yields a more quantitative policy foundation. Section 3 introduces the technical ingredients needed to axiomatize *maximin* representability of policies.

In Section 4, we introduce acyclic directed graphs and treat *maximin* representability of deterministic policies. We sharpen Theorem 3 of [1] to a Theorem D by simplifying its condition of closure under chains to acyclicity of a certain graph. We prove that Theorem 1 of [1] is a corollary of Theorem D. Actually, Theorem D itself is a corollary of Theorem G of Section 5, but for simplicity we give an independent proof. In the deterministic case with two states, we show that *maximin* representability is equivalent to the compromise rule. We finally prove that every deterministic policy that is *maximin* or *maximax* representable, is also Laplace representable.

In Section 5, we treat the general nondeterministic case and prove the main result, Theorem

G, which asserts that a policy is *maximin* representable if and only if it is closed under unions and a certain directed graph is acyclic. This is a sharpened version of Theorem 4 of [1]. In section 6, we prove a similar result for the binary case, which corresponds to Theorem 2 of [1]. This result is considerably harder to prove, but employs acyclicity of a smaller graph and is therefore, in applications, computationally cheaper. We conclude in Section 7.

**Related work.** As seen above, our main reference is [1]. We refer to [1, 3] for extensive overviews of related literature. At first sight, the inhibition relation described in [8], p. 24, looks similar to the approach investigated here. This is deceptive, however, for the inhibition relation is used to define a single function from percepts, which represent knowledge of the environment, to actions. Preference ranking based on voting procedures is the subject of social choice theory, e.g., see [6] and references given there.

**Notations.** We replace the definitions and arguments by case distinctions in natural language of [1] by more concise formulas in predicate logic, in a version as suggested by [2]. So, we use  $\equiv$  for logical equivalence,  $\wedge$  for conjunction (**and**),  $\vee$  for disjunction (**or**),  $\Rightarrow$  for implication,  $\neg$  for negation. Universal and existential quantification of a predicate  $R(x)$  with  $x$  ranging over a set  $X$  is denoted by  $(\forall x \in X :: R(x))$  and  $(\exists x \in X :: R(x))$ . If  $f$  is a numerical function on  $X$ , a similar notation is used for the minimum ( $\text{MIN } x \in X :: f(x)$ ) and the sum ( $\sum x \in X :: f(x)$ ). We write  $\#V$  for the number of elements of a set  $V$ . At some points, we also use the format of [2] to give predicate calculations with hints between braces to justify the steps.

## 2 Policies and utility functions

This section contains the main definitions. Given finite sets *States* and *Actions*, we define policies as functions that yield total preorders on *Actions*, with deterministic policies as special cases. We define utility functions as numerical functions on action-state pairs. We define three ways to associate a policy to a utility function: *maximin*, *maximax*, and Laplace. We give examples of such induced policies. We define duality of policies and introduce two important properties of policies: the compromise rule and closure under unions.

Assigning preferences is formalized by giving a total order if the policy is deterministic. Since there is sometimes no reason to prefer one action over another one, we also consider nondeterministic policies, which assign total preorders to local states.

Recall that a binary relation  $\preceq$  is called a *preorder* iff it is reflexive and transitive. The associated relations  $\prec$  and  $\approx$  are then defined by

$$\begin{aligned} x \prec y &\equiv \neg(y \preceq x), \\ x \approx y &\equiv x \preceq y \wedge y \preceq x. \end{aligned}$$

Preorder  $\preceq$  is called *total* (or *linear*) iff  $x \prec y$  implies  $x \preceq y$  for all  $x$  and  $y$ . If  $\preceq$  is total, then  $\prec$  is transitive (in other cases  $\prec$  is rather useless). The preorder  $\preceq$  is an order if and only if relation  $\approx$  equals the identity relation. For a total preorder  $T$ , we use  $\preceq_T$  to denote the corresponding infix operator, and we use  $\prec_T$  and  $\approx_T$  for the two associated relations. Total preorders are called “simple orderings” in [7] and “weak orders” in [3].

Since we aim at a general theory, we use the term *policy* for what is called a generalized s-policy in [1]. Let *States* be a nonempty finite set. The elements of *States* are called *states* and represent the possible states of the environment. We define the set *LocSt* as the set of nonempty subsets of *States*. The elements of *LocSt* are called *local states*, and represent the agent’s incomplete knowledge of the environment. Local states are called “events” in the probabilistic tradition of [7, 3].

Let *Actions* be a finite set with at least two elements. A *policy*  $P$  is defined to be a function from *LocSt* to the set of total preorders on *Actions*. For local state  $L$  and actions  $x$  and  $y$ , we write  $x \preceq_L^P y$  instead of  $x \preceq_{P(L)} y$ , and similarly with  $\preceq$  replaced by  $\prec$  and  $\approx$ . Policy  $P$  is called *deterministic* iff  $\preceq_L^P$  is a total order for every local state  $L$ , or equivalently iff  $x \approx_L^P y$  implies  $x = y$  for all actions  $x$  and  $y$ .

We define an *action-state pair* to be a pair  $(x, s)$  where  $x$  is an action and  $s$  is a state. We thus define the set *AS-pairs* as the cartesian product  $AS\text{-pairs} = \text{Actions} \times \text{States}$ . Policies are sometimes based on “utility” of action-state pairs. A *utility function*  $u$  is defined as a function  $u : AS\text{-pairs} \rightarrow \mathbb{R}$ . In order to get the best worst-case outcome for a policy, we extend utility function  $u$  to local states by giving the worst-case utility  $u_*(x, L) = (\text{MIN } s \in L :: u(x, s))$ . A policy  $P$  is said to be *maximin* represented by utility function  $u$ , cf. [1], iff for all  $L, x$  and  $y$

$$(0) \quad x \preceq_L^P y \equiv u_*(x, L) \leq u_*(y, L) .$$

It is easy to see that we then also have

$$(1) \quad x \prec_L^P y \equiv u_*(x, L) < u_*(y, L) .$$

*Example A.* Take three actions  $x, y, z$  and three states  $a, b, c$ . Let utility function  $u$  be given by the columns of  $a, b, c$  in the table

$u$	$a$	$b$	$c$	$u_* : ab$	$ac$	$bc$	$abc$
$x$	0	5	3	0	0	3	0
$y$	0	4	1	0	0	1	0
$z$	2	6	1	2	1	1	1

So, for example,  $u(x, b) = 5$ . In the columns of  $ab, ac, bc, abc$ , we give the values of  $u_*$  for the local states  $\{a, b\}, \{a, c\}$ , etc. For example,  $u_*(x, \{b, c\}) = 3$ . Let  $P$  be the policy that is *maximin* represented by utility function  $u$ . It is easy to see that  $P(\{a\})$  is the preorder with  $x \approx y \prec z$ , that  $P(\{b\})$  is the order with  $y \prec x \prec z$ , and that  $P(\{c\})$  is the preorder with  $y \approx z \prec x$ . For the larger local states, the policies are  $P(\{a, b\}) = P(\{a, c\}) = P(\{a, b, c\}) = P(\{a\})$  and  $P(\{b, c\}) = P(\{c\})$ .  $\square$

If  $P$  is an arbitrary policy, its *opposite* or *dual* is defined as the policy  $Q$  given by

$$x \preceq_L^Q y \equiv y \preceq_L^P x \text{ for all } x, y, L .$$

It is easy to see that, conversely,  $P$  is the opposite of  $Q$ .

Next to its worst-case extension  $u_*$ , every utility function  $u$  has also a best-case extension  $u^*$  given by  $u^*(x, L) = (\text{MAX } s \in L :: u(x, s))$ . A policy  $P$  is said to be *maximax* represented by utility function  $u$  iff for all  $L, x$  and  $y$

$$x \preceq_L^P y \equiv u^*(x, L) \leq u^*(y, L) .$$

The maximax representation is based on the optimistic idea to aim at the best possible outcome.

The fundamental alternative to *maximin* and *maximax* representability is Laplace’s principle of indifference [1] in which the action maximizing the sum of the values is preferred. Writing  $u^\#(x, L)$  for  $(\sum s \in L :: u(x, s))$ , we define policy  $P$  to be *Laplace* represented by  $u$  iff

$$x \preceq_L^P y \equiv u^\#(x, L) \leq u^\#(y, L) .$$

As a referee suggested, it is useful to note that Laplace representation has a probabilistic interpretation, cf. [7]: it maximizes the expected utility of actions under assumption that the states have equal probabilities.

The following duality theorem shows that the *maximin* and *maximax* principles are two sides of the same coin. It also shows that Laplace representability is self-dual.

**Duality Theorem.** (a) Policy  $P$  is *maximin* represented by utility function  $u$  if and only if its opposite policy is *maximax* represented by the utility function  $-u$ .

(b) Policy  $P$  is *Laplace* represented by utility function  $u$  if and only if its opposite policy is *Laplace* represented by the utility function  $-u$ .

*Proof.* (a) This follows from

$$\begin{aligned}
& (-u)^*(x, L) \\
&= (\text{MAX } s \in L :: -u(x, s)) \\
&= -(\text{MIN } s \in L :: u(x, s)) \\
&= -u_*(x, L) .
\end{aligned}$$

Therefore  $(-u)^*(y, L) \leq (-u)^*(x, L) \equiv u_*(x, L) \leq u_*(y, L)$ .

(b) This is proved in the same way.  $\square$

Below we show that not every *maximax* representable policy is *maximin* representable (and also the other way around). In order to prove this, we need to develop conditions necessary for *maximin* representability.

A typical property of a *maximin* representable policy  $P$  is the following *compromise rule*

$$(2) \quad y \prec_L^P x \wedge z \prec_M^P x \Rightarrow y \prec_{L \cup M}^P x \vee z \prec_{L \cup M}^P x .$$

Indeed, if  $u_*(y, L) < u_*(x, L)$  and  $u_*(z, M) < u_*(x, M)$ , then the minimum of  $u_*(y, L \cup M)$  and  $u_*(z, L \cup M)$  is less than  $u_*(x, L \cup M)$ . We call it the compromise rule, since the antecedent says that  $x$  is better than  $y$  for  $L$  and better than  $z$  for  $M$ , while the consequent says that  $L$  and  $M$  regard  $x$  as a compromise at least better than one of the alternatives  $y$  and  $z$ .

*Example B.* We show that not every *maximax* policy is a *maximin* policy by constructing a (deterministic) *maximax* policy that violates the compromise rule. We use three actions  $x, y, z$ , two states  $a$  and  $b$ , and let  $P$  be the policy that is *maximax* represented by the function  $u$  given in the table

$u$	$a$	$b$	$u^* : ab$
$x$	1	1	1
$y$	3	0	3
$z$	0	4	4

It follows that the total orders are given by  $P(\{a\}) : z \prec x \prec y$  and  $P(\{b\}) : y \prec x \prec z$  and  $P(\{a, b\}) : x \prec y \prec z$ . Policy  $P$  clearly violates the compromise rule for  $L = \{b\}$  and  $M = \{a\}$ .  $\square$

The paper [1] does not mention the compromise rule but emphasizes what is called closure under unions. A policy  $P$  is said to be *closed under unions*, cf. [1], iff for local states  $L$  and  $M$ , and actions  $x$  and  $y$ , we always have

$$\begin{aligned}
x \preceq_L^P y \wedge x \preceq_M^P y &\Rightarrow x \preceq_{L \cup M}^P y , \\
x \prec_L^P y \wedge x \prec_M^P y &\Rightarrow x \prec_{L \cup M}^P y .
\end{aligned}$$

Every *maximin* representable policy is closed under unions. This is proved in the same way as the compromise rule for such policies. It is also easy to verify that a policy  $P$  is closed under unions if and only if its opposite is. By Duality Theorem, it follows that every *maximax* representable policy is also closed under unions.

Laplace representable policies are not necessarily closed under unions. For example, if  $Q$  is the policy Laplace-represented by function  $u$  of example A, then  $z \preceq_L^Q x$  for both  $L = \{a, c\}$  and  $L = \{b, c\}$ , but not for the union  $L = \{a, b, c\}$ . They are closed under *disjoint* unions. More precisely, every Laplace representable policy  $Q$  satisfies the following ‘‘Laplace rules’’:

$$\begin{aligned}
x \preceq_L^Q y \wedge x \prec_M^Q y \wedge L \cap M = \emptyset &\Rightarrow x \prec_{L \cup M}^Q y , \\
x \approx_L^Q y \wedge x \approx_M^Q y \wedge L \cap M = \emptyset &\Rightarrow x \approx_{L \cup M}^Q y .
\end{aligned}$$

For example, policy  $P$  of example A is not Laplace representable, since the first Laplace rule fails for  $L = \{a\}$  and  $M = \{b\}$ . The policy of example B is Laplace represented by the utility function  $u$  used there. This shows that Laplace representable policies can be closed under unions and need not satisfy the compromise rule.

The axiomatization of Laplace representable policies is a nice problem, but since it seems to require completely different techniques (like linear programming), we leave it aside.

### 3 Two critical relations on action-state pairs

Our aim is to find necessary and sufficient conditions for policy  $P$  to be *maximin* representable. It seems that the compromise rule, though necessary, is not very useful. Closure under unions is also necessary, but clearly insufficient because of Example B. In this section we introduce binary relations  $\prec_{\bullet}^P$  and  $\preceq_{\bullet}^P$  on *AS-pairs*, the properties of which will determine the *maximin* representability of  $P$ .

It turns out that, if policy  $P$  is closed under unions, it is determined by its behaviour on sets of states with  $\leq 2$  elements or, more precisely, by two induced relations on *AS-pairs*. Indeed, inspired by [1], we define for an arbitrary policy  $P$ , relations  $\prec_{\bullet}^P$  and  $\preceq_{\bullet}^P$  on action-state pairs by

$$(3) \quad \begin{aligned} (x, s) \prec_{\bullet}^P (y, t) &\equiv \\ &(s = t \wedge x \prec_{\{t\}}^P y) \vee (x \prec_{\{s,t\}}^P y \wedge y \preceq_{\{t\}}^P x), \\ (x, s) \preceq_{\bullet}^P (y, t) &\equiv \\ &(s = t \wedge x \preceq_{\{t\}}^P y) \vee (x \preceq_{\{s,t\}}^P y \wedge y \prec_{\{t\}}^P x). \end{aligned}$$

For action-state pairs with the same states  $s = t$ , relation  $\prec_{\bullet}^P$  equals  $\prec_{\{t\}}^P$  and  $\preceq_{\bullet}^P$  equals  $\preceq_{\{t\}}^P$ . For pairs with the different states  $s \neq t$ , the relations indicate that the states differ in their preferences with respect to the actions involved.

Note that, in either case, the second disjunct implies  $s \neq t$ . So, in this case, the operator  $\vee$  serves as an exclusive **or**. Relation  $\prec_{\bullet}^P$  may have cycles and is not necessarily transitive, as is shown in the next example.

*Example B, continued.* Policy  $P$  of example B has the cycle  $(x, a) \prec_{\bullet}^P (y, b) \prec_{\bullet}^P (z, a) \prec_{\bullet}^P (x, a)$ . Indeed, relation  $(x, a) \prec_{\bullet}^P (y, b)$  is proved by observing that  $x \prec y$  in local state  $\{a, b\}$  whereas  $y \prec x$  in local state  $\{b\}$ . Relation  $(y, b) \prec_{\bullet}^P (z, a)$  is proved in the same way. Relation  $(z, a) \prec_{\bullet}^P (x, a)$  follows from  $z \prec x$  in local state  $\{a\}$ . Since  $(x, a) \prec_{\bullet}^P (z, a)$  is false, this also shows that  $\prec_{\bullet}^P$  need not be transitive.  $\square$

The next result justifies the definitions of  $\prec_{\bullet}^P$  and  $\preceq_{\bullet}^P$  in relation to the *maximin* principle.

(4) **Lemma.** Let policy  $P$  be *maximin* represented by utility function  $u$ . For action-state pairs  $(x, s)$  and  $(y, t)$ , we have

$$\begin{aligned} (a) \quad (x, s) \prec_{\bullet}^P (y, t) &\Rightarrow u(x, s) < u(y, t), \\ (b) \quad (x, s) \preceq_{\bullet}^P (y, t) &\Rightarrow u(x, s) \leq u(y, t). \end{aligned}$$

*Proof.* (a) Assume  $(x, s) \prec_{\bullet}^P (y, t)$ . Relation  $\prec_{\bullet}^P$  is given in (3) by a disjunction. If  $s = t$  and  $x \prec_{\{t\}}^P y$ , we have  $u_*(x, \{t\}) < u_*(y, \{t\})$  and hence  $u(x, s) < u(y, t)$  since  $s = t$ . We treat the second disjunct by observing that

$$\begin{aligned} &x \prec_{\{s,t\}}^P y \wedge y \preceq_{\{t\}}^P x \\ \equiv &\{(1) \text{ and } (0)\} \\ &u_*(x, \{s, t\}) < u_*(y, \{s, t\}) \wedge u_*(y, \{t\}) \leq u_*(x, \{t\}) \\ \equiv &\{\text{definition of } u_*\} \\ &\min(u(x, s), u(x, t)) < \min(u(y, s), u(y, t)) \wedge u(y, t) \leq u(x, t) \\ \Rightarrow &\{\text{use the second conjunct to rewrite the first conjunct}\} \\ &u(x, s) < \min(u(y, s), u(y, t)) \\ \Rightarrow &\{\text{calculus}\} \\ &u(x, s) < u(y, t). \end{aligned}$$

Case (b) is similar.  $\square$

The next result is another justification for the definition of  $\prec_{\bullet}^P$  and  $\preceq_{\bullet}^P$ . It shows that, if  $P$  is closed under unions, it is determined by  $\prec_{\bullet}^P$  and  $\preceq_{\bullet}^P$ , i.e., by its behaviour on sets of states with  $\leq 2$  elements.

(5) **Lemma.** Let policy  $P$  be closed under unions. Let  $L$  be a local state and let  $x$  and  $y$  be actions. Then we have

- (a)  $x \prec_L^P y \equiv (\forall t \in L :: (\exists s \in L :: (x, s) \prec_{\bullet}^P (y, t)))$  ,  
 (b)  $x \preceq_L^P y \equiv (\forall t \in L :: (\exists s \in L :: (x, s) \preceq_{\bullet}^P (y, t)))$  .

*Proof.* (a) The proof is by mutual implication. Assume the righthand side:

$$(6) \quad (\forall t \in L :: (\exists s \in L :: (x, s) \prec_{\bullet}^P (y, t))) .$$

Let  $Q$  be the set of the sets  $M = \{s, t\}$  such that  $s, t \in L$  and  $(x, s) \prec_{\bullet}^P (y, t)$ . Then  $L$  is the union of  $Q$  and  $x \prec_M^P y$  for all  $M$  in  $Q$ . Using closure under unions for  $\prec^P$ , we get  $x \prec_L^P y$ .

Conversely, assume that formula (6) is false. Then we can choose a state  $t_0 \in L$  such that  $\neg(x, s) \prec_{\bullet}^P (y, t_0)$  holds for all  $s \in L$ . By the rule of De Morgan,  $\neg(x, s) \prec_{\bullet}^P (y, t_0)$  is equivalent to

$$(s \neq t_0 \vee y \preceq_{\{t_0\}}^P x) \wedge (y \preceq_{\{s, t_0\}}^P x \vee x \prec_{\{t_0\}}^P y) .$$

Applying this to  $s := t_0$ , we find that  $y \preceq_{\{t_0\}}^P x$ . It follows that  $y \preceq_{\{s, t_0\}}^P x$  holds for all  $s \in L$ . We can write  $L$  as the union of the sets  $\{s, t_0\}$  with  $s \in L$ . Using closure under unions for  $\preceq^P$ , we get  $y \preceq_L^P x$ , that is  $\neg(x \prec_L^P y)$ .

(b) The proof of this is completely analogous, with  $\prec^P$  and  $\preceq^P$  interchanged consistently.  $\square$

We regard the validity of Lemma (5) as a miracle. Conversely, however, it is easy to verify that a policy is closed under unions when it satisfies (a) and (b) for all  $L$ ,  $x$ , and  $y$ . Lemma (5) is the key to the following converse of Lemma (4).

(7) **Lemma.** Let policy  $P$  be closed under unions. Let  $u$  be a utility function that, for all pairs of action-state pairs  $(x, s)$  and  $(y, t)$  with  $x \neq y$ , satisfies

- (a)  $(x, s) \prec_{\bullet}^P (y, t) \Rightarrow u(x, s) < u(y, t)$  ,  
 (b)  $(x, s) \preceq_{\bullet}^P (y, t) \Rightarrow u(x, s) \leq u(y, t)$  .

Then  $P$  is *maximin* represented by utility function  $u$ .

*Proof.* It suffices to prove, for all pairs of actions  $x \neq y$ , and all local states  $L$ ,

$$(8) \quad x \preceq_L^P y \equiv u_*(x, L) \leq u_*(y, L) .$$

We first prove that the negation of the lefthand side implies the negation of the righthand side.

$$\begin{aligned} & y \prec_L^P x \\ \equiv & \{ \text{Lemma (5)(a)} \} \\ & (\forall t \in L :: (\exists s \in L :: (y, s) \prec_{\bullet}^P (x, t))) \\ \Rightarrow & \{ x \neq y \text{ and condition (a)} \} \\ & (\forall t \in L :: (\exists s \in L :: u(y, s) < u(x, t))) \\ \Rightarrow & \{ \text{choose } t \in L \text{ with } u(x, t) = u_*(x, L) \} \\ & u_*(y, L) < u_*(x, L) . \end{aligned}$$

An analogous calculation shows that, in (8), the lefthand side implies the righthand side:

$$\begin{aligned} & x \preceq_L^P y \\ \equiv & \{ \text{Lemma (5)(b)} \} \\ & (\forall t \in L :: (\exists s \in L :: (x, s) \preceq_{\bullet}^P (y, t))) \\ \Rightarrow & \{ x \neq y \text{ and condition (b)} \} \\ & (\forall t \in L :: (\exists s \in L :: u(x, s) \leq u(y, t))) \\ \Rightarrow & \{ \text{choose } t \in L \text{ with } u(y, t) = u_*(y, L) \} \\ & u_*(x, L) \leq u_*(y, L) . \end{aligned}$$

$\square$

The lemmas (4) and (7) are the key results needed for our main theorems D and G in the sections 4 and 5 below.

## 4 Acyclicity and the deterministic cases

This section contains our results on deterministic policies. The main result is that a deterministic policy is *maximin* representable if and only if it is closed under unions and relation  $\prec_{\bullet}^P$  is acyclic. We therefore need to define and investigate acyclicity of binary relations first.

Lemma (7) suggests to construct utility function  $u$  in some way from the relations  $\prec_{\bullet}^P$  and  $\preceq_{\bullet}^P$ . Since we can reuse the following construction in a different context, we generalize the setting.

Recall that a binary relation  $R$  on a set  $X$  is called *acyclic* iff, for every sequence of elements  $x_0, \dots, x_m$  of  $X$  with  $m > 0$  and  $(x_i, x_{i+1}) \in R$  for  $0 \leq i < m$ , it holds that  $x_0 \neq x_m$ . In general, we define the  $R$ -level of an element  $x \in X$  to be the maximal number  $m$  for which there exists a sequence  $x_0, \dots, x_m$  such that  $m \geq 0$  and  $x_m = x$  and  $(x_i, x_{i+1}) \in R$  for all  $0 \leq i < m$ . The  $R$ -level of  $x$  is infinite if such sequences can be arbitrarily long. We need the following well-known result.

(9) **Lemma.** A relation  $R$  on a finite set  $X$  is acyclic if and only if all  $R$ -levels are finite.

*Proof.* Let  $Y$  be the set of elements of  $X$  with infinite  $R$ -levels. We prove that  $R$  has a cycle if and only if  $Y$  is nonempty. If relation  $R$  has a cycle, all elements of the cycle belong to  $Y$ . So  $Y$  is nonempty. Conversely, assume that  $Y$  is nonempty. For every  $y \in Y$ , finiteness of  $X$  implies the existence of an element  $z \in Y$  with  $(z, y) \in R$ . Inductively, we can therefore find an infinite backward path in  $Y$ . Since the set  $Y$  is finite, this path contains a cycle.  $\square$

The following algorithm, a variation of the classical algorithm for topological sorting, determines the  $R$ -levels of all elements of a finite set  $X$ .

```

var lev : array  $X$  of integer ;
var Rest :=  $X$  ; var i := 0 ;
var First := {  $x \in Rest \mid (\forall y \in Rest :: (y, x) \notin R)$  } ;
while First  $\neq \emptyset$  do
  for all  $x \in First$  do lev( $x$ ) := i end ;
  Rest := Rest  $\setminus$  First ; i := i + 1 ;
  First := {  $x \in Rest \mid (\forall y \in Rest :: (y, x) \notin R)$  } ;
end .

```

In the postcondition,  $lev(x)$  is the  $R$ -level of  $x$  for every  $x$  with finite level, and  $Rest$  holds all elements of  $X$  with infinite  $R$ -levels. So,  $Rest$  is empty if and only if  $R$  is acyclic.

In the remainder of this section, we consider the case that policy  $P$  is deterministic. This means that  $x \preceq_L^P y$  is equivalent to  $x = y$  or  $x \prec_L^P y$  for every pair of actions  $x, y$ , and every local state  $L$ . We now have the deterministic theorem:

**Theorem D.** Let policy  $P$  be deterministic. Then  $P$  is *maximin* representable if and only if  $P$  is closed under unions and relation  $\prec_{\bullet}^P$  on *AS-pairs* is acyclic.

*Proof.* First assume that policy  $P$  is *maximin* represented by utility function  $u$ . In Section 2, we saw that  $P$  is closed under unions. For any chain of action-state pairs  $(x_0, s_0) \prec_{\bullet}^P \dots \prec_{\bullet}^P (x_m, s_m)$ , we have  $u(x_0, s_0) < \dots < u(x_m, s_m)$  by Lemma (4)(a). Therefore, relation  $\prec_{\bullet}^P$  is acyclic.

Conversely, assume that  $P$  is closed under unions and that relation  $\prec_{\bullet}^P$  is acyclic. Let  $lev$  be the corresponding  $\prec_{\bullet}^P$ -level function on *AS-pairs*. Then  $lev$  can be used as a utility function and it satisfies

$$(x, s) \prec_{\bullet}^P (y, t) \Rightarrow lev(x, s) < lev(y, t) .$$

Since  $P$  is deterministic, it follows from (3) that

$$(x, s) \preceq_{\bullet}^P (y, t) \equiv (x, s) = (y, t) \vee (x, s) \prec_{\bullet}^P (y, t) .$$

We therefore also have

$$(x, s) \preceq_{\bullet}^P (y, t) \Rightarrow lev(x, s) \leq lev(y, t) .$$

By Lemma (7), it follows that  $P$  is *maximin* represented by utility function  $lev$ .  $\square$



*Remark.* This is a sharpened version of Theorem 3 of [1]. Assuming closure under unions, one can prove that  $(x, s) \prec_{\bullet}^P (y, t)$  is equivalent to the defining expression for  $(x, s) <_P (y, t)$  in Definition 10 of [1]. This Definition then proceeds to define  $<_P$  to be *closed under chains* if whenever  $(a_1, s_1) <_P \dots <_P (a_k, s_k)$  and either (1)  $a_1 \prec_{\{s_1\}}^P a_k$  and  $a_k \prec_{\{s_k\}}^P a_1$  or (2)  $s_1 = s_k$ , then  $(a_1, s_1) <_P (a_k, s_k)$ . Using case (2), one easily sees that relation  $<_P$  is acyclic when it is closed under chains. Theorem 3 of [1] asserts that, if policy  $P$  is deterministic and closed under unions and  $<_P$  is closed under chains, then  $P$  is *maximin* representable. This result now follows from Theorem D above. So, Theorem 3 of [1] is correct, but its definition of closure under chains is needlessly complicated, since case (1) is superfluous.  $\square$

In the special case where  $P$  is deterministic and binary, the acyclicity requirement can be lifted since it always holds. This is shown in the next result.

**Theorem D2A.** Assume that  $P$  is deterministic and that  $\#Actions = 2$ . Then  $P$  is *maximin* representable if and only if it is closed under unions.

*Proof.* In view of Theorem D, it suffices to prove that, in this case, closure under unions implies acyclicity of relation  $\prec_{\bullet}^P$ .

Consider a cycle of  $m > 0$  action-state pairs  $(x_i, s_i)$  with  $0 \leq i \leq m$  such that  $(x_i, s_i) \prec_{\bullet}^P (x_{i+1}, s_{i+1})$  for  $0 \leq i < m$  and  $(x_0, s_0) = (x_m, s_m)$ . For every  $i$  with  $0 \leq i < m$ , formula (3) implies that  $x_i \prec_M^P x_{i+1}$  with  $M = \{s_i, s_{i+1}\}$ . It follows that  $x_i \neq x_{i+1}$ . Since there are only two actions, this implies that  $m$  is even, and that  $x_{2i} = x_0$  and  $x_{2i+1} = x_1$  for all  $i$ . Consider the local state  $L = \{s_k \mid 0 \leq k < m\}$ . We can write  $L$  in two ways as a disjoint union of subsequent pairs, namely pairs  $\{s_{2i}, s_{2i+1}\}$  with  $0 \leq i < \frac{1}{2}m$ , and pairs  $\{s_{2i+1}, s_{2i+2}\}$  with  $0 \leq i < \frac{1}{2}m$ . Since  $P$  is closed under unions, the first representation implies  $x_0 \prec_L^P x_1$ . Similarly, the second representation implies  $x_1 \prec_L^P x_0$ . This contradicts the fact that  $\preceq_L^P$  is a preorder.  $\square$

Using Lemma (11) below, one can show that Theorem D2A is equivalent to Theorem 1 of [1]. As indicated in the introduction, the case with only two states has also special significance under certain interpretations of the theory. In this case, the compromise rule plays a role.

**Theorem D2S.** Assume that  $P$  is deterministic and that  $\#States = 2$ . Then  $P$  is *maximin* representable if and only if it satisfies the compromise rule (2).

*Proof.* The compromise rule was introduced in Section 2 as a property of all *maximin* representable policies. It therefore suffices to prove the other implication. Assume that  $P$  is deterministic and satisfies the compromise rule. This easily implies closure under unions. In view of Theorem D, it therefore remains to show that relation  $\prec_{\bullet}^P$  is acyclic. Let a step  $(x, s) \prec_{\bullet}^P (y, t)$  be called state-changing when  $s \neq t$ .

Assume that relation  $\prec_{\bullet}^P$  has a cycle. Since the restriction of  $\prec_{\bullet}^P$  to action-state pairs with the same state is a total order, the cycle contains at least two state-changing steps. Let  $E$  be the set of action-state pairs  $p$  for which the cycle contains some state-changing step  $p \prec_{\bullet}^P q$ . We compare the action-state pairs in  $E$  by means of the total order  $\preceq_S^P$ , where  $S = States$ . We can choose  $(i, s) \in E$  such that  $i$  is minimal with respect to  $\preceq_S^P$ . Let  $(i, s) \prec_{\bullet}^P (k, t)$  for some  $(k, t)$  with  $s \neq t$ . Since  $\#S = 2$ , we have  $S = \{s, t\}$  and  $i \prec_S^P k$  and  $k \prec_{\{t\}}^P i$ .

Let  $(p, t) \prec_{\bullet}^P (j, s)$  be the last state-changing step of the cycle before action-state pair  $(i, s)$ . Then  $(p, t) \in E$ , so  $i \preceq_S^P p$  by minimality. On the other hand,  $p \prec_S^P j$  because of  $(p, t) \prec_{\bullet}^P (j, s)$ . This implies  $i \prec_S^P j$  and, hence,  $i \neq j$ . Finally, we have  $j \prec_{\{s\}}^P i$  since the cycle goes from  $(j, s)$  to  $(i, s)$  without state changes. We thus have

$$j \prec_{\{s\}}^P i \wedge k \prec_{\{t\}}^P i \wedge i \prec_S^P j \wedge i \prec_S^P k.$$

This contradicts the compromise rule (2) for  $L = \{s\}$  and  $M = \{t\}$ .  $\square$

In this result, the assumption  $\#States = 2$  is essential, as is shown by the next example.

*Example C.* We give a deterministic policy  $P$  for a system with three states  $a, b, c$  and four actions  $0, 1, 2, 3$ , that satisfies the compromise rule and is not *maximin* representable. Policy  $P$  is given by listing the local states as strings of states with the corresponding orders:

local states	order
$a, ab :$	$3 \prec 0 \prec 2 \prec 1 .$
$b :$	$3 \prec 1 \prec 0 \prec 2 .$
$c, ac :$	$0 \prec 2 \prec 3 \prec 1 .$
$bc :$	$0 \prec 3 \prec 1 \prec 2 .$
$abc :$	$0 \prec 3 \prec 2 \prec 1 .$

The policy is not *maximin* representable because of the four-cycle  $(0, a) \prec_{\bullet}^P (1, b) \prec_{\bullet}^P (2, c) \prec_{\bullet}^P (3, a) \prec_{\bullet}^P (0, a)$ . The verification of the cycle and of the compromise rule can be done by hand, but we used the computer to generate this case by an exhaustive search with the functional language Haskell [4]. One can verify that this policy is Laplace representable with utility function  $u$  given by

$u$	$a$	$b$	$c$	$u^{\#} : ab$	$ac$	$bc$	$abc$
0	1	2	0	3	1	2	3
1	8	1	5	9	13	6	14
2	2	6	1	8	3	7	9
3	0	0	4	0	4	4	4

□

In example A, we gave a nondeterministic *maximin* representable policy that was not Laplace representable. We now show that nondeterminism was essential.

**Theorem DL.** Let  $P$  be a deterministic policy that is *maximin* or *maximax* representable. Then  $P$  is Laplace representable.

*Proof.* We first treat the case that  $P$  is *maximax* represented by utility function  $u$ . Let  $N$  be the number of states. Since the *maximax* principle only uses the order of the values of  $u$  and since the number of action-state pairs is finite, we may assume that  $u(x, s) \geq 0$  for all action-state pairs  $(x, s)$  and that, for all pairs of action-state pairs  $(x, s)$  and  $(y, t)$ , we have

$$(*) \quad u(x, s) < u(y, t) \Rightarrow N \cdot u(x, s) < u(y, t) .$$

For actions  $x$  and  $y$ , and local state  $L$ , we now observe

$$\begin{aligned} & u^*(x, L) < u^*(y, L) \\ \Rightarrow & \{ \text{definition } u^* \} \\ & (\exists t \in L :: (\forall s \in L :: u(x, s) < u(y, t))) \\ \Rightarrow & \{ (*) \} \\ & (\exists t \in L :: (\forall s \in L :: N \cdot u(x, s) < u(y, t))) \\ \Rightarrow & \{ \text{calculus and definition } u^{\#} \} \\ & (\exists t \in L :: u^{\#}(x, L) < u(y, t)) \\ \Rightarrow & \{ \text{calculus and definition } u^{\#} \} \\ & u^{\#}(x, L) < u^{\#}(y, L) . \end{aligned}$$

Since policy  $P$  is deterministic and *maximax* represented by  $u$ , it follows that

$$x \preceq_L^P y \equiv u^{\#}(x, L) \leq u^{\#}(y, L) .$$

This proves that  $P$  is Laplace represented by  $u$ .

Now, assume that  $P$  is *maximin* representable. Then its opposite is *maximax* representable and therefore Laplace representable. By the Duality Theorem, it follows that  $P$  itself is Laplace representable. □

## 5 Acyclicity and the nondeterministic Theorem G

In this section, we treat the general case of nondeterministic policies.

We first generalize the concept of acyclicity in the following way. Let a pair of binary relations  $(R, S)$  be called *acyclic* iff, for every sequence of elements  $x_0, \dots, x_m$  of  $X$  with  $m > 0$  and  $(x_i, x_{i+1}) \in R \cup S$  for all  $0 \leq i < m$  and  $(x_i, x_{i+1}) \in R$  for at least one index  $i$ , it holds that  $x_0 \neq x_m$ . It is easy to see that a pair of binary relations  $(R, S)$  is acyclic if and only if the relation  $R \circ S^*$  is acyclic, where

$$R \circ S^* = \{(x, y) \mid (\exists z :: (x, z) \in R \wedge (z, y) \in S^*)\}$$

is the relational composition of relation  $R$  with the reflexive transitive closure  $S^*$  of relation  $S$ .

This concept of acyclicity replaces the three concepts of “closure under chains” of [1], Definitions 7, 10, 14.

(10) **Lemma.** Let  $(R, S)$  be a pair of binary relations on a finite set  $X$ . The pair  $(R, S)$  is acyclic if and only if there is a numerical function  $f$  on  $X$  with

- (a) If  $(x, y) \in R$ , then  $f(x) < f(y)$ .
- (b) If  $(x, y) \in S$ , then  $f(x) \leq f(y)$ .

*Proof.* Let us write  $T = R \circ S^*$ . First, assume that a function  $f$  with the proposed properties exist. Then  $f(x) < f(y)$  for all  $(x, y) \in T$ . It follows that relation  $T$  is acyclic, so that the pair  $(R, S)$  is acyclic. Conversely, assume that the pair is acyclic, so that  $T$  is acyclic. The function  $T$ -level then satisfies the conditions (a) and (b).  $\square$

**Theorem G.** A policy  $P$  is *maximin* representable if and only if  $P$  is closed under unions and the pair of relations  $(\prec_{\bullet}^P, \preceq_{\bullet}^P)$  on  $AS$ -pairs is acyclic.

*Proof.* First, assume that  $P$  is *maximin* represented by utility function  $u$ . In Section 2, we saw that this implies that  $P$  is closed under unions. It follows from Lemma (4) that function  $u$  satisfies the condition of Lemma (10), so that the pair  $(\prec_{\bullet}^P, \preceq_{\bullet}^P)$  is acyclic.

Conversely, assume that  $P$  is closed under unions and that the pair  $(\prec_{\bullet}^P, \preceq_{\bullet}^P)$  is acyclic. By Lemma (10), there is a utility function  $f$  that satisfies the requirements of Lemma (7) and hence yields a *maximin* representation of policy  $P$ .  $\square$

*Remarks.* The utility function  $f$  obtained via the proof of Lemma (10) is the  $T$ -level function for relation  $T = (\prec_{\bullet}^P) \circ (\preceq_{\bullet}^P)^*$ .

Theorem G is inspired by Theorem 4 of [1], which however does not give a necessary condition for *maximin* representability. In fact, it requires that a certain relation  $=_P$  is *closed under chains* which property implies that, whenever  $(a_1, s_1) =_P \dots =_P (a_k, s_k)$  and  $a_1 \neq a_k$  and  $s_1 \neq s_k$  and  $a_1 \preceq_{\{s_1\}}^P a_k$  and  $a_k \preceq_{\{s_k\}}^P a_1$  then  $(a_1, s_1) =_P (a_k, s_k)$ . This condition is not necessary for *maximin* representability. Relation  $=_P$  is defined in [1] in such a way that, in our terms,  $(x, s) =_P (y, t)$  is equivalent to

$$(s = t \wedge x \approx_{\{s\}}^P y) \vee (s \neq t \wedge x \neq y \wedge x \preceq_{\{s\}}^P y \wedge y \preceq_{\{t\}}^P x \wedge x \approx_{\{s,t\}}^P y).$$

One can verify that, in example A, we have the relations  $(x, a) =_P (y, b) =_P (z, c)$ . Moreover  $x \neq z$  and  $a \neq c$  and  $x \preceq_{\{a\}}^P z$  and  $z \preceq_{\{c\}}^P x$ . But the conclusion of closure under chains that  $(x, a) =_P (z, c)$  should hold, is false. This shows that closure under chains does not follow from *maximin* representability, as is asserted in [1].  $\square$

*Example D.* We now construct a nondeterministic binary policy  $P$ , that is closed under unions and satisfies the compromise rule but is neither *maximin* nor *maximax* representable. We use two actions  $x$  and  $y$ , and four states  $a, b, c, d$ . Policy  $P$  is given by listing the local states as strings of states with the corresponding preorders:

local states	preorder
$a, c, ab, ac, abc :$	$x \prec y .$
$b, d, bc, bd, bcd :$	$y \prec x .$
$ad, cd, abd, acd, abcd :$	$x \approx y .$

In this case, a tedious verification is needed to verify that  $P$  is closed under unions. Since there are only two actions, it is easy to see that the compromise rule (2) follows from closure under unions. A tedious verification shows that we have a four-cycle  $(x, a) \prec_{\bullet}^P (y, b) \prec_{\bullet}^P (x, c) \preceq_{\bullet}^P (y, d) \preceq_{\bullet}^P (x, a)$ . Therefore,  $P$  is not *maximin* representable.

In order to show that  $P$  is not *maximax* representable, it suffices by Duality Theorem to prove that its opposite is not *maximin* representable. Let  $Q$  be the opposite of  $P$ . In general, for opposite policies  $P$  and  $Q$ , we have the rules

$$\begin{aligned} (x, s) \prec_{\bullet}^P (y, t) &\equiv (y, s) \prec_{\bullet}^Q (x, t) , \\ (x, s) \preceq_{\bullet}^P (y, t) &\equiv (y, s) \preceq_{\bullet}^Q (x, t) . \end{aligned}$$

Therefore, the above four-cycle induces the four-cycle  $(y, a) \prec_{\bullet}^Q (x, b) \prec_{\bullet}^Q (y, c) \preceq_{\bullet}^Q (x, d) \preceq_{\bullet}^Q (y, a)$ . As above, this shows that  $Q$  is not *maximin* representable so that  $P$  is not *maximax* representable.  $\square$

## 6 The nondeterministic binary case

This section is devoted the nondeterministic binary case and in particular to Theorem G2A, which sharpens Theorem 2 of [1]. So, we assume  $\#Actions = 2$ . Of course, Theorem G is applicable to this case and relates *maximin* representability to acyclicity of relations on *AS-pairs*. In the binary case, Theorem 2 of [1] relates *maximin* representability to acyclicity of relations on the smaller set *States*.

Since  $\#Actions = 2$ , there are only three total preorders on *Actions*: either element can be the greatest one, or they are equivalent. In other words, the total preorder  $\preceq$  is determined by the set of the greatest elements  $Gr(\preceq) = \{x \mid (\forall y :: y \preceq x)\}$ . It follows that a policy  $P$  is determined by the values  $P[L] = Gr(\preceq_L^P)$ . Note that  $P[L]$  always consists of one or two actions, i.e., we have  $\#P[L] = 1$  or  $\#P[L] = 2$ .

Since we want to retain the general definitions in the special case, we first prove that the special definition of [1] indeed corresponds to the general definition:

(11) **Lemma.** The following conditions on  $P$  are equivalent:

- (a)  $P$  is closed under unions.
- (b) For all  $L, M \in LocSt$ , we have  $P[L] \cap P[M] \subseteq P[L \cup M] \subseteq P[L] \cup P[M]$ .
- (c) For all  $L, M \in LocSt$ , the set  $P[L \cup M]$  equals  $P[L]$  or  $P[M]$  or  $P[L] \cup P[M]$ .

*Proof.* We first prove (a) $\equiv$ (b). Closure under unions is a conjunction of two implications universally quantified over actions  $x$  and  $y$  and local states  $L$  and  $M$ . Both implications are trivial for  $x = y$ . So, assume that  $x \neq y$ . Since  $\#Actions = 2$ , the implications are equivalent to

$$\begin{aligned} y \in P[L] \wedge y \in P[M] &\Rightarrow y \in P[L \cup M] , \\ x \notin P[L] \wedge x \notin P[M] &\Rightarrow x \notin P[L \cup M] . \end{aligned}$$

These are the two set inclusions of (b).

Since  $P[K]$  is a nonempty subset of *Actions* for every local state  $K$ , and since  $\#Actions = 2$ , the verification of (b) $\equiv$ (c) is a matter of case distinctions.  $\square$

For simplicity, we abbreviate  $P[\{s\}]$  and  $P[\{s, t\}]$  to  $P[s]$  and  $P[s, t]$  for all states  $s$  and  $t$ .

Following [1], we define relations  $>_P$  and  $=_P$  on *States* as follows

$$\begin{aligned} s >_P t &\equiv P[s] = P[s, t] \neq P[t] , \\ s =_P t &\equiv P[s] \cap P[t] = \emptyset \wedge P[s, t] = Actions . \end{aligned}$$

As a first justification of these definitions, we prove

(12) **Lemma.** Assume that  $P$  is *maximin* represented by utility function  $u$ . Let function  $m$  on states be given by  $m(s) = (\text{MIN } x :: u(x, s))$  where  $x$  ranges over the two actions.

(a) Assume  $s >_P t$  and  $\#P[s] = 1$ . Then  $m(s) < m(t)$  and  $\#P[t] \leq 2$ .

(b) Assume  $s >_P t$  and  $\#P[s] = 2$ . Then  $m(s) \leq m(t)$  and  $\#P[t] \leq 1$ .

(c) Assume  $s =_P t$ . Then  $m(s) = m(t)$  and  $\#P[s] = \#P[t] = 1$ .

*Proof.* We assume that  $x$  and  $y$  are the two elements of  $A$ . In all three cases, the assertion on  $\#P[t]$  (and  $\#P[s]$ ) is easy and can be left to the reader.

(a) We may assume that  $\{x\} = P[s] = P[s, t] \neq P[t]$ . It follows that  $m(s) = u(y, s) < u(x, s)$  and  $u_*(y, \{s, t\}) < u_*(x, \{s, t\})$  and  $m(t) = u(x, t) \leq u(y, t)$ . Therefore,  $u_*(y, \{s, t\}) < u_*(x, \{s, t\}) \leq u(x, t) \leq u(y, t)$ , which implies  $u(y, s) = u_*(y, \{s, t\})$  and hence  $m(s) < u_*(x, \{s, t\}) \leq u(x, t) = m(t)$ .

(b) We may assume that  $\{x\} = P[t]$ , while  $P[s] = P[s, t] = A$ . It follows that  $m(t) = u(y, t) < u(x, t)$  and  $m(s) = u(x, s) = u(y, s)$  and  $u_*(y, \{s, t\}) = u_*(x, \{s, t\})$ . Therefore,  $u_*(x, \{s, t\}) \leq u(y, t) < u(x, t)$  and hence  $m(s) = u(x, s) = u_*(x, \{s, t\}) = u_*(y, \{s, t\}) \leq u(y, t) = m(t)$ .

(c) We may assume that  $P[s] = \{x\}$  and  $P[t] = \{y\}$ . Then  $m(s) = u(y, s) < u(x, s)$  and  $m(t) = u(x, t) < u(y, t)$ , while  $u_*(y, \{s, t\}) = u_*(x, \{s, t\})$ . It follows that  $u_*(y, \{s, t\}) \leq u(x, t) < u(y, t)$  so that  $u(y, s) = u_*(y, \{s, t\}) = u_*(x, \{s, t\}) \leq u(x, t)$ , that is  $m(s) \leq m(t)$ . By symmetry, we get  $m(s) = m(t)$ .  $\square$

(13) **Corollary.** Assume that policy  $P$  is *maximin* representable. Then the pair of relations  $(>_P, =_P)$  on *States* is acyclic.

*Proof.* We use function  $m$  of Lemma (12). Since the set *States* is finite, there is a real number  $\varepsilon > 0$  such that  $m(s) < m(t)$  implies  $m(s) + 2 \cdot \varepsilon \leq m(t)$  for all pairs of states  $s$  and  $t$ . We now define function  $f$  on states by  $f(s) = m(s) - \varepsilon \cdot \#P[s]$ . It follows from Lemma (12) that, for all  $s$  and  $t$ ,

$$\begin{aligned} s >_P t &\Rightarrow f(s) < f(t) , \\ s =_P t &\Rightarrow f(s) = f(t) . \end{aligned}$$

Now Lemma (10) implies that the pair  $(>_P, =_P)$  is acyclic.  $\square$

Conversely, assume that the pair  $(>_P, =_P)$  is acyclic. We use the function *level* associated to relation  $T = (>_P) \circ (=P)^*$  and define the utility function *plev* on *AS-pairs* by

$$\begin{aligned} \text{plev}(x, s) &= \text{level}(s) \text{ if } x \notin P[s] \vee P[s] = A, \\ \text{plev}(x, s) &= \text{Large} \text{ if } x \in P[s] \neq A, \end{aligned}$$

where *Large* is a number larger than all values of *level*.

(14) **Lemma.** Assume that  $P$  is closed under unions and that relation  $(>_P, =_P)$  is acyclic. Let  $x \neq y$  in *Actions*.

(a) If  $(x, s) \prec_{\bullet}^P (y, t)$ , then  $\text{plev}(x, s) < \text{plev}(y, t)$ .

(b) If  $(x, s) \preceq_{\bullet}^P (y, t)$ , then  $\text{plev}(x, s) \leq \text{plev}(y, t)$ .

*Proof.* (a) By the definitions of  $P$  and  $\prec_{\bullet}^P$ , the assumption of (a) is equivalent to

$$(s = t \wedge x \notin P[t]) \vee (x \notin P[s, t] \wedge x \in P[t]) .$$

First, assume that  $s = t$ . Then we have  $x \notin P[s]$  and  $y \in P[s] \neq A$  and hence  $\text{plev}(x, s) = \text{level}(s) < \text{Large} = \text{plev}(y, t)$ . Secondly, assume  $s \neq t$ . Since  $P$  is closed under unions, we have  $x \notin P[s]$ . Therefore,  $P[s] = P[s, t] \neq P[t]$ , that is  $s >_P t$ . This implies that  $\text{plev}(x, s) = \text{level}(s) < \text{level}(t) \leq \text{plev}(y, t)$ .

(b) By the definitions of  $P$  and  $\preceq_{\bullet}^P$ , the assumption of (b) is equivalent to

$$(s = t \wedge y \in P[t]) \vee (y \in P[s, t] \wedge y \notin P[t]) .$$

If  $s = t$ , this means that  $x \notin P[t]$  or  $P[t] = A$ . This implies that  $\text{plev}(x, s) = \text{level}(s)$  and  $\text{plev}(y, t) \geq \text{level}(s)$ . Now, assume  $s \neq t$ . Then we have  $P[t] = \{x\}$ . Since  $P$  is closed under unions, there are three cases:

- (b1)  $P[s] = P[s, t] = \{y\} \neq P[t]$ ,
- (b2)  $P[s] = P[s, t] = \{x, y\} \neq P[t]$ ,
- (b3)  $P[s] = \{y\} \wedge P[s, t] = \{x, y\}$ .

In the cases (b1) and (b2), we have  $s >_P t$  and  $plev(x, s) = level(s)$ , so that  $plev(x, s) < level(t) \leq plev(y, t)$ . In case (b3), we have  $s =_P t$ , so that  $plev(x, s) = level(s) \leq level(t) \leq plev(y, t)$ .  $\square$

**Theorem G2A.** Assume  $\#Actions = 2$ . Policy  $P$  is *maximin* representable if and only if it is closed under unions and the pair of relations  $(>_P, =_P)$  on *States* is acyclic.

*Proof.* The necessity of closure under unions was seen above. The necessity of acyclicity was proved in Lemma (13). Conversely, assume that  $P$  is closed under unions and that  $(>_P, =_P)$  is acyclic. Then  $P$  is *maximin* represented by utility function  $plev$  introduced above because of Lemmas (7) and (14).  $\square$

*Remark.* Theorem 2 of [1] uses the same relations  $>_P$  and  $=_P$ . Instead of acyclicity it requires “closure under chains”, a weak form of transitivity, and “respect of domination”, a condition we do not encounter.  $\square$

## 7 Concluding remarks

The paper [1] was the starting point for a beautiful mathematical theory, but we needed predicate calculus to limit the number of case distinctions and enhance the reliability of the proofs. Our main results were as follows. First duality: an arbitrary policy  $P$  is *maximin* representable if and only if its dual is *maximax* representable. Theorem D: a deterministic policy  $P$  is *maximin* representable if and only if  $P$  is closed under unions and relation  $\prec_{\bullet}^P$  on action-state pairs is acyclic. Theorem G: an arbitrary policy  $P$  is *maximin* representable if and only if  $P$  is closed under unions and the pair of relations  $(\prec_{\bullet}^P, \preceq_{\bullet}^P)$  on action-state pairs is acyclic. Theorem DL: every deterministic policy that is *maximin* or *maximax* representable is Laplace representable.

We treated the binary cases (those with only two actions) as corollaries of the general cases. In particular, we confirmed the result of [1] that a deterministic binary policy is *maximin* representable if and only if it is closed under unions. We proved that an arbitrary binary policy  $P$  is *maximin* representable if and only if it is closed under unions and the pair of relations  $(>_P, =_P)$  on *States* is acyclic.

In the case with an arbitrary number of actions and only two states, we proved that a deterministic policy is *maximin* representable if and only if it satisfies the compromise rule.

Investigation of a number of nontrivial examples was indispensable for a fruitful development of the theory. Inspired by Theorem DL, we wonder whether the Laplace rules mentioned in section 2 are sufficient to imply Laplace representability.

## References

- [1] Brafman, R.I., Tennenholz, M.: An axiomatic treatment of three qualitative decision criteria. J. ACM **47** (2000) 452–482
- [2] Dijkstra, E.W., Scholten, C.S.: Predicate calculus and program semantics. Springer V. 1990.
- [3] Dubois, D., Fargier, H., Perny, P., Prade, H.: Qualitative decision theory: from Savage’s axioms to non-monotonic reasoning. J ACM ([www.acm.org/jacm/Upcoming.html](http://www.acm.org/jacm/Upcoming.html), to appear).
- [4] [www.haskell.org](http://www.haskell.org)
- [5] Lakatos, I.: Proofs and refutations: the logic of mathematical discovery. (eds. J. Worrall, E. Zahar) Cambridge University Press 1976.

- [6] Nurmi, H.: Voting procedures and how to deal with them. Springer V. 1999. ISBN 3-540-66236-7.
- [7] Savage, L.J.: The foundations of statistics. Dover Publ. inc., New York 1972.
- [8] Wooldridge, M.: Intelligent agents. In: Weiss, G. (ed.): Multiagent Systems. The MIT Press, April 1999, pp. 3–51.