THE REVIEW OF SYMBOLIC LOGIC
Volume 5, Number 1, March 2012

# PREFERENCE BASED ON REASONS

DANIEL OSHERSON
Department of Psychology, Princeton University
SCOTT WEINSTEIN
Department of Philosophy, University of Pennsylvania

**Abstract.** We describe a logic of preference in which modal connectives reflect reasons to desire that a sentence be true. Various conditions on models are introduced and analyzed.

**§1. Introduction.** Sometimes preferences are the result of identifiable reasons, as when you install a fire alarm for concern about safety. This suggests that studying reasons along with preference might illuminate both. An obstacle to such a project is the multitude of reasons for and against a given action that combine subliminally to yield a decision, as when you go ahead with the fire alarm despite the bother, cost, and false alerts; the underlying calculus of reason aggregation seems largely hidden from introspection. The centrality of reasons to action and rationality is nonetheless sufficient motive to persevere in their analysis despite the difficulty. In this spirit, we here advance a modal logic in which different reasons for a preference can be aggregated in various ways.

Our inquiry is preceded by several studies of the logic of preference, beginning with von Wright (1963). More contemporary work includes systems designed to elucidate the interaction between choice and epistemic possibility (see Lang *et al*., 2003; van Benthem *et al*., 2009). Of particular relevance is Liu (2008, chap. 3). This work introduces "priorities" (which function like reasons in the present setting) that are ordered by importance, and integrated into a formal language of preference and belief. Several ways of extracting preferences from priorities are explored. The interplay of preferences and beliefs is also analyzed, along with the impact of updating belief and preference. Liu's work is closest to the approach taken here inasmuch as it develops a modal language and associated semantics. The conceptual framework is nonetheless different from ours, as will become clear presently. Another fruitful perspective on the integration of preferences issues from the graph-theoretic approach advanced in Andréka *et al*. (2002); different graphs represent alternative orderings of the alternatives in play, and might be considered separate reasons for choice among them. Within a yet different tradition, multiattribute utility theory (Keeney & Raiffa, 1993) bears directly on reason aggregation through the combination of utilities based on separate dimensions. The theory has revealed exact conditions under which aggregation can proceed additively but it does not explore the logical structure of reasons and preference, as we shall do here. The insight achievable by formal analysis of reasons is illustrated by Dietrich & List (2009). These authors demonstrate a representation theorem relating choice to the respective bundles of reasons that apply to the options in play; the axioms needed for their result are remarkably weak. Several issues are thereby

Received: September 17, 2011.



doi:10.1017/S1755020311000244

clarified, among them the significance of combining reasons (their analysis rests not on individual reasons but on sets of them).

To keep the present project manageable, conceptual issues about the nature of reasons and their role in rational discourse will be set aside. An entry to this literature is provided by Dietrich & List (2009), and sustained discussion is available in Pettit (2002). Of course, the reasons that come to mind are not necessarily those that govern choice (see, e.g., Messick, 1985; Haidt, 2001). Our theory is indifferent to this distinction but it will be more natural to limit examples to conscious effective reasons. The case of the fire alarm serves to convey the character of our theory. Specifically, we picture an agent who imagines a world *that resembles the actual one* but with a fire alarm, and another world (possibly his own) without one. The agent then compares the two worlds according to various utility scales (one that measures safety, another cost, and so forth), as well as a distinct utility scale that takes all the individual scales into account. Our formalism is designed to capture this picture.

We proceed by first introducing the language under investigation. Informal glosses for some of its formulas will clarify the ideas in play. Next the semantics of our logic is presented, followed by consideration of subclasses of models that meet various conditions. We then turn to decidability issues. A discussion of open questions is provided at the end.

**§2. Language.** The present section introduces a family of modal languages, and discusses the intended meaning of the modality. A *language of reason-based preference* is determined by its *signature*, which consists of:

(a) a nonempty set $\mathbb{P}$ of propositional variables
(b) a nonempty collection $\mathbb{S}$ of nonempty subsets of $\mathbb{N}$ (the set $\{0, 1, \dots\}$ of natural numbers)

The language of reason-based preference determined by signature $(\mathbb{P}, \mathbb{S})$ is denoted $\mathcal{L}(\mathbb{P}, \mathbb{S})$, and is built from the following symbols.

(a) the set $\mathbb{P}$ of propositional variables
(b) the unary connective $\neg$
(c) the binary connective $\wedge$
(d) for every set $X \in \mathbb{S}$, the binary connective $\succeq_X$
(e) the two parentheses

Formulas are defined inductively via:

$$p \in \mathbb{P} \quad | \quad \neg\varphi \quad | \quad (\varphi \wedge \psi) \quad | \quad (\varphi \succeq_X \psi) \text{ for } X \in \mathbb{S}.$$

Moreover, we rely on the following abbreviations.

$$\begin{aligned}
(\varphi \vee \psi) &\quad \text{for} \quad \neg(\neg\varphi \wedge \neg\psi) \\
(\varphi \to \psi) &\quad \text{for} \quad (\neg\varphi \vee \psi) \\
(\varphi \leftrightarrow \psi) &\quad \text{for} \quad ((\varphi \to \psi) \wedge (\psi \to \varphi)) \\
(\varphi \succeq_{1\dots k} \psi) &\quad \text{for} \quad (\varphi \succeq_{\{1\dots k\}} \psi) \\
(\varphi \succ_X \psi) &\quad \text{for} \quad (\varphi \succeq_X \psi) \wedge \neg(\psi \succeq_X \varphi) \\
(\varphi \approx_X \psi) &\quad \text{for} \quad (\varphi \succeq_X \psi) \wedge (\psi \succeq_X \varphi)
\end{aligned}$$

$$(\varphi \preceq_X \psi) \quad \text{for} \quad (\psi \succeq_X \varphi)$$
$$(\varphi \prec_X \psi) \quad \text{for} \quad (\psi \succ_X \varphi)$$
$$\top \quad \text{for} \quad (p \rightarrow p)$$
$$\bot \quad \text{for} \quad \neg\top$$

The formula $\varphi \succ_1 \psi$ is to be understood along the following lines. Fix an agent $\mathcal{A}$ whose reasoning is at issue. Let $u_1$ be a utility scale that reflects some dimension of interest to $\mathcal{A}$. Then $\varphi \succ_1 \psi$ is true just in case:

> $\mathcal{A}$ envisions a situation in which $\varphi$ is true and that otherwise differs little from his actual situation (if $\varphi$ is already true then $\mathcal{A}$'s actual situation may well be the one he envisions). Likewise, $\mathcal{A}$ envisions a second situation that is like his actual situation except that $\psi$ is true. Finally, the utility according to $u_1$ of the first imagined situation exceeds that of the second.

In the fire alarm example, $\mathcal{A}$ envisions his home with a new fire alarm, but with the same furniture, cat, and fireplace as before. Home with no fire alarm is the actual situation, hence especially easy to envision. If $u_1$ measures safety, and $p$ is "$\mathcal{A}$ will purchase a fire alarm" then $p \succ_1 \neg p$ holds inasmuch as the alarm improves safety. (Since $\top$ is also true in $\mathcal{A}$'s situation, $p \succ_1 \neg p$ is materially equivalent to $p \succ_1 \top$.) If $\mathcal{A}$ is short on cash, and $u_2$ reflects finances then $p \prec_2 \neg p$ is true, whereas the status of $p \succ_{1,2} \neg p$ depends on the manner in which utilities are aggregated (e.g., averaging, minimum, etc.). More generally, we allow preferences $\varphi \succ_X \psi$ between arbitrary formulas $\varphi, \psi$ in view of the (possibly multiple) utilities in $X \in \mathbb{S}$. The formula $\varphi \succ_X \psi$ thus represents $\mathcal{A}$'s preference for $\varphi$ over $\psi$ when $\mathcal{A}$ brings to mind just the reasons indexed in $X$. If $\bigcup \mathbb{S} \in \mathbb{S}$ then preference *tout court* for $\varphi$ over $\psi$ is represented by $\varphi \succ_{\bigcup \mathbb{S}} \psi$, that is, taking account of all reasons in play.

Of course, the greater utility of a given situation compared to another is just one way of expressing a reason for preferring the former to the latter. More generality can be achieved by representing each kind of reason by an arbitrary binary relation over situations, instead of insisting on numerical comparisons of cardinal utility. Recourse to such relations will be raised again in the Discussion section. For now, we develop our theory in the context of utility, with the expectation that most readers will find this setting conceptually familiar.

If our agent is presumed to be *moral* then reasons are meant, very roughly, to be *good* (at least, not *bad*). Morality will here be left unexplored, however. Instead, $\mathcal{A}$ is conceived as logically empowered but otherwise like the rest of us. Also notice how little any of this has to do with *reasons to believe* (except for odd cases like being rewarded for reaching genuine religious conviction). Only reasons for preference will be at issue. There is nonetheless one connection to belief that bears comment.

The appeal to situations that differ minimally from the actual one, except for satisfying a given formula, is familiar from well-known theories of counterfactual conditionals (Stalnaker, 1968; Lewis, 1973). It thus risks bedevilment from a similar range of cases. Suppose, for example, that $p$ is "Winter ends a little earlier than last year." Then too many $p$-worlds offer themselves as alternatives to the actual world (since the set of shorter winters has no member closest to last year's winter). The present endeavor, however, may not be as vulnerable as the earlier one to such cases. For it here suffices that the reasoning agent bring to mind a cognitively salient situation that satisfies the formula in question (e.g., winter a week shorter), not necessarily the maximally similar one. Indeed, the agent may not be prepared to identify the maximally similar $p$-world, or even to understand such an idea. Consistent with this relaxed attitude, to each consistent proposition our semantics assigns a world that represents life were the proposition true, where the choice of world may depend on the agent's current position. Some constraints on the choice will be examined, but otherwise the reasoning agent is on his own. We take all this to be a rough idealization of what happens in actual decision making. One imagines an alternative situation that satisfies the proposition at issue, then evaluates it along various dimensions (i.e., utility scales).

The utility scales that determine the truth of modal formulas are intended to measure the impact on choice of specific considerations, for example, cost, health, professional advancement. Because deliberation is assumed to transpire in a single mind (the agent's), aggregation of different scales into an overall value seems feasible; indeed, people do it all the time. For simplicity, the scales express *expected* utilities, that is, with probabilities already factored in. Thus, the safety improvements envisioned from installing a fire alarm already integrate the agent's confidence that the device will work as advertised.

Even when utility scales are kept separate, languages of reason-based preference allow interesting interactions. For an illustration, first observe that $\varphi \succ_i \top$ means (roughly) that the $u_i$-utility of the envisioned $\varphi$-world exceeds that of the actual world. Now consider:

$$(p \succ_1 \top) \succ_2 \top.$$

This says that the agent has a $u_2$-reason for there being a $u_1$-reason in favor of $p$. For example, let $p$ be the assertion that you buy a low-power automobile. Let $u_2$-utility be pecuniary: $u_2(w_1) > u_2(w_2)$ iff you have more cash in $w_1$ compared to $w_2$. Let $u_1$-utility reflect personal safety: $u_1(w_1) > u_1(w_2)$ iff you incur less risk traveling in $w_1$ than in $w_2$. Then the formula asserts that it's in your financial interest that your buying a low-power automobile is in your safety interest—which might well be true inasmuch as low-power vehicles are cheaper.

We conclude this section with another illustration of the interaction of individual utility scales. Consider:

$$\neg q \succ_1 (p \succ_2 q).$$

This says that the agent $u_1$-prefers that $q$ be false rather than $u_2$-prefer $p$ over $q$. For example, let $q$ be the assertion that your brother runs for mayor, and let $p$ be that Miss Smith (no relation) also runs. Let $u_1$-utility measure family pride, and let $u_2$-utility measure political value to an ailing municipality. Then the formula asserts that from the point of view of family pride, you'd rather that your brother not run for mayor than that Miss Smith be the superior candidate.

**§3. Semantics.** We now provide a formal semantics designed to capture the intuitive picture elaborated in the preceding section. Several preliminary concepts are needed. Fundamental is the choice of a nonempty set $\mathbb{W}$ to embody the imaginative possibilities ("worlds") available to an agent in the course of practical deliberation. Subsets of $\mathbb{W}$ are called *propositions*. As discussed above, given a nonempty proposition $A$ and a world $w$, an agent envisions a salient alternative to $w$ among the worlds in $A$. (If $w \in A$ then the "alternative" might be $w$ itself.) We formalize this idea as follows.

DEFINITION 3.1. *A selection function $s$ over $\mathbb{W}$ is a mapping from $\mathbb{W} \times \{A \subseteq \mathbb{W} \mid A \neq \emptyset\}$ to $\mathbb{W}$ such that for all $w \in \mathbb{W}$ and $\emptyset \neq A \subseteq \mathbb{W}$, $s(w, A) \in A$.*

Thus, $s(w, A)$ is a choice of world to represent $A$, where the choice depends on $w$. (The idea is that $s$ chooses a member of $A$ that is similar to $w$.)

Next, recall that each world can be evaluated according to various utility scales, each involving one or more dimensions of value. All the scales are indexed by members of $\mathbb{S}$.

DEFINITION 3.2. *A utility function $u$ over $\mathbb{W}$ and $\mathbb{S}$ is a mapping from $\mathbb{W} \times \mathbb{S}$ to $\Re$ (the reals).*

For $w \in \mathbb{W}$ and $\{i\}, X \in \mathbb{S}$, we write $u(w, \{i\})$ as $u_i(w)$, and $u(w, X)$ as $u_X(w)$.

Let $\mathbb{P}$ be a nonempty set of propositional variables. Our last preliminary is the assignment of a proposition to each variable in $\mathbb{P}$.

DEFINITION 3.3. *A truth-assignment (over $\mathbb{W}$ and $\mathbb{P}$) is a mapping from $\mathbb{P}$ to the power set of $\mathbb{W}$.*

For a truth-assignment $t$, the idea is that $p \in \mathbb{P}$ is true in $w \in \mathbb{W}$ just in case $w \in t(p)$ (and otherwise false). This is all we need to introduce models.

DEFINITION 3.4. *A model for a signature $(\mathbb{P}, \mathbb{S})$ is a quadruple $(\mathbb{W}, s, u, t)$ where*

*(a) $\mathbb{W}$ is a nonempty set of worlds;*
*(b) $s$ is a selection function over $\mathbb{W}$;*
*(c) $u$ is a utility function over $\mathbb{W}$ and $\mathbb{S}$;*
*(d) $t$ is a truth-assignment over $\mathbb{W}$ and $\mathbb{P}$.*

It remains to specify the proposition (set of worlds) expressed by a formula $\varphi$ in a model $\mathcal{M}$. This proposition is denoted $\varphi[\mathcal{M}]$, and defined inductively as follows.

DEFINITION 3.5. *Let signature $(\mathbb{P}, \mathbb{S})$, $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$, and model $\mathcal{M} = (\mathbb{W}, s, u, t)$ for $(\mathbb{P}, \mathbb{S})$ be given.*

*(a) If $\varphi \in \mathbb{P}$ then $\varphi[\mathcal{M}] = t(\varphi)$.*
*(b) If $\varphi$ is the negation $\neg\theta$ then $\varphi[\mathcal{M}] = \mathbb{W} \setminus \theta[\mathcal{M}]$.*
*(c) If $\varphi$ is the conjunction $(\theta \wedge \psi)$ then $\varphi[\mathcal{M}] = \theta[\mathcal{M}] \cap \psi[\mathcal{M}]$.*
*(d) If $\varphi$ has the form $(\theta \succeq_X \psi)$ for $X \in \mathbb{S}$, then $\varphi[\mathcal{M}] = \emptyset$ if either $\theta[\mathcal{M}] = \emptyset$ or $\psi[\mathcal{M}] = \emptyset$. Otherwise:*

$$\varphi[\mathcal{M}] = \{w \in \mathbb{W} \mid u_X(s(w, \theta[\mathcal{M}])) \geq u_X(s(w, \psi[\mathcal{M}]))\}.$$

Observe that $(\theta \succeq_X \psi)[\mathcal{M}]$ is defined to be empty if there is no world that satisfies $\theta$ or none that satisfies $\psi$. Thus, we read $(\theta \succeq_X \psi)$ with existential import ("the $\theta$-world is weakly $X$-better than the $\psi$-world," where the definite description is Russellian). In the nontrivial case, let $A \neq \emptyset$ be the proposition expressed by $\theta$ in $\mathcal{M}$, and $B \neq \emptyset$ the one expressed by $\psi$. Then (intuitively) world $w$ satisfies $(\theta \succeq_X \psi)$ in $\mathcal{M}$ iff the world selected from $A$ as closest to $w$ has utility no less than that of the world selected from $B$ as closest to $w$. A word of caution: the existential requirement on the truth of $(\theta \succeq_X \psi)$ allows $\neg(\theta \succeq_X \psi)[\mathcal{M}] \neq (\theta \prec_X \psi)[\mathcal{M}]$. Indeed, if $\theta[\mathcal{M}] = \emptyset$ then $\neg(\theta \succeq_X \psi)[\mathcal{M}] = \mathbb{W}$ but $(\theta \prec_X \psi)[\mathcal{M}] = \emptyset$.

The following definition imports standard terminology and notation to the present context.

DEFINITION 3.6. *Let $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ and model $\mathcal{M} = (\mathbb{W}, s, u, t)$ for $(\mathbb{P}, \mathbb{S})$ be given.*

*(a) $\mathcal{M}$ satisfies $\varphi$ just in case $\varphi[\mathcal{M}] \neq \emptyset$.*

*(b)* $\varphi$ *is valid in* $\mathcal{M}$ *just in case* $\varphi[\mathcal{M}] = \mathbb{W}$.
*(c)* $\varphi$ *is valid just in case* $\varphi$ *is valid in every model.*
*(d)* $\varphi$ *is valid in a given class* $C$ *of models just in case* $\varphi$ *is valid in every model of* $C$.

We use related expressions (like "satisfiable") in the obvious way. It is noteworthy that our language allows expression of the global modality (see Blackburn *et al.*, 2001, sec. 2.1). Choose any $X \in \mathbb{S}$, and for $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ let:

$$\Box\varphi \stackrel{\text{def}}{=} \neg(\neg\varphi \succeq_X \neg\varphi) \quad \text{and} \quad \Diamond\varphi \stackrel{\text{def}}{=} (\varphi \succeq_X \varphi). \tag{3.7}$$

Then unwinding clause Definition 3.5d of our semantic definition yields:

PROPOSITION 3.8. *For all* $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ *and models* $\mathcal{M} = (\mathbb{W}, s, u, t)$*:*

*(a)* $\Box\varphi[\mathcal{M}] \neq \emptyset$ *iff* $\Box\varphi[\mathcal{M}] = \mathbb{W}$ *iff* $\varphi[\mathcal{M}] = \mathbb{W}$.
*(b)* $\Diamond\varphi[\mathcal{M}] \neq \emptyset$ *iff* $\Diamond\varphi[\mathcal{M}] = \mathbb{W}$ *iff* $\varphi[\mathcal{M}] \neq \emptyset$.

It follows from Proposition 3.8 that the axioms of S5 are valid for $\Box$ and $\Diamond$. Other valid formulas of our language include the following (proofs are easy). For all $X \in \mathbb{S}$, and $\varphi, \psi, \theta \in \mathcal{L}(\mathbb{P}, \mathbb{S})$:

$\models ((\varphi \succeq_X \psi) \wedge (\psi \succeq_X \theta)) \rightarrow (\varphi \succeq_X \theta)$

$\models (\Diamond\varphi \wedge \Diamond\psi) \rightarrow ((\varphi \succeq_X \psi) \vee (\psi \succeq_X \varphi))$

$\models (\Diamond\varphi \wedge \Diamond\psi) \leftrightarrow (\neg(\varphi \succeq_X \psi) \leftrightarrow (\psi \succ_X \varphi))$

$\models \neg(\bot \succeq_X \varphi)$ and $\models \neg(\varphi \succeq_X \bot)$

$\models \Diamond\varphi \rightarrow (\varphi \approx_X \psi)$ if $\varphi$ and $\psi$ are equivalent.

The next two sections introduce subclasses of structures which conform to various hypotheses; we explore the logical principles validated thereby. The hypotheses considered in Section §4 are called *frame properties* because their definition depends on just the worlds and selection function of a model, that is, on just its "frame." The remaining hypotheses (Section §5) involve utility and the assignment of propositions to variables. Several of the properties discussed below have already appeared within order-theoretic approaches to preference, for example, in Levi (1986, chap. 6).

**§4. Stronger theories based on frame properties of models.** For this section, let model $\mathcal{M} = (\mathbb{W}, s, u, t)$ have signature $(\mathbb{P}, \mathbb{S})$.

***4.1. Reflexivity.*** If a world $w$ satisfies a formula $\varphi$ then the "nearest" $\varphi$-world is intuitively $w$ itself. This condition is not imposed on selection functions by Definition 3.1 but can be added as follows.

DEFINITION 4.1. $\mathcal{M}$ *is reflexive just in case for all* $w \in \mathbb{W}$ *and* $A \subseteq \mathbb{W}$*, if* $w \in A$*, then* $s(w, A) = w$.

The formula exhibited in the following proposition illustrates the impact of reflexivity. It says that a given proposition is at least as good as the status quo or its negation is.

PROPOSITION 4.2. *Let* $\varphi$ *be* $(p \succeq_X \top) \vee (\neg p \succeq_X \top)$. *Then* $\varphi$ *is invalid but valid in the class of reflexive models.*

*Proof.* To verify the invalidity of $\varphi$, suppose that $\mathbb{W} = \{w_0, w_1, w_2\}$, $t(p) = \{w_0, w_1\}$, $s(w_0, \{w_0, w_1\}) = s(w_0, p[\mathcal{M}]) = w_1$, $s(w_0, \{w_2\}) = s(w_0, \neg p[\mathcal{M}]) = w_2$, $s(w_0, \mathbb{W}) = s(w_0, \top[\mathcal{M}]) = w_0$, and $u_X(w_0) > u_X(w_1), u_X(w_2)$. Then it is easy to see that $w_0 \notin \varphi[\mathcal{M}]$ hence $\varphi$ is not valid.

On the other hand, suppose that $\mathcal{M}$ is reflexive, and let $w_0 \in \mathbb{W}$. Then either $w_0 \in p[\mathcal{M}]$ or $w_0 \in \neg p[\mathcal{M}]$, say the former (the other case is parallel). By reflexivity, $s(w_0, p[\mathcal{M}] = w_0$. Likewise, $w_0 \in \top[\mathcal{M}] = \mathbb{W}$, so again by reflexivity, $s(w_0, \top[\mathcal{M}] = w_0$. Since $u_X(w_0) \geq u_X(w_0)$, $w_0 \in \varphi[\mathcal{M}]$. □

Reflexivity entails that some formulas are satisfied only by infinite models.

PROPOSITION 4.3. *There is $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ such that $\varphi$ is satisfied by some infinite reflexive model but by no finite reflexive model.*

*Proof.* Let $\varphi$ be the conjunction of the following formulas.

(4.4) *(a)* $\Box(p \to (p \prec_X \neg p))$
*(b)* $\Box(\neg p \to (\neg p \prec_X p))$

It is easy to verify that $\varphi$ is satisfied by a model whose worlds form an $\omega$-sequence when ordered by $u_X$, and which alternate between satisfying $p$ and $\neg p$. On the other hand, suppose for a contradiction that $\varphi$ is satisfied by finite model $\mathcal{M} = (\mathbb{W}, s, u, t)$. Then some $w_0 \in \mathbb{W}$ has maximum $u_X$ utility. Suppose that $w_0$ satisfies $p$ (the other case is parallel). Then (4.4)a and Reflexivity imply that there is $w_1 \in \mathbb{W}$ satisfying $\neg p$ such that $u_X(w_0) < u_X(w_1)$. This contradicts the choice of $w_0$ as having maximum $u_X$ utility. □

***4.2. Regularity.*** If you think that living in Boston is most similar to your current situation among the set of all addresses in New England then shouldn't you think that living in Boston is most similar to your current situation among the set of all addresses in Massachusetts? A similar principle is standardly applied to choice (Sen, 1971) even though its violation has been documented in several empirical studies (e.g., Payne & Puto, 1982; Tentori *et al.*, 2001). In the present setting, we are led to the following constraint on selection.

DEFINITION 4.5. *$\mathcal{M}$ is regular just in case for all $w \in \mathbb{W}$, nonempty $A \subseteq B \subseteq \mathbb{W}$, and $w_1 \in A$: If $s(w, B) = w_1$ then $s(w, A) = w_1$.*

Regularity validates the formula appearing in the next proposition. An instance is this: If buying either a Ford or a Chevy makes more sense than buying a Toyota then either it makes more sense to buy a Ford than a Toyota, or it makes more sense to buy a Chevy than a Toyota (or both).

PROPOSITION 4.6. *Let $\varphi$ be $((p \vee q) \succ_X r) \to ((p \succ_X r) \vee (q \succ_X r))$. Then $\varphi$ is invalid but valid in the class of regular models.*

*Proof.* A countermodel for $\varphi$ is easy to devise. To show validity in the regular models, suppose that $\mathcal{M}$ is regular, and let $w \in ((p \vee q) \succ_X r)[\mathcal{M}]$ be given. Then there are $w_1, w_2 \in \mathbb{W}$ with:

(4.7) *(a)* $w_1 = s(w, (p \vee q)[\mathcal{M}])$,
*(b)* $w_2 = s(w, r[\mathcal{M}])$, *and*
*(c)* $u_X(w_1) > u_X(w_2)$.

By (4.7)a, either $w_1 \in t(p)$ or $w_1 \in t(q)$, say the latter (the other case is parallel). Since $q[\mathcal{M}] \subseteq (p \vee q)[\mathcal{M}]$, it follows from regularity that $w_1 = s(w, q[\mathcal{M}])$. In view of (4.7)b and c, $w \in (q \succ_X r)[\mathcal{M}]$. □

The combination of reflexivity and regularity validates the following formula, which exhibits modal embedding.

$$((p \prec_1 \top) \succ_2 (q \prec_1 \top)) \rightarrow (\neg p \succ_2 \neg q). \qquad (3.7)$$

For an instance, suppose that $p, q$ represent plans for new shopping malls, and that $u_1, u_2$ measure their political and ecological interest, respectively. Then (3.7) asserts: If it is ecologically better for $p$ than for $q$ to politically backfire then abstaining from $p$ is ecologically better than abstaining from $q$.

PROPOSITION 4.9. *Formula (3.7) is invalid but valid in the class of models that are reflexive and regular.*

*Proof.* The invalidity of (3.7) is easy to verify. For validity in the class of reflexive regular models, let reflexive regular model $\mathcal{M} = (\mathbb{W}, s, u, t)$ and $w \in \mathbb{W}$ be given. Suppose that:

(4.10) $w \in ((p \prec_1 \top) \succ_2 (q \prec_1 \top))[\mathcal{M}]$.

We must show:

(4.11) $w \in (\neg p \succ_2 \neg q)[\mathcal{M}]$.

By (4.10), there are $w_1, w_2 \in \mathbb{W}$ with:

(4.12) *(a)* $\succ_1 \top)\mathcal{M}) \neq \emptyset$,
*(b)* $w_1 = s(w, (p \prec_1 \top)[\mathcal{M}])$,
*(c)* $w_2 = s(w, (q \prec_1 \top)[\mathcal{M}])$,
*(d)* $u_2(w_1) > u_2(w_2)$.

By reflexivity, it is easy to verify:

(4.13) *(a)* $(p \prec_1 \top)[\mathcal{M}] \subseteq \neg p[\mathcal{M}]$,
*(b)* $(q \prec_1 \top)[\mathcal{M}] \subseteq \neg q[\mathcal{M}]$.

So by (4.12)a,b and (4.13), we have $\neg p[\mathcal{M}] \neq \emptyset$ and $\neg q[\mathcal{M}] \neq \emptyset$. Hence there are $w_1^*, w_2^* \in \mathbb{W}$ with:

(4.14) *(a)* $w_1^* = s(w, \neg p[\mathcal{M}])$,
*(b)* $w_2^* = s(w, \neg q[\mathcal{M}])$.

But by (4.12)a, (4.13)a, (4.14)a, and regularity, $w_1^* = w_1$. Likewise, by (4.12)b, (4.13)b, (4.14)b, and regularity, $w_2^* = w_2$. Thus, (4.12)c implies $u_2(w_1^*) > u_2(w_2^*)$ which together with (4.14) yields (4.11). □

An alternative formulation of regularity is given by the following definition and proposition.

DEFINITION 4.15. *$\mathcal{M}$ is locally lexicographic just in case for every $v \in \mathbb{W}$ there is a well order $R_v$ of $\mathbb{W}$ such that for all $A \subseteq \mathbb{W}$, $s(v, A)$ is the $R_v$-least member of $A$.*

PROPOSITION 4.16. *A model is regular if and only if it is locally lexicographic.*

*Proof.* The right-to-left direction is immediate. For the other direction we proceed as follows. Suppose $(\mathbb{W}, s, u, t)$ is regular and $v \in \mathbb{W}$. Define by transfinite recursion a well-ordering $R_v$ of $\mathbb{W}$ as follows. For every ordinal $\alpha$ let $w_\alpha = s(v, \mathbb{W} - \{w_\beta | \beta < \alpha\})$. [So $w_0 = s(v, \mathbb{W})$.] Now let $s'$ be the following locally lexicographic selector. For every nonempty proposition $A$, $s'(v, A) = w_\alpha$ where $\alpha$ is the least ordinal $\gamma$ such that $w_\gamma \in A$.

It suffices to show that for every nonempty proposition $A$, $s(v, A) = s'(v, A)$. Let $s'(v, A) = w_\alpha$. By our construction, $A \subseteq (\mathbb{W} - \{w_\beta | \beta < \alpha\})$. By definition, $s(v, \mathbb{W} - \{w_\beta | \beta < \alpha\}) = w_\alpha$. But $w_\alpha \in A$, so by the regularity of $s$, $s(v, A) = w_\alpha$. □

***4.3. Metric selection.*** A natural way to express the thought that selection chooses "nearby" worlds is via a metric on $\mathbb{W}$. The following definition gives form to this idea.

DEFINITION 4.17. *A selection function $s$ over $\mathbb{W}$ is metrizable just in case there is a metric $d : \mathbb{W} \times \mathbb{W} \to \Re$ such that for all $w \in \mathbb{W}$ and $\emptyset \neq A \subseteq \mathbb{W}$, $s(w, A)$ is the unique $d$-closest member of $A$ to $w$. $\mathcal{M} = (\mathbb{W}, s, u, t)$ is metric just in case $s$ is metrizable.*

Note that $s$ is metrizable only if $d$-closest worlds exist (there are no chains of worlds ever $d$-closer to $w$).

EXAMPLE 4.18. *The following model is reflexive and regular but not metric. Let $\mathbb{W} = \{w_0, w_1, w_2\}$, and let $s$ be the (unique) reflexive selection function that satisfies:*

$$\begin{aligned} s(w_0, \{w_1, w_2\}) &= w_1 \\ s(w_1, \{w_0, w_2\}) &= w_2 \\ s(w_2, \{w_0, w_1\}) &= w_0 \end{aligned}$$

*It is easy to see that $s$ must also be regular. But $s$ is not metrizable since otherwise the three selections imply $d(w_0, w_1) < d(w_0, w_2)$, $d(w_1, w_2) < d(w_0, w_1)$, and $d(w_0, w_2) < d(w_1, w_2)$, for some distance metric $d$. But these inequalities yield $d(w_0, w_2) < d(w_0, w_2)$, contradiction.*

As a straightforward consequence of Definition 4.17, we have:

PROPOSITION 4.19. *Every metric model is reflexive and regular. Also, every metric model is countable.*

More consequentially, the next theorem shows that notwithstanding the fact that the metric models are a proper subset of the reflexive and regular models, metric validity reduces to reflexive regular validity. See the Appendix for proof.

THEOREM 4.20. *Every formula which is valid in the class of metric models, is valid in the class of reflexive and regular models.*

***4.4. Lexicographic ordering.*** Let us consider another way to strengthen regularity.

DEFINITION 4.21. *$\mathcal{M}$ is lexicographic just in case there is a well order $R$ of $\mathbb{W}$ such that for all $w \in \mathbb{W}$ and $A \subseteq \mathbb{W}$, $s(w, A)$ is the $R$-least member of $A$.*

Thus, the definition provides a uniform version of the local lexicographic property formulated in Definition 4.15. All lexicographic models are regular but not vice versa. The added constraint imposed by the lexicographic property validates some additional formulas as shown by the following proposition (whose proof is elementary).

PROPOSITION 4.22. *Let $\varphi$ be $(p \succeq_X q) \rightarrow \Box(p \succeq_X q)$. Then $\varphi$ is false in some regular model but valid in the class of lexicographic models.*

A natural generalization of lexicographic ordering may be defined as follows.

DEFINITION 4.23. *A selection function $s$ over $\mathbb{W}$ is proposition driven just in case for all $w_1, w_2 \in \mathbb{W}$ and $\emptyset \neq A \subseteq \mathbb{W}$, $s(w_1, A) = s(w_2, A)$.*

That is, proposition driven selection functions ignore their first arguments. Lexicographic ordering implies proposition drivenness; the next proposition shows the former to be a stronger condition than the latter.

PROPOSITION 4.24. *There is a formula satisfiable in the class of proposition driven models but not in the class of lexicographic models.*

*Proof.* Let $\varphi$ be the conjunction of the following formulas.

$$(p \vee q) \succ_X \top \qquad p \prec_X \top \qquad q \prec_X \top.$$

It is easy to verify that no regular model satisfies $\varphi$ but that some proposition driven model does. Since lexicographic ordering implies regularity, the proposition follows immediately. □

The next proposition is a corollary to Proposition 4.16.

PROPOSITION 4.25. *A model is proposition driven and regular if and only if it is lexicographic.*

**§5. Stronger theories that are not based on frames.** We now consider properties of models that cannot be defined just in terms of $(\mathbb{W}, s)$, the frame of a model. The background signature $(\mathbb{P}, \mathbb{S})$ may thus be expected to interact with the validity of formulas in classes of models specified by these properties.

***5.1. Proximity.*** Intuitively, a selection function applied to a world $w$ and nonempty proposition $A$ should pick a member $w_1$ of $A$ that is "near" or "similar" to $w$. One way to articulate this idea is to require that the two worlds differ minimally in the sets of propositional variables that each makes true. The following notation helps us formulate this idea. For $w \in \mathbb{W}$, let $t^{-1}(w) = \{p \in \mathbb{P} \mid w \in t(p)\}$. That is, $t^{-1}(w)$ is the set of propositional variables that $\mathcal{M}$ satisfies at $w$. For sets $S, T$, let $S \triangle T$ denote their symmetric difference $(S \setminus T) \cup (T \setminus S)$. Then the idea of selecting "nearby worlds" can be rendered as follows.

DEFINITION 5.1. *$\mathcal{M}$ is proximal just in case the following condition is met, for all $w \in \mathbb{W}$ and all nonempty propositions $A \subseteq \mathbb{W}$.*

> *If $s(w, A) = w_1$ then there is no $w_2 \in A$ such that $t^{-1}(w) \triangle t^{-1}(w_2) \subset t^{-1}(w) \triangle t^{-1}(w_1)$.*

For example, suppose that $t^{-1}(w) = \{p, q\}$, $t^{-1}(w_1) = \{p, r\}$, and $t^{-1}(w_2) = \{p, q, r\}$. Let $A = \{w_1, w_2\}$. Then $s$ violates proximity if $s(w, A) = w_1$ since $t^{-1}(w) \triangle t^{-1}(w_2) = \{r\} \subset \{q, r\} = t^{-1}(w) \triangle t^{-1}(w_1)$.

For the next two propositions, we rely on an hypothesis about our signature, namely, that $\mathbb{P} = \{p, q, r\}$. In conjunction with regularity, proximity validates a formula reminiscent of the *sure thing principle* (Savage, 1954).

PROPOSITION 5.2. *Let $\varphi$ be*

$$(((p \wedge r) \succ_X (q \wedge r)) \wedge ((p \wedge \neg r) \succ_X (q \wedge \neg r))) \rightarrow (p \succ_X q).$$

*Then $\varphi$ is invalid in the class of regular and in the class of proximal models but valid in the class of models that are both regular and proximal.*

An instance of $\varphi$ is the following. If one has better reason to vacation in Florence during an Italian transport strike than to vacation in Rome during such a transport strike, and if one has better reason to vacation in Florence with no transport strike than to vacation in Rome with no such strike then one has better reason to vacation in Florence than in Rome.

*Proof of Proposition 5.2.* Construction of the needed countermodels is left for the reader. Suppose that $\mathcal{M}$ is regular and proximal with $w \in \mathbb{W}$. Either $w \in t(r)$ or $w \notin t(r)$; assume the former (the argument is parallel in the other case). There is nothing left to prove unless the following statements are true [since otherwise the left conjunct in the antecedent of $\varphi$ is false; see Definition 3.5d].

(5.3) *(a)* $t(p) \cap t(r) \neq \emptyset$
*(b)* $t(q) \cap t(r) \neq \emptyset$.

By (5.3), $p[\mathcal{M}] \neq \emptyset$ and $q[\mathcal{M}] \neq \emptyset$. So let $w_1, w_2 \in \mathbb{W}$ be such that:

(5.4) *(a)* $w_1 = s(w, p[\mathcal{M}])$
*(b)* $w_2 = s(w, q[\mathcal{M}])$.

Since $w \in t(r)$, (5.3)a, (5.4)a, and proximity imply $w_1 \in t(p) \cap t(r)$. Hence, $w_1 \in (p \wedge r)[\mathcal{M}] \subseteq p[\mathcal{M}]$, so regularity implies $w_1 = s(w, (p \wedge r)[\mathcal{M}])$. Likewise, $w_2 = s(w, (q \wedge r)[\mathcal{M}])$. From $(p \wedge r) \succ_X (q \wedge r)$ we infer $u_X(w_1) > u_X(w_2)$ which in view of (5.4) implies $p \succ_X q$. Thus $w \in \varphi[\mathcal{M}]$. □

Similar reasoning suffices to prove:

PROPOSITION 5.5. *Let $\varphi$ be*

$$(p \wedge ((p \wedge q) \succ_X r)) \rightarrow (q \succ_X r).$$

*Then $\varphi$ is invalid in the class of regular and in the class of proximal models but valid in the class of models that are both regular and proximal.*

For an instance of this formula, suppose that you have a greater gustatory interest in ham and eggs than just oatmeal. Then if you already have ham, you'll be more interested in eggs than just oatmeal.

***5.2. Extensionality, saturation, and perfection.*** We next consider the relation between worlds and the propositional variables they satisfy. The following condition requires that distinct worlds don't make the same variables true.

DEFINITION 5.6. *$\mathcal{M}$ is extensional just in case for all $w_1, w_2 \in \mathbb{W}$, $\{v \in \mathbb{P} \mid w_1 \in t(v)\} = \{v \in \mathbb{P} \mid w_2 \in t(v)\}$ implies $w_1 = w_2$.*

Observe that if $\mathbb{P}$ is finite then every extensional model is finite. Also, it is easy to see that every proximal extensional model is reflexive. Hence, in a finite signature, no such model satisfies the conjunction of Formulas (4.4)(a) and (b). If $\mathbb{P} = \{p\}$ then obviously the invalid formula $(p \approx_X \neg p) \rightarrow (p \approx_X \top)$ is valid in the extensional models.

If every subset of variables inhabits some world, the model may be called "saturated."

DEFINITION 5.7. *$\mathcal{M}$ is saturated just in case for all $T \subseteq \mathbb{P}$ there is $w \in \mathbb{W}$ with $\{v \in \mathbb{P} \mid w \in t(v)\} = T$.*

DEFINITION 5.8. *$\mathcal{M}$ is perfect just in case $\mathcal{M}$ is both extensional and saturated.*

In a perfect model, $\mathbb{W}$ can be identified with the power set of $\mathbb{P}$. The combination of perfection and proximity has consequences for the "contraposition" of reasons, as in $(p \succ_X q) \rightarrow (\neg q \succ_X \neg p)$. This formula is plausible at first sight; for it seems that if $p$ is $u_X$-superior to $q$ then $u_X$ also favors $q$ rather than $p$ *failing* to hold. Thus, keeping a promise is morally superior to teasing the infirm hence not teasing the infirm should be morally superior to not keeping a promise, which it is. Closer inspection, however, reveals that only a weaker form of contraposition can be maintained.

PROPOSITION 5.9. *Let $C$ be the class of perfect and proximal models. Then $(p \succ_X q) \rightarrow (\neg q \succ_X \neg p)$ is not valid in $C$. However, $((\neg p \wedge \neg q) \rightarrow (p \succ_X q))$ is valid in a given model of $C$ iff $((p \wedge q) \rightarrow (\neg q \succ_X \neg p))$ is valid in the same model.*

*Proof.* We demonstrate the left-to-right direction in the second part of the proposition. Let $\mathcal{M} \in C$ be given, and suppose that:

(5.10) $(\neg p \wedge \neg q) \rightarrow (p \succ_X q)$ *is valid in* $\mathcal{M}$.

By saturation, let $w \in t(p) \cap t(q)$. By saturation again, there are $w_1, w_2 \in \mathbb{W}$ with:

(5.11) *(a)* $w_1 = s(w, \neg q[\mathcal{M}])$, *and*
*(b)* $w_2 = s(w, \neg p[\mathcal{M}])$.

To complete the proof it suffices to show that:

(5.12) $u_X(w_1) > u_X(w_2)$.

By (5.11), proximity, and perfection:

(5.13) *(a)* $w_1$ *satisfies the same variables as* $w$, *except for* $q$.
*(b)* $w_2$ *satisfies the same variables as* $w$, *except for* $p$.

By perfection, there is $w^* \in \mathbb{W}$ that satisfies the same subset of $\mathbb{P}$ as $w$ except for $p, q$. That is, $w^*$ falsifies $p$ and $q$ but otherwise agrees with $w$. Hence by (5.10), $w^* \in (p \succ_X q)[\mathcal{M}]$. So there are $w'_1, w'_2 \in \mathbb{W}$ with:

(5.14) *(a)* $w'_1 = s(w^*, p[\mathcal{M}])$,
*(b)* $w'_2 = s(w^*, q[\mathcal{M}])$, *and*
*(c)* $u_X(w'_1) > u_X(w'_2)$.

By (5.14)a,b, proximity, and perfection:

(5.15) *(a)* $w'_1$ *satisfies the same variables as* $w^*$, *except for* $p$.
*(b)* $w'_2$ *satisfies the same variables as* $w^*$, *except for* $q$.

From (5.13), (5.15), and perfection, $w_1 = w'_1$ and $w_2 = w'_2$. Therefore, (5.12) follows from (5.14)c. □

***5.3. Conditions on the utility function.*** We now consider different ways that utilities can be combined. This topic is at the heart of the relation between reasons and preference. For as noted earlier, we conceive preference for $\varphi$ over $\psi$ to be represented by $\varphi \succ_{\bigcup \mathbb{S}} \psi$, that is, taking account of all reasons in play. (Here it is assumed that $\bigcup \mathbb{S} \in \mathbb{S}$.) We start with the most basic condition on utility aggregation, namely, that $u_X$ depends on just the $u_i$ indexed by $X$.

DEFINITION 5.16. *Let finite $X \in \mathbb{S}$ be given. Model $\mathcal{M}$ is local for $X$ just in case:*

*(a) for all $i \in X$, $\{i\} \in \mathbb{S}$ and*
*(b) there is a function $g$ from finite subsets of $\Re$ to $\Re$ such that for all $w \in \mathbb{W}$, $u_X(w) = g(\langle u_i(w) \mid i \in X\rangle)$.*

*In this case, we call $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ $g$-valid if $\varphi$ is true in the class of models for which $u_X$ is computed via $g$.*

For example, locality prevents $u_{\{1,2\}}(w)$ from depending on $u_3(w)$. It is easy to see that the following formula is valid in the class of $\{1, 2\}$-local models but false in some non$\{1, 2\}$-local model.

$$((p \approx_1 p') \wedge (q \approx_1 q') \wedge (p \approx_2 p') \wedge (q \approx_2 q')) \rightarrow ((p \approx_{\{1,2\}} q) \leftrightarrow (p' \approx_{\{1,2\}} q')).$$

Candidates for $g$ in Definition 5.16 include:

$$u_X(w) = \begin{array}{|l|l|} \hline \text{average}\{u_i(w) \mid i \in X\} & \text{median}\{u_i(w) \mid i \in X\} \\ \hline \text{minimum}\{u_i(w) \mid i \in X\} & \text{maximum}\{u_i(w) \mid i \in X\} \\ \hline \end{array}$$

Formulas separate some of these locality classes. For example, the following schema is average-valid but neither min- nor max-valid with respect to $\{i, j\}$.

$$((\varphi \succ_i \psi) \wedge (\varphi \approx_j \psi)) \rightarrow (\varphi \succ_{\{i,j\}} \psi).$$

To see that the schema is not min-valid, take $u_j$ to assign identical numbers to all worlds, much smaller than the numbers that $u_i$ assigns. Do the reverse for a countermodel to max-validity.

Next is a schema that is min-valid and max-valid but not average-valid.

$$(\varphi \approx_{\{i,j,k\}} \psi) \rightarrow ((\varphi \approx_{\{i,j\}} \psi) \vee (\varphi \approx_{\{i,k\}} \psi) \vee (\varphi \approx_{\{j,k\}} \psi)).$$

For a countermodel to the formula with respect to averaging, let $w_1, w_2$ be the worlds attained through $\varphi, \psi$, respectively, and let the $i, j, k$ utilities be given in the accompanying table.

| | $w_1$ | $w_2$ |
|---|---|---|
| $i$ | 2 | 0 |
| $j$ | 2 | 3 |
| $k$ | 2 | 3 |

Observe that utility aggregation has so far been monotonic in the following sense.

DEFINITION 5.17. *Let $X = \{x_1 \dots x_n\} \in \mathbb{S}$ be given, where also $\{x_1\}, \dots, \{x_n\} \in \mathbb{S}$. A model $\mathcal{M}$ is monotone for $X$ just in case for all $\varphi, \psi$,*

$$((\varphi \succ_{x_1} \psi) \wedge \cdots \wedge (\varphi \succ_{x_n} \psi)) \rightarrow (\varphi \succ_X \psi)$$

*is valid in $\mathcal{M}$.*

The four functions discussed above are consistent with monotonicity but it is easy to imagine circumstances in which nonmonotonic aggregation takes place. For example, you might prefer to spend time with people of luxuriant life style (they're more fun), encoded in $u_1$, and also prefer people who espouse asceticism and self-restraint (they're more admirable), encoded in $u_2$. The two utility functions considered individually might order Jim above Jack as dinner partners but $u_{1,2}$ will reverse the preference if it is sensitive to Jim's hypocrisy.

**§6. Decidability and compactness.** The present section offers three theorems about the compactness and decidability of satisfiability (hence, about the decidability of validity as well). For this purpose, we fix a signature $(\mathbb{P}, \mathbb{S})$ in which $\mathbb{P}$ is an initial segment of $\mathbb{N}$, and $\mathbb{S}$ is a set of finite subsets of $\mathbb{N}$. The first theorem concerns satisfiability with respect to the class of all models.

THEOREM 6.1. *The set of satisfiable formulas of* $\mathcal{L}(\mathbb{P}, \mathbb{S})$ *is decidable.*

Adjustments to the proof of Theorem 6.1 verify the following corollaries.

COROLLARY 6.2. *If a formula of* $\mathcal{L}(\mathbb{P}, \mathbb{S})$ *is satisfiable then it is satisfied in a finite model (i.e., in a model with finitely many worlds).*

COROLLARY 6.3. *The set of formulas of* $\mathcal{L}(\mathbb{P}, \mathbb{S})$ *that are satisfiable in the class of reflexive models is decidable.*

Corollary 6.2 may be contrasted with Proposition 4.3, stating that some formulas can be satisfied by a reflexive model only if the model contains infinitely many worlds.

The second theorem bears on lexicographic ordering in the sense of Definition 4.21, and on proposition drivenness in the sense of Definition 4.23.

THEOREM 6.4. *The set of formulas of* $\mathcal{L}(\mathbb{P}, \mathbb{S})$ *that are satisfiable in the class of lexicographic models is decidable, as is the set of formulas that are satisfiable in the class of proposition driven models. Indeed, both sets of formulas are* ***NP****-complete.*

The final theorem affirms that satisfiability with respect to the class of all models is countably compact. We call a collection $\Sigma \subseteq \mathcal{L}$ of formulas "satisfiable" just in case there is a model $\mathcal{M}$ that satisfies every member of $\Sigma$ at a common point, that is, just in case:

$$\bigcap\{\varphi[\mathcal{M}] \mid \varphi \in \Sigma\} \neq \emptyset.$$

THEOREM 6.5. *Suppose that signature* $(\mathbb{P}, \mathbb{S})$ *is countable, and let* $\Sigma \subseteq \mathcal{L}(\mathbb{P}, \mathbb{S})$ *be given. Then* $\Sigma$ *is satisfiable if and only if every finite subset of* $\Sigma$ *is satisfiable.*

On the other hand, if either $\mathbb{P}$ or $\mathbb{S}$ is uncountable then compactness breaks down. Proofs of the theorems are provided in the Appendix.

**§7. Generalized frames.** In our theory, $\varphi \succeq_X \psi$ can be understood as asserting that $u_X$ assigns at least as much value to the proposition expressed by $\varphi$ as to the proposition expressed by $\psi$. The latter two propositions are represented by elements of each, picked out as a function of the world at which the formula is evaluated. A natural generalization is to compare the value of propositions directly, without recourse to individual worlds as representatives. We explore this idea in the present section. Let $(\mathbb{P}, \mathbb{S})$ be our background signature, and recall that a *total preorder* is transitive, connected, and reflexive over its domain.

DEFINITION 7.1. *Let a set* $\mathbb{W}$ *of worlds be given.*

*(a) By a value-ordering for* $\mathbb{W}$ *and* $\mathbb{S}$ *is meant a function* $v$ *from* $\mathbb{W} \times \mathbb{S}$ *to the set of total preorders over the class of nonempty subsets of* $\mathbb{W}$*. We call the pair* $(\mathbb{W}, v)$ *a generalized frame.*

*(b) Let a truth-assignment* $t$ *and a value-ordering* $v$ *for* $\mathbb{W}$ *and* $\mathbb{S}$ *be given. Then* $(\mathbb{W}, t, v)$ *is a generalized model.*

Intuitively, a value-ordering arranges propositions by utility, relative to index $X \in \mathbb{S}$ and vantage point $w \in \mathbb{W}$. The semantics of generalized models is given by Definition 3.5 with the following substitution for Clause 3.5d. Let $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ and generalized model $\mathcal{M} = (\mathbb{W}, t, v)$ for $(\mathbb{P}, \mathbb{S})$ be given.

3.5d′ If $\varphi$ has the form $(\theta \succeq_X \psi)$ for $X \in \mathbb{S}$, then $\varphi[\mathcal{M}] = \emptyset$ if either $\theta[\mathcal{M}] = \emptyset$ or $\psi[\mathcal{M}] = \emptyset$. Otherwise:

$$\varphi[\mathcal{M}] = \{w \in \mathbb{W} \mid \theta[\mathcal{M}] \text{ comes no earlier than } \psi[\mathcal{M}] \text{ in } v(w, X)\}.$$

Now let model $\mathcal{M} = (\mathbb{W}, s, u, t)$ be given. Then a value-ordering $v$ is induced by the following condition. For $w \in \mathbb{W}$, $X \in \mathbb{S}$, and nonempty $A, B \subseteq \mathbb{W}$, $A$ is (weakly) ordered after $B$ iff $u_X(w_A) \geq u_X(w_B)$ where $w_A = s(w, A)$ and $w_B = s(w, B)$. (The truth-assignment $t$ plays no role.) We call $(\mathbb{W}, v)$ *the generalized frame induced by* $\mathcal{M}$.

Given a model $(\mathbb{W}, s, u, t)$, $w \in \mathbb{W}$, and nonempty $A \subseteq \mathbb{W}$, there is $w_0 \in \mathbb{W}$ with $u_X(s(w, A)) = u_X(s(w, \{w_0\}))$, namely, $w_0 = s(w, A)$. So we have:

LEMMA 7.2. *Let value-ordering* $v$ *be induced by model* $(\mathbb{W}, s, u, t)$*. Then for all* $w \in \mathbb{W}$ *and* $X \in \mathbb{S}$*, every equivalence class in* $v(w, X)$ *contains a singleton set.*

We have the following immediate consequence, which shows that some generalized frames cannot be induced by models.

PROPOSITION 7.3. *Let* $\mathbb{W}$ *contain at least two worlds. Let value-ordering* $v$ *be such that for some* $w \in \mathbb{W}$ *and* $X \in \mathbb{S}$*, either*

*(a)* $v(w, X)$ *refines* $\subset$ *over the field of nonempty subsets of* $\mathbb{W}$ *or*

*(b)* $v(w, X)$ *is a strict linear order over the nonempty subsets of* $\mathbb{W}$*.*

*Then* $(\mathbb{W}, v)$ *is not induced by any model.*

On the other hand, the next proposition shows that some interesting classes of generalized frames can be characterized in $\mathcal{L}(\mathbb{P}, \mathbb{S})$. We say that $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ is *valid in a generalized frame* $(\mathbb{W}, v)$ in case $\varphi$ is true in every generalized model of form $(\mathbb{W}, t, v)$.

PROPOSITION 7.4. *Let* $X \in \mathbb{S}$ *be given. There are* $\varphi_1, \varphi_2 \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ *such that*

*(a)* $\varphi_1$ *is valid in a generalized frame* $(\mathbb{W}, v)$ *if and only if for all* $w \in \mathbb{W}$*,* $v(w, X)$ *refines* $\subset$ *over the nonempty subsets of* $\mathbb{W}$*;*

*(b)* $\varphi_2$ *is valid in a generalized frame* $(\mathbb{W}, v)$ *if and only if for all* $w \in \mathbb{W}$*,* $v(w, X)$ *is a strict linear order over the nonempty subsets of* $\mathbb{W}$*.*

*Proof.* It is easy to verify the proposition with the following choices of $\varphi_1, \varphi_2$, respectively.

$$(\Box(p \to q) \wedge \neg\Box(q \to p) \wedge \Diamond p) \to p \prec_X q$$

$$\neg\Box(p \leftrightarrow q) \to ((p \prec_X q) \vee (q \prec_X p))$$

□

In Osherson & Weinstein (To appear), we provide an axiomatization of the set of formulas that are valid in generalized frames and establish that a formula is valid in the class of generalized frames if and only if it is valid in the class of frames.

**§8. Discussion.** The foregoing investigation raises many questions and avenues for further research. We indicate some directions.

***8.1. Utility.*** Suppose that distinct $\{i\}, \{j\}, \{k\} \in \mathbb{S}$. For all $\varphi, \theta \in \mathcal{L}(\mathbb{P}, \mathbb{S})$, let:

$$(\varphi \veebar \theta) \stackrel{\text{def}}{=} ((\varphi \succ_i \theta) \wedge (\varphi \succ_j \theta)) \vee ((\varphi \succ_i \theta) \wedge (\varphi \succ_k \theta)) \vee ((\varphi \succ_j \theta) \wedge (\varphi \succ_k \theta)).$$

Then $\varphi \veebar \theta$ is true if a majority of the utility scales $i, j, k$ are favorable to $\varphi$ compared to $\theta$. Observe that $((\varphi \veebar \theta) \wedge (\theta \veebar \psi)) \to (\varphi \veebar \psi)$ (transitivity) is not guaranteed in a given model inasmuch as the utility scales $u_i, u_j, u_k$ might embody a voting cycle (see Johnson, 1998). Therefore, $\veebar$ cannot itself be represented by a utility scale. The following matter thus merits exploration.

OPEN QUESTION 8.1. *Suppose that $\{i, j, k\} \in \mathbb{S}$. Under what conditions does $\varphi \veebar \theta$ imply $\varphi \succ_{i,j,k} \theta$, and vice versa?*

The voting operator $\veebar$ might best be analyzed in the context of a generalization of our approach to utility. Instead of utility scales corresponding to each $X \in \mathbb{S}$, we may posit relations $R_X \subseteq \mathbb{W} \times \mathbb{W}$. In this setup, $\theta \succeq_X \psi$ is true at $w \in \mathbb{W}$ just in case $(s(w, \theta[\mathcal{M}]), s(w, \psi[\mathcal{M}])) \in R_X$. Such relations $R_X$ could vary in their order-theoretic properties (e.g., transitivity) as well as in their connection to relations $R_i$ with $i \in X$. This perspective might allow the remarkable results developed in Andréka *et al.* (2002), about combining preference relations, to shed light on the logic of reasons. In Osherson & Weinstein (Forthcoming), we axiomatize the sets of valid formulas that arise under various choices of relation $R$.

Questions also remain about the classes of utility functions defined in Section 5.3. Can any of them be uniquely characterized by a set of formulas? Even the less ambitious problem of separating utility functions is currently unresolved. For example, the following question was left open.

OPEN QUESTION 8.2. *Is there $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ that is minimum-valid but not maximum-valid (and vice versa)?*

***8.2. Selection.*** Additional conditions on selection functions remain to be investigated. For example, an alternative concept of selection allows more than one world to be "nearest" to a target. To express this idea, we replace Definition 3.1 with the following.

DEFINITION 8.3. *A wide selection function $s$ over $\mathbb{W}$ is a mapping from $\mathbb{W} \times \{A \subseteq \mathbb{W} \mid A \neq \emptyset\}$ to the power set of $\mathbb{W}$ such that for all $w \in \mathbb{W}$ and $\emptyset \neq A \subseteq \mathbb{W}$, $\emptyset \neq s(w, A) \subseteq A$.*

Selection functions in the original sense of Definition 3.1 can now be seen as the special case in which only singleton sets are returned. To satisfy a formula $\theta \succeq_X \psi$ in the context of a wide selection function, we may require that some nearby $\theta$-world is weakly $X$-better than some nearby $\psi$-world, or that all of them are, etc. The consequences of these options have yet to be explored.

***8.3. Updating.*** Suppose you live in a model $\mathcal{M} = (\mathbb{W}, s, u, t)$ but wish to take on board $\varphi \in \mathcal{L}$ as an assumption. We take this to mean that $\varphi$ will be made true in all worlds of some successor model $\mathcal{M}' = (\mathbb{W}', s', u', t')$ that is the natural *$\varphi$-update* to $\mathcal{M}$. (Updating is analyzed from a graph-theoretic perspective in Andréka *et al.*, 2002; van Benthem & Liu, 2007.)

If $\varphi$ is boolean, updating $\mathcal{M}$ seems easy: set $\mathbb{W}' = \{w \in \mathbb{W} \mid w \models \varphi\}$, and let $s'$, $u'$, $t'$ be the obvious reducts of $s$, $u$, and $t$ to $\mathbb{W}'$. (Updating in this sense is not defined if $\mathbb{W}' = \emptyset$.) But if $\varphi$ has a modal connective, matters are not straightforward. Consider the following choice for $\mathcal{M}$, where $(\mathbb{P}, \mathbb{S}) = (\{p, q\}, \{\{i\}\})$.

$$\mathbb{W} = \{w_1, w_2, w_3, w_4\}$$

$$t(p) = \{w_2, w_4\} \quad t(q) = \{w_3\}$$

$$u_i(w_4) < u_i(w_3) < u_i(w_2) < u_i(w_1)$$

$$s(w_1, p[\mathcal{M}]) = w_2$$

$$s(w_1, q[\mathcal{M}]) = w_3$$

$$s(w_2, p[\mathcal{M}]) = s(w_3, p[\mathcal{M}]) = s(w_4, p[\mathcal{M}]) = w_4$$

$$s(w_2, q[\mathcal{M}]) = s(w_3, q[\mathcal{M}]) = s(w_4, q[\mathcal{M}]) = w_3$$

For $\varphi := p \prec_i q$ to be true throughout $\mathcal{M}'$, it suffices to remove $w_1$ from $\mathbb{W}$. But since $s(w_1, \{w_4\})$ must equal $w_4$, it is easy to verify that removing $w_2$ from $\mathbb{W}$ also suffices for the same purpose. Updating in the general case thus requires choice among successor models, in a sense familiar from the theory of belief revision (Gärdenfors, 1988). Investigation of the matter might usefully address the following issue. Given a proposed updating operator $\ddagger$ and a class $C$ of models with (say) the regularity property, for which $\varphi \in \mathcal{L}$ (if any) is $\{\mathcal{M} \ddagger \varphi \mid \mathcal{M} \in C\}$ guaranteed to be regular?

## §9. Appendix: Proof of Theorem 4.20.

The theorem is an immediate corollary to the following lemma, the proof of which involves the notion of *modal depth*, defined as follows.

DEFINITION 9.1. *We define $\mu(\varphi)$, the modal depth of $\varphi$, by recursion on $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ as follows.*

$$\mu(\varphi) = \begin{cases} 0 & \text{if } \varphi \in \mathbb{P} \\ \mu(\psi) & \text{if } \varphi = \neg\psi \\ \max\{\mu(\psi), \mu(\theta)\} & \text{if } \varphi = (\psi \wedge \theta) \\ \max\{\mu(\psi), \mu(\theta)\} + 1 & \text{if } \varphi = (\psi \preceq_X \theta) \end{cases}$$

LEMMA 9.2. *Suppose $\mathbb{P}$ and $\mathbb{S}$ are finite. For every ref and regular model $\mathcal{M} = (\mathbb{W}, s, u, t)$, $w \in \mathbb{W}$, and $n \in \mathbb{N}$, there is a metric model $\mathcal{M}^* = (\mathbb{W}^*, s^*, u^*, t^*)$ and a world $w^* \in \mathbb{W}^*$ such that for every $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$, if $\mu(\varphi) \leq n$, then*

$$w \in \varphi[\mathcal{M}] \text{ if and only if } w^* \in \varphi[\mathcal{M}^*].$$

*Proof.* Let ref and regular model $\mathcal{M} = (\mathbb{W}, s, u, t)$, $w \in \mathbb{W}$, and $n \in \mathbb{N}$ be given. To prove the proposition, we construct from $\mathcal{M}$ a metric model $\mathcal{M}^* = (\mathbb{W}^*, s^*, u^*, t^*)$ and a world $w^* \in \mathbb{W}^*$ such that for every $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$, if $\mu(\varphi) \leq n$, then

$$w \in \varphi[\mathcal{M}] \text{ if and only if } w^* \in \varphi[\mathcal{M}^*].$$

To reduce notational clutter, we suppress the subscripts on occurrences of $\preceq$ and suppress them likewise on utility functions $u$. The construction proceeds in three steps. In the first step, we *unravel* $\mathcal{M}$ at the world $w$ with respect to the modalities generated by the selection function $s$ and the propositions expressed by formulas of modal depth $\leq n$. (Unraveling is a standard technique in modal logic, see Blackburn *et al.* 2001, sec. 2.1.) This leads to a "partial" model $\mathcal{M}^\dagger = (\mathbb{W}^*, s^\dagger, u^*, t^*)$ and a world $w^* \in \mathbb{W}^*$ such that for every $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ with $\mu(\varphi) \leq n$,

(9.3)

$$w \in \varphi[\mathcal{M}] \textit{ if and only if } w^* \in \varphi[\mathcal{M}^\dagger],$$

in which $s^\dagger$ is only defined in its second argument on propositions expressed by formulas of modal depth $< n$. In the second step, we define a metric $d$ on $\mathbb{W}^*$ with respect to which the partial selector $s^\dagger$ is metric. Finally, we extend $s^\dagger$ to a total selector $s^*$ and verify that $d$ is a metricization of the resulting model $\mathcal{M}^*$. □

STEP 1: For every $k \in \mathbb{N}$, let $\Omega_k = \{\varphi[\mathcal{M}] \mid \mu(\varphi) \leq k\} - \{\emptyset\}$. It follows immediately from the finiteness of $(\mathbb{P}, \mathbb{S})$ that for every $k \in \mathbb{N}$, $\Omega_k$ is finite. For each $Z \in \Omega_{n-1}$, define a binary relation $R_Z$ on $\mathbb{W}$ as follows.

$$\text{For all } v, v' \in \mathbb{W}\text{ , } R_Z(v, v') \text{ if and only if } s(v, Z) = v'.$$

We use the "accessibility" relations $R_Z$ to define the unraveling of $\mathcal{M}$ at $w$. A finite sequence $\langle w_0, w_1, \ldots, w_m\rangle$ from $\mathbb{W}$ is a world in $\mathbb{W}^*$ if and only if $w_0 = w$ and for each $0 \leq i < m$, there is a $Z \in \Omega_{n-1}$ such that $R_Z(w_i, w_{i+1})$. If $v^* = \langle w_0, w_1, \ldots, w_m\rangle \in \mathbb{W}^*$, we write $\mathsf{last}(v^*)$ for $w_m$. For $w' \in \mathbb{W}$, we use $v^*w'$ to denote $\langle w_0, w_1, \ldots, w_m, w'\rangle$, the concatenation of $w'$ to the end of $v^*$. Define $u^*$ and $t^*$ by emulating $u$ and $t$, namely:

(9.4) *For all* $v^* \in \mathbb{W}^*$,

*(a)* $u^*(v^*) = u(\mathsf{last}(v^*))$ *and*

*(b)* $v^* \in t^*(p)$ *if and only if* $\mathsf{last}(v^*) \in t(p)$, *for every* $p \in \mathbb{P}$.

In order to complete the construction of the partial model $\mathcal{M}^\dagger$ it remains only to define the partial selector $s^\dagger(v^*, \psi[\mathcal{M}^\dagger])$ for all $v^* \in \mathbb{W}^*$ and all $\psi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ of modal depth $< n$. We do this by a recursion on $j \leq n$ which simultaneously defines a sequence of partial selectors $s^\dagger_j$ and partial models $\mathcal{M}^\dagger_j = (\mathbb{W}^*, s^\dagger_j, u^*, t^*)$. We begin the recursion by setting $s^\dagger_0 = \emptyset$. Then for $0 \leq j < n$ and for $\psi$ of modal depth $\leq j$, we let

$$\text{(9.5)}\quad s^\dagger_{j+1}(v^*, \psi[\mathcal{M}^\dagger_j]) = \begin{cases} v^*s(\mathsf{last}(v^*), \psi[\mathcal{M}]) & \textit{if } s(\mathsf{last}(v^*), \psi[\mathcal{M}]) \neq \mathsf{last}(v^*) \\ v* & \textit{otherwise.} \end{cases}$$

In order to justify this inductive definition we must show that for each $0 \leq j \leq n$, $\psi[\mathcal{M}^\dagger_j]$ is well-defined by Definition 3.5 for each $\psi$ of modal depth $\leq j$. For this, it suffices to show that for each $0 \leq j < n$, $s^\dagger_{j+1}$ is a partial selector, that is, we must show that

(9.6) *for every* $v^* \in \mathbb{W}^*$ *and* $\psi, \chi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ *of modal depth* $\leq j$,

(a) $\psi[\mathcal{M}_j^\dagger] = \chi[\mathcal{M}_j^\dagger]$ *if and only if* $\psi[\mathcal{M}] = \chi[\mathcal{M}]$ *and*

(b) *if* $\psi[\mathcal{M}_j^\dagger] \neq \emptyset$, *then* $s_{j+1}^\dagger(v^*, \psi[\mathcal{M}_j^\dagger]) \in \psi[\mathcal{M}_j^\dagger]$.

Condition (9.6)a guarantees that the definition of $s_{j+1}^\dagger$ via (9.5) is consistent, while (9.6)b insures that $s_{j+1}^\dagger$ is a selector for propositions expressed by sentences of modal depth $\leq j$. We will establish (9.6) by induction and simultaneously prove by induction that for all $0 \leq j \leq n$,

(9.7) *for all* $v^* \in \mathbb{W}^*$ *and* $\psi, \chi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ *of modal depth* $\leq j$,

(a) $v^* \in \psi[\mathcal{M}_j^\dagger]$ *if and only if* $\mathsf{last}(v^*) \in \psi[\mathcal{M}]$, *if and only if* $\psi[\mathcal{M}] = \chi[\mathcal{M}]$ *and*

(b) $\psi[\mathcal{M}_j^\dagger] = \emptyset$ *if and only if* $\psi[\mathcal{M}] = \emptyset$.

To begin the induction, let $j = 0$. Then for all $\psi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ of modal depth $\leq j$, $\psi[\mathcal{M}_j^\dagger]$ is well-defined, since, by Definition 3.5, its value does not depend on $s_j^\dagger$. Moreover, in this case, (9.7)a is a direct consequence of (9.4)b. Next, note that the right-to-left direction of (9.7)b follows immediately from (9.7)a. For the left-to-right direction, suppose that $\psi[\mathcal{M}] \neq \emptyset$. It then follows from the definition of $\mathbb{W}^*$ that there is a $v^* \in \mathbb{W}^*$ such that $\mathsf{last}(v^*) \in \psi[\mathcal{M}]$. Therefore, by (9.7)a, $\psi[\mathcal{M}_j^\dagger] \neq \emptyset$. Observe next that the right-to-left direction of (9.6)a also follows immediately from (9.7)a. For the left-to-right direction, suppose that $\psi[\mathcal{M}] \neq \chi[\mathcal{M}]$, say, $(\psi \wedge \neg\chi)[\mathcal{M}] \neq \emptyset$; the other case $(\chi \wedge \neg\psi)[\mathcal{M}] \neq \emptyset$ is handled the same way. Since $\psi$ and $\chi$ are sentences of modal depth $\leq j$, so is $(\psi \wedge \neg\chi)$. It follows from the definition of $\mathbb{W}^*$ that there is a $v^* \in \mathbb{W}^*$ such that $\mathsf{last}(v^*) \in \psi[\mathcal{M}] - \chi[\mathcal{M}]$. Therefore, by (9.7)a, $\psi[\mathcal{M}_j^\dagger] \neq \chi[\mathcal{M}_j^\dagger]$. Finally, (9.6)b follows immediately from (9.5) and (9.7)a.

For the induction step, suppose that (9.6)a,b and (9.7)a,b hold for all $v^* \in \mathbb{W}^*$ and $\psi, \chi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ of modal depth $\leq j$ for some $0 < j < n$, and let $\theta \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ be a sentence of modal depth $j + 1$. It is easy to see from the definition of the models under construction that $\psi[\mathcal{M}_j^\dagger] = \psi[\mathcal{M}_{j+1}^\dagger]$ for $\psi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ of modal depth $\leq j$. It follows at once, from the induction hypothesis, that $\theta[\mathcal{M}_{j+1}^\dagger]$ is well-defined by Definition 3.5. In order to conclude the argument, it suffices to show that (9.7)a holds for $\theta$, since the proofs of (9.6)a,b and (9.7)b from (9.7)a are exactly parallel to those in the basis of the induction. Moreover, since $\theta$ has modal depth $j + 1$, $\theta$ is a boolean combination of sentences of the form $\psi \preceq \chi$ where $\psi$ and $\chi$ have modal depth $\leq j$. Since it is easy to see that (9.7)a is preserved under boolean combinations, we may suppose that $\theta$ is $\psi \preceq \chi$ for some $\psi$ and $\chi$ of modal depth $\leq j$. In case either $\psi[\mathcal{M}_{j+1}^\dagger] = \emptyset$ or $\chi[\mathcal{M}_{j+1}^\dagger] = \emptyset$, it follows immediately from the induction hypothesis (9.7)b that (9.7)a holds for $\theta$. So suppose that $\psi[\mathcal{M}_{j+1}^\dagger] \neq \emptyset$ and $\chi[\mathcal{M}_{j+1}^\dagger] \neq \emptyset$. It then follows immediately from (9.4)a, (9.5), and the induction hypothesis (9.7)a for $\psi$ and $\chi$ that (9.7)a holds for $\theta$.

Finally, let $\mathcal{M}^\dagger = \mathcal{M}_n^\dagger$ and $w^* = \langle w \rangle$. It follows at once that (9.3) holds for all $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ of modal depth $\leq n$.

STEP 2: We now define a metric $d$ on $\mathbb{W}^*$ which is a metricization of the partial structure $\mathcal{M}^\dagger$, that is,

(9.8) *for each* $\psi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ *of modal depth* $< n$, *and each* $v^* \in \mathbb{W}^*$, $s^\dagger(v^*, \psi[\mathcal{M}^\dagger])$ *is the unique* $d$*-nearest world to* $v^*$.

Let $\Omega^\dagger = \{\varphi[\mathcal{M}^\dagger] \mid \mu(\varphi) \leq (n-1)\} - \{\emptyset\}$. For $v \in \mathbb{W}^*$, let $C_v = \{s^\dagger(v, Z) \mid Z \in \Omega^\dagger\}$, and let $k_v$ be the cardinality of $C_v$. Note that $\mathbb{W}^*$ is a finitely branching directed tree $T$ where $v'$ is a child of $v$ if and only if $v' = v\mathsf{last}(v')$. For all $v, v' \in \mathbb{W}^*$, $v' \in C_v$ if and only if $v'$ is a child of $v$ or $v' = v$, hence the out-degree of each world is $k_v - 1$. We first show that for every $v \in \mathbb{W}^*$ there is a strict linear ordering $x_1^v, \ldots, x_{k_v}^v$ of $C_v$ such that

(9.9) *for every $Z \in \Omega^\dagger$, $s^\dagger(v, Z)$ is the $x_i^v \in Z$ of lowest index.*

For the proof of (9.9), observe that the regularity of $\mathcal{M}$ implies that for every $v' \in C_v$, there is a $Z \in \Omega^\dagger$ such that $Z \cap C_v = \{v'\}$. Since $\Omega^\dagger$ is closed under boolean combinations, it is immediate that for every nonempty $Y \subseteq C_v$, there is a $Z_Y \in \Omega^\dagger$ such that $Z_Y \cap C_v = Y$. Now, for each $v \in \mathbb{W}^*$, we define the $x_i^v$'s by induction on $i \in \mathbb{N}$ up to $k_v$. Let $x_1^v = v$ and for $1 \leq i < k_v$, let $x_{i+1}^v = s^\dagger(v, Z_{C_v - \{x_1^v, \ldots, x_i^v\}})$. It follows directly from the definition and the reflexivity and regularity of $\mathcal{M}$ that (9.9) holds.

As general background, for $T'$ an edge-weighted directed tree and $v, v'$ vertexes of $T'$, we write $d(v, v')$ for the the ordinary (i.e., weighted path-length) distance between $v$ and $v'$ in the edge-weighted symmetrized tree derived from $T'$. Returning to our tree $T$, we write $T_m$ for the restriction of $T$ to its vertexes of height $\leq m$ and $V_m$ for its set of vertexes of height exactly $m$. We proceed to introduce weights on the edges of $T$ via recursion on $T_m$. The basis is trivial, since $T_0$ contains no edges. At stage $m > 0$, for all $v \in V_{m-1}$ and $1 < i \leq k_v$, we choose weights $\lambda_i^v$ on the edges from $v$ to $x_i^v$ to satisfy the following conditions.

(9.10) *(a) if $i < k_v$, then $\lambda_i^v < \lambda_{i+1}^v$;*
*(b) $\lambda_i^v \in (1 + 2^{-(m+1)}, 1 + 2^{-m})$;*
*(c) for all $v, v', v'' \in T_m$, if $v' \neq v''$, then $d(v, v') \neq d(v, v'')$.*

It is clear that such weights can be chosen. Now (9.8) follows directly from (9.9) and (9.10)a,b.
STEP 3: By (9.10)c, we may define a total selector $s^*$ extending $s^\dagger$ such that for every $v \in \mathbb{W}^*$ and every nonempty $X \subseteq \mathbb{W}^*$, $s^*(v, X)$ is the unique $d$-nearest member of $X$ to $v$. This completes our definition of the metric structure $\mathcal{M}^*$. It follows at once from (9.3) that for all $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ of modal depth $\leq n$, $w \in \varphi[\mathcal{M}]$ if and only if $w^* \in \varphi[\mathcal{M}^*]$. □

**§10. Appendix: proof of Theorem 6.1.** To demonstrate that the set of satisfiable formulas of $\mathcal{L}(\mathbb{P}, \mathbb{S})$ is decidable, we apply the well-known "method of mosaics" (see Blackburn *et al*. 2001, sec. 6.4). We carry out the construction in some detail.

Let $\theta \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ be given. Let $\Sigma$ be the collection of subformulas of $\theta$, and let $Z$ be the set of utility indices that appear in $\theta$. We close $\Sigma$ under one application of negation, followed by one application of $\leftrightarrow$, followed by one application of negation, followed by one application of $\Diamond$, followed by one application of negation. The resulting set of formulas will be called $\Omega$. We say that $\Delta \subseteq \Omega$ is a *Hintikka set* (abbreviated *H-set*) if and only if

(10.1) *(a) for every $\neg\varphi \in \Omega$, $\varphi \in \Delta$ iff $\neg\varphi \notin \Delta$ and*
*(b) for every $(\varphi \wedge \psi) \in \Omega$, $(\varphi \wedge \psi) \in \Delta$ iff both $\varphi \in \Delta$ and $\psi \in \Delta$.*

We let $\Xi$ be the collection of all H-sets. Note that if $n$ is the length of $\theta$, then the size $c$ of $\Omega$ (and thus of every H-set) is $O(n^2)$. Therefore, the size $d$ of $\Xi$ is $O(2^{n^2})$. For the purposes of the next definition, we establish the notational convention that if $f$ is the graph of a partial function, we write $f(a)$ for the $b$ such that $\langle a, b\rangle \in f$, when $a$ is in the domain of $f$. A *brick* is a triple $\langle \Delta, \sigma, \{\upsilon_X \mid X \in Z\}\rangle$ where

(10.2) *(a)* $\Delta$ *is an H-set;*

*(b)* $\sigma$ *is the graph of a partial function from* $\Sigma$ *into* $\Xi$ *such that* $\varphi \in \sigma(\varphi)$ *for every* $\varphi \in \Sigma$ *on which* $\sigma$ *is defined;*

*(c) for each* $X \in Z$, $\upsilon_X$ *is a function from* $\sigma$ *to* $\{i \mid 1 \leq i \leq \mathsf{card}(\sigma)\}$*;*

*(d) if* $\Diamond\varphi \in \Delta$*, then for some* $\Delta' \in \mathsf{range}(\sigma)$, $\varphi \in \Delta'$*;*

*(e) if* $\Diamond\varphi \notin \Delta$*, then for all* $\Delta' \in \mathsf{range}(\sigma) \cup \{\Delta\}$, $\varphi \notin \Delta'$*;*

*(f)* $\Box(\varphi \leftrightarrow \psi) \in \Delta$ *if and only if* $\sigma(\varphi) = \sigma(\psi)$*;*

*(g)* $(\varphi \preceq_X \psi) \in \Delta$ *iff* $\upsilon_X(\langle\varphi, \sigma(\varphi)\rangle) \leq \upsilon_X(\langle\psi, \sigma(\psi)\rangle)$.

Let $z$ be the size of $Z$. Note that the number $b$ of bricks is $O(d^{c+1} \cdot c^{cz})$.

If $\beta$ is a brick, we write $\beta_1$, $\beta_2$, and $\beta_3$ for the first, second, and third coordinates of $\beta$. A set $B$ of bricks is a *mosaic* if and only if

(10.3) *(a) for all* $\beta, \beta' \in B$, $\{\varphi \mid \Diamond\varphi \in \beta_1\} = \{\varphi \mid \Diamond\varphi \in \beta_1'\}$ *and*

*(b) for all* $\beta \in B$ *and for all* $\Delta \in \mathsf{range}(\beta_2)$ *there is a* $\beta' \in B$ *such that* $\beta_1' = \Delta$.

A set $B$ of bricks is a mosaic for $\theta \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ if and only if $B$ is a mosaic and for some $\beta \in B$, $\theta \in \beta_1$. Note that the number of mosaics is $O(2^b)$ and that it is decidable in time polynomial in the size of a set $B$ of bricks whether $B$ is a mosaic. It follows that the decision problem "Does there exist a mosaic for $\theta$" is in $\mathsf{NTIME}(b)$. Theorem 6.1 is thus a corollary to the following.

PROPOSITION 10.4. *For every* $\theta \in \mathcal{L}(\mathbb{P}, \mathbb{S})$, $\theta$ *is satisfiable if and only if there is a mosaic for* $\theta$.

To prove the left to right direction of Proposition 10.4, let satisfiable $\theta \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ be given. For notational convenience we assume that $Z = \{X\}$. The generalization to multiple utility indices is routine. Let model $\mathcal{M} = (\mathbb{W}, s, u, t)$ satisfy $\theta$ and suppose that $w_0 \in \theta[\mathcal{M}]$. For each $w \in \mathbb{W}$, let $\Delta_w = \{\varphi \in \mathbf{\Omega} \mid w \in \varphi[\mathcal{M}]\}$. Note that for every $w \in \mathbb{W}$, $\Delta_w$ is an H-set. Now for each $w \in \mathbb{W}$, we construct a brick $\beta_w = \langle\Delta_w, \sigma_w, \upsilon_w\rangle$, where for each $\varphi \in \Sigma$,

$$\sigma_w(\varphi) = \Delta_{s(w, \varphi[\mathcal{M}])}$$

and for all $\varphi, \psi \in \Delta_w$

$$\upsilon_w(\langle\varphi, \sigma_w(\varphi)\rangle) \leq \upsilon_w(\langle\psi, \sigma_w(\psi)\rangle) \text{ iff } u(s(w, \varphi[\mathcal{M}])) \leq u(s(w, \psi[\mathcal{M}])).$$

Let $B = \{\beta_w \mid w \in \mathbb{W}\}$. It is easy to verify that $B$ is a mosaic for $\theta$.

For the right to left direction of Proposition 10.4, suppose that $B$ is a mosaic for $\theta$. We show how to use the bricks of $B$ to construct a tree-like infinite model $\mathcal{M} = (\mathbb{W}, s, u, t)$, the root world $w_0$ of which satisfies $\theta$. The tree underlying the model is both node-labeled and edge-labeled. We call the node labels "brick-labels" and we call the edge labels "selector-labels." The construction of the tree proceeds by induction; at stage $i$, we construct the worlds of depth $i$.

At stage 0 we introduce a world $w_0$ and label $w_0$ with some brick $\beta \in B$ such that $\theta \in \beta_1$ (such a brick-label exists, since $B$ is a mosaic for $\theta$).

Let $W_n$ be the set of worlds constructed at stage $n$. At stage $n+1$ we proceed as follows. For each $w \in W_n$ we construct the children of $w$ as follows. Let $\beta_w = \langle\Delta_w, \sigma_w, \upsilon_w\rangle$ be the brick-label of $w$. For each $\Delta \in \mathsf{range}(\sigma_w)$ we introduce a child $w'$ of $w$; we label the edge from $w$ to $w'$ with selector-label $\sigma_w^{-1}[\Delta](= \{\varphi \mid \sigma_w(\varphi) = \Delta\})$ and we choose as the the brick-label of $w'$ some brick $\beta \in B$ with $\beta_1 = \Delta$ (by (10.3)ii, such a brick exists since $B$ is a mosaic). This completes the definition of $W_{n+1}$. We let $\mathbb{W} = \bigcup_n W_n$. Next we define $u$ and $t$.

(a) For every $p \in \mathbb{P}$, $t(p) = \{w \mid p \in (\beta_w)_1\}$.
(b) For every $w' \in \mathbb{W}$ $u(w') = v_w(\Delta)$, where $w$ and $\Delta$ are the unique world and Hintikka set such that $w'$ is a child of $w$ and $(\beta_{w'})_1 = \Delta$.

To complete the definition of the model $\mathcal{M}$ we must specify the selection function $s$. Recall that $\Sigma$ is the set of subformulas of $\theta$. For nonempty subsets $T$ of $\mathbb{W}$ not of the form $\varphi[\mathcal{M}]$ for some $\varphi \in \Sigma$, and for each $w$, we choose $s(w, T)$ to be an arbitrary element of $T$. For $\varphi \in \Sigma$ we wish to define $s$ so that for every $w \in \mathbb{W}$, $s(w, \varphi[\mathcal{M}])$ is the unique child $w'$ of $w$ such that $\varphi$ is an element of the selector-label of the edge from $w$ to $w'$, provided that $\varphi[\mathcal{M}]$ is nonempty. Since, in general, $\varphi[\mathcal{M}]$ depends on $s$, we will define $s$ on sets of the form $\varphi[\mathcal{M}]$ by recursion on the logical complexity of $\varphi$. Simultaneously, we will prove by induction on logical complexity that for every $w \in \mathbb{W}$ and $\varphi \in \Sigma$,

$$w \in \varphi[\mathcal{M}] \quad \text{iff} \quad \varphi \in (\beta_w)_1,$$

thereby completing the proof of the theorem.

Basis: It follows immediately from the definition of $t$ that for every $p \in \Sigma$ and $w \in \mathbb{W}$:

(10.5) $w \in p[\mathcal{M}]$ *iff* $p \in (\beta_w)_1$.

Now, for each $p \in \Sigma$ and $w \in \mathbb{W}$, we let $s(w, p[\mathcal{M}])$ be the unique child $w'$ of $w$ such that $p$ is an element of the selector-label of the edge from $w$ to $w'$, provided that $p[\mathcal{M}]$ is nonempty. In order to secure the legitimacy of this definition we need to verify that

(10.6) *(a)* $s(w, p[\mathcal{M}]) \in p[\mathcal{M}]$*;*
*(b)* $p[\mathcal{M}] = \emptyset$ *iff no selector-label of an edge exiting* $w$ *contains* $p$*;*
*(c) for all* $q \in \Sigma$*,* $p[\mathcal{M}] = q[\mathcal{M}]$ *iff* $p$ *and* $q$ *are contained in exactly the same selector-labels of edges exiting* $w$ *(i.e., either none of them, or a unique one containing both).*

Condition (10.6)a follows immediately from (10.5). In order to establish (10.6)b, suppose first that $p[\mathcal{M}] = \emptyset$. By (10.5), it follows that for all $w' \in \mathbb{W}$, $p \notin (\beta_{w'})_1$. Hence, for no child $w'$ of $w$ is $p \in (\beta_{w'})_1$. Therefore, no selector-label of an edge exiting $w$ contains $p$. For the converse, suppose that no selector-label of an edge exiting $w$ contains $p$. Note that since $p \in \Sigma$, $\neg\Diamond p \in \Omega$. It follows from (10.2)d that $\Diamond p \notin (\beta_w)_1$, and thence from (10.3)a, that for every $w' \in \mathbb{W}$, $\Diamond p \notin (\beta_{w'})_1$. Hence, by (10.1)a, for every $w' \in \mathbb{W}$, $\neg\Diamond p \in (\beta_{w'})_1$. Therefore, by (10.2)e and (10.1)a, for every $w' \in \mathbb{W}$, $p \notin (\beta_{w'})_1$. Hence, by (10.5), $p[\mathcal{M}] = \emptyset$.

In order to establish (10.6)c, suppose first that $q \in \Sigma$ and $p[\mathcal{M}] = q[\mathcal{M}]$. By (10.5), we may conclude that for all $w' \in \mathbb{W}$, $p \in (\beta_{w'})_1$ iff $q \in (\beta_{w'})_1$; the RHS of (10.6)c now follows immediately from the definition of $\mathcal{M}$ (the argument parallels that for the left to right direction of (10.6)b above). Finally, suppose that $p$ and $q$ are contained in exactly the same selector-labels of edges exiting $w$. We may suppose that this set is nonempty, for otherwise the result follows from (10.6)b. It follows at once that $\sigma_w(p) = \sigma_w(q)$. Note that since $p, q \in \Sigma$, $\Box(p \leftrightarrow q) \in \Omega$; we may then conclude, by (10.2)6, that $\Box(p \leftrightarrow q) \in (\beta_w)_1$. Hence, by (10.3)a, for all $w' \in \mathbb{W}$, $\Box(p \leftrightarrow q) \in (\beta_{w'})_1$. But then, by (10.1)a,b and (10.2)e, for all $w' \in \mathbb{W}$, $p \in (\beta_{w'})_1$ iff $q \in (\beta_{w'})_1$. We may conclude, by (10.5), that $p[\mathcal{M}] = q[\mathcal{M}]$.

*Induction Hypothesis*: Suppose that for all $w \in \mathbb{W}$,

(10.7) *(a)* $w \in \varphi[\mathcal{M}]$ *iff* $\varphi \in (\beta_w)_1$*;*
*(b)* $w \in \psi[\mathcal{M}]$ *iff* $\psi \in (\beta_w)_1$*;*
*(c)* $s(w, \varphi[\mathcal{M}])$ *and* $s(w, \psi[\mathcal{M}])$ *are determined.*

Induction Step: It follows immediately from (10.1)a,b and (10.7)a,b that for all $w \in \mathbb{W}$,

$$w \in (\varphi \wedge \psi)[\mathcal{M}] \quad \text{iff} \quad (\varphi \wedge \psi) \in (\beta_w)_1$$

and

$$w \in (\neg\varphi)[\mathcal{M}] \quad \text{iff} \quad (\neg\varphi) \in (\beta_w)_1.$$

It remains to show that

$$w \in (\varphi \preceq \psi)[\mathcal{M}] \quad \text{iff} \quad (\varphi \preceq \psi) \in (\beta_w)_1.$$

Suppose first that $w \in (\varphi \preceq \psi)[\mathcal{M}]$. Then, $s(w, \varphi[\mathcal{M}])$ and $s(w, \psi[\mathcal{M}])$ are both defined and $v_w(s(w, \varphi[\mathcal{M}])) = u(s(w, \varphi[\mathcal{M}])) \leq u(s(w, \psi[\mathcal{M}])) = v_w(s(w, \psi[\mathcal{M}]))$. It follows at once, by (10.2)g and (10.7)a,b, that $(\varphi \preceq \psi) \in (\beta_w)_1$. Suppose, on the other hand, that $w \notin (\varphi \preceq \psi)[\mathcal{M}]$. Then either at least one of $\varphi[\mathcal{M}]$ or $\psi[\mathcal{M}]$ is empty, or $v_w(s(w, \psi[\mathcal{M}])) = u(s(w, \psi[\mathcal{M}])) < u(s(w, \varphi[\mathcal{M}])) = v_w(s(w, \varphi[\mathcal{M}]))$. In either case, it follows from (10.7)a,b, (10.1)a, and (10.2)g, that $(\varphi \preceq \psi) \notin (\beta_w)_1$.

The extension of the definition of $s$ to $(\varphi \wedge \psi)[\mathcal{M}]$, $(\neg\varphi)[\mathcal{M}]$, and $(\varphi \preceq \psi)[\mathcal{M}]$ is justified exactly as in the basis of the induction. □

The above argument may be adapted to establish Corollaries 6.2 and 6.3.

*Proof of Corollary 6.2.* We modify the construction of $(\mathbb{W}, s, u, t)$ in the argument for the right to left direction of Proposition 10.4 to build a finite model satisfying $\theta$ from a mosaic $B$ for $\theta$. Let $\mathbb{W} = \bigcup_n W_n$ be the set of worlds constructed in the proof above, and let $n$ be the first stage such that for every $w \in \mathbb{W}$ there is an $m < n$ and a $w' \in W_m$ such that $\beta_w = \beta_{w'}$. We now close the construction of $(\mathbb{W}, s, u, t)$ at stage $n + 1$ by choosing the children of each $w \in W_n$ to be suitable worlds in $\bigcup_{m \leq n} W_m$ that satisfy the conditions in the construction of $(\mathbb{W}, s, u, t)$. □

*Proof of Corollary 6.3.* We modify the definition of brick as follows. A *reflexive brick* $B = \langle \Delta, \sigma, v \rangle$ is a brick that satisfies the following additional condition:

(10.8) *for all* $\varphi \in \Delta$, $s(\varphi) = \Delta$.

A *reflexive mosaic* is a mosaic composed of reflexive bricks; a reflexive mosaic for $\theta$ is defined similarly. Corollary 6.3 follows from the next proposition.

PROPOSITION 10.9. *For every* $\theta \in \mathcal{L}(\mathbb{P}, \mathbb{S})$*,* $\theta$ *is satisfiable in a reflexive model if and only if there is a reflexive mosaic for* $\theta$*.*

The proof of Proposition 10.9 is a straightforward adaptation of the proof of Proposition 10.4. The only subtlety is that in the definition of the tree-like model $(\mathbb{W}, s, u, t)$ we can no longer define $u$ in advance, but must define it by recursion following the recursive construction of the tree. We extend the definition of $u$ to the children of a world $w$ by choosing rational values for the children in such a way as to establish an isomorphism with the order induced by $v_w$, remembering that $w$ will be chosen as a child of $w$ in the obvious way, so that $u$ must retain the value for $w$ that was determined at stage $n$. □

**§11. Appendix: proof of Theorem 6.4.** Both sets of formulas are **NP**-hard since the satisfiability problem for nonmodal sentential logic is ptime-reducible to each. A straightforward application of the mosaic method yields the conclusion that each is in **NP**. □

**§12. Appendix: proof of Theorem 6.5.** We derive the compactness theorem for $\mathcal{L}(\mathbb{P}, \mathbb{S})$, where the signature $(\mathbb{P}, \mathbb{S})$ is countable, as a corollary to the compactness theorem for first-order logic. Our argument follows a standard strategy which proceeds via translating a modal language into a (fragment of) first-order logic. The translation essentially codifies the definition of satisfaction over some class of relational frames. In our case, this direct strategy requires modification since our "frames" are not (first-order) relational structures, in particular, the selection function has a "type 1" argument, the proposition from which a salient confirming representative is chosen. Moreover, the utility functions, whose range is the set of real numbers, present another obstacle to smooth "first-orderization" in the context of our compactness argument. To overcome these difficulties, the first step in our compactness proof is to "compile" a structure $\mathcal{M} = (\mathbb{W}, s, u, t)$ into a relational structure $\mathcal{F}_{\mathcal{M}}$ and to translate each sentence $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ to a first-order formula $\varphi^{\dagger}(x)$ with one free variable such that for all $w \in \mathbb{W}$,

(12.1) $w \in \varphi[\mathcal{M}]$ *iff* $\mathcal{F}_{\mathcal{M}} \models \varphi^{\dagger}[w]$.

Given $\mathcal{M} = (\mathbb{W}, s, u, t)$ of signature $(\mathbb{P}, \mathbb{S})$, we define $\mathcal{F}_{\mathcal{M}}$. The signature of $\mathcal{F}_{\mathcal{M}}$ consists of a unary relation symbol $Q_p$, for each $p \in \mathbb{P}$; a binary relation symbol $\leq_X$, for each $X \in \mathbb{S}$; and a binary relation symbol $R_{\varphi}$, for each sentence $\varphi$ of $\mathcal{L}(\mathbb{P}, \mathbb{S})$.

(12.2) *The interpretation of each relation symbol in the signature of $\mathcal{F}_{\mathcal{M}}$ is defined as follows (we suppress the superscript $\mathcal{F}_{\mathcal{M}}$ on each relation symbol):*

*(a)* $Q_p = t(p)$;
*(b) for all* $w, w' \in \mathbb{W}$, $w \leq_X w'$ *iff* $u_X(w) \leq u_x(w')$;
*(c) for all* $w, w' \in \mathbb{W}$, $R_{\varphi}(w, w')$ *iff* $w' = s(w, \varphi[\mathcal{M}])$.

Note that (12.2)c implies that $R_{\varphi}$ is the empty relation if and only if $\varphi[\mathcal{M}] = \emptyset$.

(12.3) *We now define, by recursion, for each sentence $\varphi$ of $\mathcal{L}(\mathbb{P}, \mathbb{S})$, its translation $\varphi^{\dagger}(x)$, a formula with one free variable in the first-order language of $\mathcal{F}_{\mathcal{M}}$.*

*(a)* $p^{\dagger} = Q_p(x)$;
*(b)* $(\varphi \wedge \psi)^{\dagger} = \varphi^{\dagger}(x) \wedge \psi^{\dagger}(x)$;
*(c)* $(\neg\varphi)^{\dagger} = \neg\varphi^{\dagger}(x)$;
*(d)* $(\varphi \preceq_X \psi)^{\dagger} = (\exists y)(\exists z)(\varphi^{\dagger}(y) \wedge \psi^{\dagger}(z) \wedge R_{\varphi}(x, y) \wedge R_{\psi}(x, z) \wedge y \leq_X z)$.

On the basis of (12.2) and (12.3), it is now easy to verify (12.1).
Next, we describe a first-order theory $D$ in the signature of $\mathcal{F}_{\mathcal{M}}$ such that

(12.4) *for every* $\mathcal{M}$, $\mathcal{F}_{\mathcal{M}} \models D$

and

(12.5) *for every countable first-order structure* $A$, *if* $A \models D$, *then for some* $\mathcal{M}$, $A = \mathcal{F}_{\mathcal{M}}$.

We proceed to describe $D$.

(12.6) *D consists of the following first-order sentences.*

*(a)* $(\forall x)(\forall y)(\forall z)(x \leq_X y \rightarrow (y \leq_X z \rightarrow x \leq_X z))$*, for each* $X \in \mathbb{S}$*;*

*(b)* $(\forall y)(y \leq_X y)$*, for each* $X \in \mathbb{S}$*;*

*(c)* $(\forall y)(\forall z)(y \leq_X z \vee z \leq_X y)$*, for each* $X \in \mathbb{S}$*;*

*(d)* $(\forall x)(\forall y)(R_\varphi(x, y) \rightarrow \varphi^\dagger(y))$*, for each* $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$*;*

*(e)* $(\exists x)\varphi^\dagger(x) \rightarrow (\forall x)(\exists y)(\forall z)(R_\varphi(x, z) \leftrightarrow y = z)$*, for each* $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$*;*

*(f)* $(\forall x)(\varphi(x) \leftrightarrow \psi(x)) \rightarrow (\forall x)(\forall y)(R_\varphi(x, y) \leftrightarrow R_\psi(x, y))$*, for each* $\varphi, \psi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$.

It is easy to verify (12.4) by direct inspection of the clauses of (12.6). In order to verify (12.5), we argue as follows. Let $A$ be a countable relational structure satisfying $D$. We define a structure $\mathcal{M} = (\mathbb{W}, s, u, t)$. First, let $\mathbb{W} = |A|$ and let $t(p) = Q_p^A$, for each $p \in \mathbb{P}$. Next, by (12.6)a–c, for each $X \in \mathbb{S}$, $\leq_X^A$ is a countable linear preorder. It follows from the universality of the rational numbers among countable linear orders that a utility function $u_X$ may be chosen so that for all $w, w' \in \mathbb{W}$, $w \leq_X w'$ if and only if $u_X(w) \leq u_X(w')$. For each $\varphi \in \mathcal{L}(\mathbb{P}, \mathbb{S})$ and $w, w' \in \mathbb{W}$, let $s(w, \varphi[\mathcal{M}]) = w'$ if and only if $R_\varphi(w, w')$. Finally, let $s(w, P)$ be an arbitrarily chosen element of $P$ for any proposition $P \subseteq \mathbb{W}$ which is not expressed by a sentence. It is easy to see that the structure $\mathcal{M}$ satisfies (12.5).

We now derive Theorem 6.5 from the compactness theorem for first-order logic. Let $T$ be a set of sentences of $\mathcal{L}(\mathbb{P}, \mathbb{S})$ and suppose that every finite subset $T' \subseteq T$ is satisfiable. It follows at once from (12.1) and (12.4) that for every finite $T' \subseteq T$, $\{\varphi^\dagger(c) \mid \varphi \in T'\} \cup D$ is satisfiable (here $c$ is a constant symbol). Therefore, by the Compactness and Löwenheim–Skolem Theorems for first-order logic, there is a countable structure $A$ such that $A \models \{\varphi^\dagger(c) \mid \varphi \in T\} \cup D$. Hence, by (12.1) and (12.5), $T$ is satisfiable. □

## BIBLIOGRAPHY


Andréka, H., Ryan, M., & Schobbens, P.-Y. (2002). Operators and laws for combining preferential relations. *Journal of Logic and Computation*, **12**, 12–53.

Blackburn, P., de Rijke, M., & Venema, Y. (2001). *Modal Logic*. New York, NY: Cambridge University Press.

Dietrich, F., & List, C. (2009). *A Reason-Based Theory of Rational Choice*. Technical Report, London School of Economics.

Gärdenfors, P. (1988). *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Cambridge, MA: MIT Press.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, **108**, 814–834.

Johnson, P. E. (1988). *Social Choice: Theory and Research*. New York, NY: Sage Publications.

Keeney, R. L., & Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*. Cambridge, UK: Cambridge University Press.

Lang, J., van der Torre, L., & Weydert, E. (2003). Hidden uncertainty in the logical representation of desires. In *Proceedings of Eighteenth International Joint Conference on Artificial Intelligence (IJCAI03).* San Francisco, CA: Morgan Kaufmann.

Levi, I. (1986). *Hard Choices: Decision Making under Unresolved Conflict*. Cambridge, UK: Cambridge University Press.

Lewis, D. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Liu, F. (2008). Changing for the better: Preference dynamics and agent diversity. PhD Thesis, ILLC, University of Amsterdam.

Messick. D. M. (1985). Social interdependence and decision making. In Wright, G., editor. *Behavioral Decision Making*. New York, NY: Plenum, pp. 87–109.

Osherson, D., & Weinstein, S. (To appear). Modal logic for preference based on reasons. In Tannen, V., editor. *Lecture Notes in Computer Science*. New York, NY: Springer.

Osherson, D., & Weinstein, S. (Forthcoming). Two approaches to the logic of preference.

Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *The Journal of Consumer Research*, **9**, 90–95.

Pettit, P. (2002). *Rules, Reasons, and Norms*. Oxford, UK: Oxford University Press.

Savage, L. J. (1954). *The Foundations of Statistics*. New York, NY: Wiley.

Sen, A. (1971). Choice functions and revealed preference. *Review of Economic Studies*, **38**, 307–317.

Stalnaker, R. (1968). A theory of conditionals. In Rescher, N., editor. *Studies in logical theory*. Oxford, UK: Blackwell.

Tentori, K., Osherson, D., Hasher, L., & May, C. (2001). Wisdom and aging: Irrational preferences in college students but not older adults. *Cognition*, **81**, B87–B96.

van Benthem, J., Girard, P. K., & Roy, O. (2009). Everything else being equal: A modal logic for *Ceteris Paribus* preferences. *Journal of Philosophical Logic*, **38**, 83–125.

van Benthem, J., & Liu, F. (2007). Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logic*, **17**, 157–182.

von Wright, G. H. (1963). *The Logic of Preference*. Edinburgh, UK: Edinburgh University Press.

DEPARTMENT OF PSYCHOLOGY
PRINCETON UNIVERSITY
PRINCETON NJ 08540 USA
*E-mail*: osherson@princeton.edu

DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA PA 19104 USA
*E-mail*: weinstein@cis.upenn.edu