Rank-based decision theory and conditional deontic logic

Wolfgang Spohn[1]

[1]Department of Philosophy, University of Konstanz, Germany

Abstract

Ranking theory is a well-established account of rational belief (= taking to be true) and its dynamics. Each formal representation of epistemic states should be extendable to a decision theory; ultimately, beliefs manifest themselves in rational action. However, the decision-theoretic extension of ranking theory is wanting. Giang and Shenoy's "A Qualitative Linear Utility Theory for Spohn's Theory of Epistemic Beliefs" (2000) is the only proposal so far. This paper will modify this proposal. The modification allows building a bridge to the field of (conditional) deontic logic, indeed advancing this field and thus instructing legal theory. This is to show that the modification is a fruitful one deserving further investigation and application.

Keywords. qualitative decision theory, ranking theory, ranking-theoretic utility, conditional deontic logic

1. Introduction

This paper is a hybrid between decision theory and deontic logic and thus potentially interesting to both economics and legal theory and potentially building a theoretical connection between the two disciplines.

More precisely, ranking theory (Spohn [18]) is by now a well-established epistemological theory. It is formally equivalent with possibility theory. Both were independently developed at around the same time (Spohn [17], Dubois and Prade [2]). In line with their respective authors, possibility theory is mainly known in computer science, and ranking theory in philosophy. In my view, ranking theory has the advantage of a firm interpretation (see below).

Now, for every proposed epistemic format—and there are many in the meantime (see, e.g., Halpern [7]) and they may all be understood as attempts at explicating so-called Knightian uncertainty—the issue arises how it might be extended to a decision theory. Decision theorists are not dogmatically attached to standard probability and utility theory (though, of course, these are the by far dominating accounts). The field often called qualitative decision theory is indeed growing considerably. However, before decision theorists get seriously interested in a novel epistemic format, they want to see this issue plausibly addressed, and justifiably so. Dubois, Prade, and Sabbadin [3] have attempted to satisfy this demand regarding possibility theory (see my comments below). Regarding ranking theory, the only attempt known to me is Giang and Shenoy [5]. This attempt is hardly perceived, though, presumably because it provably leads to quite a peculiar format of the utility function, which was apparently not suited to arouse the interest of decision theorists in ranking theory. I confess, originally I was not enthused, either.

In the meantime, however, I think that their approach is good and fruitful. Only one of their axioms needs to be changed and amended, and then we arrive at a very well interpretable utility function. It will be quite obvious then that a full-fledged decision theory can be developed in these terms in perfect analogy to standard decision theory.

More surprising is that we thereby gain a promising approach to conditional deontic logic. Deontic logic, originating from von Wright [22], is well established by now. But conditional deontic logic treating conditional obligations is still a contested field. Lewis [9, sect. 5.1] is perhaps most widely acknowledged, but also controversial, not in the least because there is still great uncertainty about how to best deal with the paradox of Chisholm [1], prima facie obligations, normative conflicts, etc. Finally, alethic-deontic logic, i.e., a logic studying the interaction of epistemic and deontic operators, particularly in their conditional form, is completely underdeveloped. There are few attempts, not leading far enough in my view.[1] Here, the rank-based decision theory introduced in this paper has the potential of changing the game. This is of potential interest also to legal theorists who in general are skeptical of using deontic logic precisely because their concerns are not well addressed in the standard parts and only confusingly addressed in the tentative parts of deontic logic.

The paper introduces ranking theory in Section 2. Sec-

[1]The most elaborate account is provided by Schurz [16], ch. 7, where he discusses alethic-deontic logics, and ch. 11, where he explores the possibility of what he calls analytic bridge principles (between is and ought).

tion 3 proceeds to explain the procedure of Giang and Shenoy [5] of extending ranking theory to a decision theory—with the important modification announced. The resulting decision theory is briefly explained and exemplified in Section 4. In Section 5 I develop the connection to and the potential enrichment of deontic logic. The resulting logic is summarized in the final Section 6.

2. Ranking theory

"Rank" and "ranking" are terms multiply used in formal epistemology and cognitive psychology.[2] The specific use employed here was introduced by Pearl [13], who chose the term "ranking function" for what I originally called "ordinal conditional functions" in Spohn [17]. Ranking functions are very much like probability functions. In Spohn [18, sect. 10.2] I provide a kind of formal translation of the latter into the former and claim that, in a way, the only thing to be done was to pursue this translation ever further, leading to one interesting application of ranking theory after the other. However, I insist that ranks must not be understood as probabilities or even as infinitesimal probabilities. They have their own good interpretation not to be assimilated to probabilities. Here they are:

Definition 2.1. κ is a *negative ranking function* for the finite set of possible worlds W iff κ is a function from $\mathcal{P}(W)$, the power set of W, into $\mathbb{N} \cup \{\infty\}$, the set of natural numbers plus infinity,[3] such that for all $A, B \subseteq W$:

(1) $\kappa(W) = 0$ and $\kappa(\emptyset) = \infty$,

(2) $\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\}$ (*the law of disjunction*).

If we abbreviate $\kappa(\{w\})$ by $\kappa(w)$, (2) entails that

(3) $\kappa(A) = \min\{\kappa(w) \mid w \in A\}$.

This is how the set function κ on $\mathcal{P}(W)$ results from a point function on W.

Negative ranks represent degrees of *disbelief* (this is why they are called negative). That is, $\kappa(A) = 0$ says that A is not disbelieved, and $\kappa(A) = n > 0$ says that A is disbelieved (to degree n). According to (1) and (2) we have $\min\{\kappa(A), \kappa(\bar{A})\} = \kappa(W) = 0$; that is, at least one of $\kappa(A)$ and $\kappa(\bar{A})$ must be 0. This means that you cannot take both, A and $\bar{A}$, to be false; this is a basic consistency requirement. But we may have $\kappa(A) = \kappa(\bar{A}) = 0$, in which case κ is unopinionated concerning A. *Belief* in A is the same as disbelief in $\bar{A}$ and is thus represented by $\kappa(\bar{A}) > 0$.[4]

This representation of belief via double negation may have hampered the reception of ranking functions. It is straightforward, though, to define a positive ranking function β by $\beta(A) = \kappa(\bar{A})$, which represents belief and degrees of belief directly. It is also possible to define a two-sided ranking function τ by $\tau(A) = \beta(A) - \kappa(A) = \kappa(\bar{A}) - \kappa(A)$, which represents both belief (by positive values) and disbelief (by negative values) simultaneously. This is perhaps the intuitively most appealing version. However, in formal treatments it is best to stick to negative ranking functions, as we do here.

The analogy of (1) and (2) to the probability axioms is salient. It is thus quite popular (see, e.g., Tan and Pearl [21], Goldszmidt and Pearl [6]) to interpret (negative) ranks as orders of magnitude of probabilities; if these orders of magnitude are taken relative to an infinitesimal base, the interpretation is even formally adequate. However, it is materially *inadequate*. Believing is taking to be true. This is a very familiar notion deeply entrenched in ordinary language. Still, belief can be firmer and less firm. Ranks represent both, belief and its degrees. However, these degrees differ from probabilities. Belief is not probability 1 (or maximal certainty by probabilistic standards[5]). Neither is it probability 1 minus some infinitesimal. This is still much too much certainty. Finally, it is not probability $1 - \epsilon$ for some positive real ϵ, as the famous lottery paradox shows. Probabilities do not seem to be able to grasp belief. We better take beliefs as what they are and capture them by ranks instead of trying to inadequately understand them through the probabilistic lens.[6]

This firm interpretation is also what distinguishes ranking functions from possibility measures. If we define Π by $\Pi(A) = q^{\kappa(A)}$ for some q with $0 < q < 1$, then Π is a possibility measure. However, degrees of possibility are intuitively quite indeterminate, and less than full possibility 1 does not sound like disbelief (= taking to be false). Because of this indeterminacy, I discuss here only ranking functions.[7]

We can also conditionalize ranking functions. If $\kappa(A) < \infty$, then the *conditional rank* of B *given* A is defined as

(4) $\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A)$.

This is equivalent to

(5) $\kappa(A \cap B) = \kappa(B \mid A) + \kappa(A)$ (*the law of conjunction*).

[2] See, e.g., Wakker [23], who uses the terms "rank" and "rank-dependence" throughout the book, but not in the sense used here.

[3] One may extend the definition to infinite possibility spaces. Let's keep things simple here. And one may consider other ranges of these functions, e.g., real numbers or ordinal numbers (see Spohn [18, p. 72]). For the present purpose, natural numbers are suited best.

[4] One may conceive of belief more strictly and define belief in A as $\kappa(\bar{A}) > z$ for some $z \geq 0$. This does not change the logic of belief. But it extends the range of neutrality in a perhaps intuitively desirable way, since this definition makes κ unopinionated concerning A iff $\kappa(A), \kappa(\bar{A}) \leq z$.

[5] By the way, if belief is not probability 1 and if knowledge is something like justified true belief (plus something else), then knowledge isn't probability 1, either, as is commonly assumed in economic texts.

[6] Of course, there are many attempts to solve the lottery paradox. Here is no place for entering a critical discussion of them.

[7] In Spohn [19] I provide a somewhat more detailed overview over various attempts to approach Knightian uncertainty.

The analogy of (5) to the probabilistic law of multiplication is obvious. $\kappa(B \mid A) > 0$ represents *disbelief* in B given A; and $\kappa(\bar{B} \mid A) > 0$ represents *belief* in B given A. Thus,

(6) either $\kappa(B \mid A) = 0$ or $\kappa(\bar{B} \mid A) = 0$ or both (*the conditional law of negation*).

That is, you cannot have contradictory beliefs under any condition A which you do not take to be impossible, i.e., for which $\kappa(A) < \infty$. Given definition (4), (6) is indeed equivalent with (2). This means that ranking theory essentially assumes nothing but conditional consistency of beliefs—and thus has strong normative foundations.[8]

With this notion of conditional ranks and conditional belief we can also explain belief change, namely by assuming conditionalization rules analogous to the probabilistic case. Indeed, since these rules can be applied several times, we can also explain iterated belief change, an issue with which other accounts of belief change had great difficulties (see Spohn [18, sect. 5.1 and 5.6] and Rott [14]). This is my main reason for preferring ranking theory over other formal representations of belief.

Ranks do not rely on vague intuitive feelings of strength. They come with a rigorous measurement theory. If you can answer any pair of questions of the form "to which beliefs do you still stick after giving up the belief in A?" in a coherent way, i.e., if your twofold belief contractions satisfy certain plausible requirements, then your ranking function is uniquely determined on a ratio scale (see Hild and Spohn [8]).

This must suffice as a short introduction to ranking theory. Everyone versed in probability theory can fathom that similar developments are forthcoming in ranking theory. Bayesian confirmation theory has a ranking analog. The entire theory of Bayes nets can be developed in ranking terms. This leads to a theory of deterministic causation in parallel to causal Bayes net theory. And so on. For all this see Spohn [18].

3. Preferences among Ranking Prospects

Alas, all of this is pure epistemology. How beliefs manifest in rational action is not addressed at all so far. Here, Giang and Shenoy [5] had a simple but ingenious idea: just copy the utility theory of Luce and Raiffa [10, sect. 2.5], i.e., replace probabilistic lotteries by what may be called ranking prospects (i.e., prospects with uncertain alternative outcomes the uncertainty of which is measured by ranks, not by probabilities[9]), translate the axioms for preferences over lotteries into axioms for preferences over such prospects, and then prove, as in standard utility theory, the existence of a utility function so that the preferences over prospects represent the expected utilities of these prospects (where expectation is taken in the ranking-theoretic sense), indeed the unique existence up to some class of transformations. This works surprisingly smoothly. As announced, however, at one point the exact copy led them astray in my view. So, let me explain my modified axiomatics; I shall indicate what is their original work and where I deviate. In the next section, I shall give a small numerical example of how this ranking-theoretic decision theory works.

Let $W = \{w_1, \dots, w_n\}$ be a finite set of possible outcomes. This is the same W as before. Outcomes may also be understood as entire possible worlds (as far as represented in the model). For $w_1, \dots, w_n \in W$ let $p = [r_1 w_1, \dots, r_n w_n]$ be a *simple prospect* if $r_j \in \mathbb{N} \cup \{\infty\}$ $(j = 1, \dots, n)$ and $r_i = 0$ for some i. Simple prospects are prospects. And if $p_1, \dots, p_m$ are prospects, then $[s_1 p_1, \dots, s_m p_m]$ is a prospect in turn, where $s_i \in \mathbb{N} \cup \{\infty\}$ $(i = 1, \dots, m)$ and $s_i = 0$ for some i. Prospects that are not simple are *compound*. Let P be the set of all prospects. In denoting a prospect, we may omit all terms with $r_i = \infty$. For $i < j$, $[rw_i, 0w_j]$ is called a *basic prospect*. For $i = 1$ and $j = n$, a basic prospect is called *primitive*. The prospect $[0w_j]$ also counts as basic, and $[0w_1]$ as primitive.

Now we assume a (weak) preference relation $\succcurlyeq$ over $W \cup P$ (where $\succ$ is the corresponding strict relation and $\approx$ the corresponding equivalence relation) and characterize it as follows:

Axiom 1 (ordering of alternatives)**.** $\succcurlyeq$ is a weak order over W.

Axiom 2 (transitivity)**.** $\succcurlyeq$ is a transitive relation over $W \cup P$.

These axioms go without saying. Without loss of generality we shall henceforth assume that $w_1 \succcurlyeq \dots \succcurlyeq w_n$. Now a crucial axiom is:

Axiom 3 (reduction of compound prospects)**.** Let $p^i = [r^i_1 w_1, \dots, r^i_n w_n]$ be simple prospects $(i = 1, \dots, m)$ and let $p = [s_1 p^1, \dots, s_m p^m]$ be a compound prospect. Define the reduced prospect p^* as $p^* = [r^*_1 w_1, \dots, r^*_n w_n]$, where $r^*_j = \min_{i \le m}(r^i_j + s_i)$. Then $p^* \approx p$.

Here, the definition of the r^*_j represents the ranking-theoretic form of expectation. Generally, the product of probabilities translates into the sum of ranks, and the sum of probabilities translates into the minimum of ranks.[10]

Axiom 4 (substitutability)**.** If for $p_i, p^*_i \in W \cup P$, $p^*_i \approx p_i$, then

[8]This way of justifying the formalism is not available to possibility theory.

[9]Hence the label "prospect". The term "lottery" sounds too heavily probabilistically loaded.

[10]Here, the crucial difference to the decision-theoretic extension of possibility theory emerges. The reduction of lotteries of Dubois, Prade, and Sabbadin [3, p. 463, Axiom 4] as well as their pessimistic and their optimistic criterion of forming 'expected utilities' deploy only ordinal operations, i.e., they take either the minimum (infimum) or the maximum (supremum). This is where the possibilistic and the ranking-theoretic approach to decision theory become incomparable.

$$[s_1 p_1, \ldots, s_m p_m] \approx \\ [s_1 p_1, \ldots, s_{i-1} p_{i-1}, s_i p_i^*, s_{i+1} p_{i+1}, \ldots, s_m p_m].$$

Obviously, axioms 3 and 4 reduce any iterated compound prospect to an equivalent simple prospect, since any compound prospect is ultimately built from simple ones.

Axiom 5 (identity)**.** For each $j = 1, \ldots, n$, $w_j \approx [0w_j]$.

This also seems to go without saying. Each outcome is equivalent to the prospect in which all other outcomes are impossible. Giang and Shenoy [5] are able to prove this, while I can't, as far as I see.

Axiom 6 (qualitative weak monotonicity for simple prospects)**.** If $p = [r_1 w_1, \ldots, r w_j, \ldots, r_n w_n]$ and $p' = [r_1 w_1, \ldots, s w_j, \ldots, r_n w_n]$ are simple prospects and $r \leq s \in \mathbb{N} \cup \{\infty\}$, then $p \succcurlyeq p'$.

This seems obvious as well. By making an outcome in a simple prospect less credible, i.e., more firmly disbelieved, the prospect cannot improve (and often, but not necessarily worsens—see Lemmas 3.1 and 3.2 below).

We also need an axiom asserting strict monotonicity:

Axiom 7 (restricted strict monotonicity for basic prospects)**.** If $[rw_i, 0w_j] \succ [0w_j]$ and $s > r$, then $[rw_i, 0w_j] \succ [sw_i, 0w_j]$.

We should add:

Axiom 8 (discreteness)**.** If $[rw_i, 0w_j] \succ \big[(r+1)w_i, 0w_j\big]$, then there is no prospect p such that $[rw_i, 0w_j] \succ p \succ \big[(r+1)w_i, 0w_j\big]$.

This is to ensure that the 'utility' representation to be constructed will be integer-valued.

Axioms 6–8 are a first slight deviation from Giang and Shenoy [5]. Instead of basic prospects they consider 'standard lotteries' between the (or a) best and the (or a) worst outcome, and they state their monotonicity condition only for their standard lotteries. But they assume strict monotonicity across all standard lotteries, while Axiom 7 postulates it only in a restricted way. And in their representation result they assume the representing utility function to be integer-valued, while we can prove this with Axiom 8. So, these seem marginal differences so far.

The crucial difference is in our final axioms. Giang and Shenoy [5] translate the continuity assumption of Luce and Raiffa [10, p. 27] into the assertion that each outcome is equivalent to a standard lottery between the best and the worst outcome. This is the literal translation. However, it has the consequence that the best outcome gets utility $+\infty$ and the worst $-\infty$. This may be defended as an artifact of the construction, but it indicates that we have somehow been led astray in the translation. The idea of weighing between the best and the worst outcome is still too probabilistic in spirit. Negative epistemic ranks measure a distance from the epistemic optimum, rank 0. And in this spirit, a utility function better compares each possible outcome only with the best one. This will have the effect that utilities will not weigh between positive and negative outcomes, as Giang and Shenoy have it, but only measure a distance from the desired optimum that we will fix to be 0. This idea leads us to

Axiom 9 (quasi-continuity)**.** For each $j = 1, \ldots, n$ there is an $r \in \mathbb{N}$ such that $[0w_j] \approx [rw_1, 0w_j]$.

And we require a final

Axiom 10 (indifference)**.** For all $i, j = 1, \ldots, n$ with $i < j$, $[0w_i] \approx [0w_i, 0w_j]$.

In other words, only the better alternative counts in a prospect with equal weights. This contradicts probabilistic intuitions. There, one would expect that the desirability of a mixture of two outcomes lies somewhere between the desirabilities of these outcomes. However, the ranking-theoretic conception is different. The epistemic rank (= degree of disbelief) of a disjunction is the minimum of the ranks of the disjuncts. And similarly then for a ranking-theoretically conceived 'disutility' function, where the 'disutility' of w_i *or* w_j equals the smaller 'disutility' of the two. This is what Axiom 10 expresses.

These axioms entail a series of lemmas:

Lemma 3.1. *For each $j = 1, \ldots, n$ there is a minimal $r \in \mathbb{N}$ such that, for all $s \geq r$, $w_j \approx [sw_1, 0w_j]$.*

Proof. Suppose that $[0w_j] \approx [rw_1, 0w_j]$ for some $r \in \mathbb{N}$, as guaranteed by Axiom 9, and $s \geq r$. Then we have $[sw_1, 0w_j] \approx \big[s[0w_1], 0[rw_1, 0w_j]\big]$ (according to Axioms 9 and 4) $\approx [rw_1, 0w_j]$ (according to Axiom 3). If this holds for some r and all $s \geq r$, then there must exist a minimal such r. □

Hence define q_j, for all $j = 1, \ldots, n$, to be the minimal s for which $w_j \approx [sw_1, 0w_j]$.

Lemma 3.2. *Each outcome is equivalent to a primitive prospect; i.e., for each $j = 1, \ldots, n$, $w_j \approx [q_j w_1, 0w_n]$.*

Proof. First, consider the case $w_j \approx w_n$. Then the assertion follows trivially from Lemma 3.1. So assume $w_j \succ w_n$ and hence $q_j < q_n$. Then we have $[q_j w_1, 0w_n] \preccurlyeq [q_j w_1, 0w_j, 0w_n]$ (Axiom 6) $\preccurlyeq [q_j w_1, 0w_j]$ (Axioms 10 and 4) $\approx w_j$. Similarly, we have $\big[(q_j - 1)w_1, 0w_n\big] \preccurlyeq w_j$. However, $[q_j w_1, 0w_j] \prec \big[(q_j - 1)w_1, 0w_n\big]$ (Axiom 7), and there is no prospect strictly between the two (Axiom 8). This entails that $[q_j w_1, 0w_n] \approx w_j$. □

Lemma 3.3. *Each simple or compound prospect is equivalent to a primitive prospect; i.e., for each $p \in P$ there is an $r \in \mathbb{N}$ such that $p \approx [rw_1, 0w_n]$.*

Proof. As mentioned, each compound prospect can be reduced to a simple prospect by Axioms 3 and 4. And with Lemma 3.2 each simple prospect reduces to a primitive prospect, again via Axioms 3 and 4. □

Lemma 3.3 corresponds to Lemma 1 of Giang and Shenoy [5], which states that, given a preference relation satisfying their axioms, each lottery is equivalent to a standard lottery in their sense. They then move to a representation theorem analogous to the one of utility theory, and we can do exactly the same:

Theorem 3.1. *Let $\succcurlyeq$ be a preference relation on $W \cup P$ satisfying Axioms 1–10. Then there exists exactly one 'disutility' function ν from $W \cup P$ into $\mathbb{N}$ such that for all prospects $p = [s_1 p^1, \ldots, s_m p^m]$, $\nu(p) = \min_{i \le m}(s_i + \nu(p_i))$ and $\succcurlyeq$ represents ν in the sense that for all p, $p^* \in W \cup P$, $p \succcurlyeq p^*$ iff $\nu(p) \le \nu(p^*)$.*[11]

Proof. It is clear that on the basis of Lemma 3.2 we can define a 'disutility' function ν for the outcomes (we will immediately discuss that other labels are more fitting) by setting $\nu(w_j) = q_j$. As in Lemma 3.3 we can extend ν to all prospects. And then the proof of Giang and Shenoy [5] of their theorem applies mutatis mutandis to our assertion that the ν thus constructed is the only one satisfying the conditions stated.[12] □

4. A RANK-BASED DECISION THEORY

How can we understand the construction of ν just given? "Utility", or even "disutility", is a misleading term, too closely wedded to the familiar constructions since von Neumann & Morgenstern. A better picture goes like this:

Outcomes, or worlds, are more or less desirable; they are as you want them to be, or they are more or less wanting. They are normatively ideal, i.e., as they should be, or they are normatively more or less deviant. If ideal outcomes are excluded for some reason, then at least one of the second best outcomes ought to be realized. If they are excluded as well, a third best outcome is required. And so on. ν measures the *normative deviancy* of outcomes or worlds.

The epistemic assessment of the outcomes of a prospect can be understood in a similar way. The outcomes with rank 0 are the only ones not excluded; they conform to the beliefs contained in the prospect. If I err and such an outcome does not obtain, then a least disbelieved outcome is believed to obtain, and so on again. That is, the ranks of the outcomes measure their *epistemic deviancy*. Recall my remark above that these deviancies are not merely ordered, as my explanation may suggest. They have numerical values, and these values are supported by a rigorous measurement theory.

What Lemma 3.1 tells us then is that there is a trade-off between epistemic and normative deviancy. If $\nu(w_j) = q_j$ and if s is the epistemic deviancy of w_1 in the prospect $[sw_1, 0w_j]$, then, for $s < q_j$, the epistemically deviant w_1 is still preferred to w_j; the epistemic deviancy is still small enough. $s = q_j$ results in indifference; there is an exact trade-off between the epistemic deviancy of w_1 and the normative deviancy of w_j. And for $s > q_j$ the prospect no longer worsens; the epistemic deviancy of w_1 becomes so large that only the normative deviancy of w_j counts in the prospect. This is how ranking-theoretic weighing works. This explanation should make our modified axioms 9 and 10 and Lemma 3.1 intelligible.

It should be clear how we can build a decision theory on our theorem, namely in perfect analogy to standard decision theory. Actions have prospects. Or, rather, we can identify actions with prospects, since we need not distinguish between actions opening identical prospects. Thus we have defined an expected disutility, or rather an expected deviancy, of actions, and the decision rule is to minimize this magnitude. We may unfold this in decision trees, introduce strategies, define their expected disutility, and so on.

Let me illustrate this with a toy example.[13] I have a moderately satisfying permanent job at my university. I am pleased to get an offer from another university for a really exciting, but demanding, research position, which, however, is not yet tenured. If I do well, I get tenure after five years (T); if not, I will have to look for a new job (N). My present university attempts to counter this offer and grants me access to a promotion program. If I am evaluated positively, I get the promotion (P); if not, I stay in my present job (J). Should I accept the offer (A)?

My preferences are clear: $T \succ P \succ J \succ N$. Let us quantify this in a plausible way by a 'disutility' function ν: $\nu(T) = 0$ (T is the 'normative ideal' in this situation), $\nu(P) = 3$, $\nu(J) = 5$, and $\nu(N) = 8$. Now it all depends on my beliefs, or rather my ranking function κ. I have doubts that I would do well enough in the research position. Thus, $\kappa(T \mid A) = x$ and $\kappa(N \mid A) = 0$ (x may be 0, too; but if $x > 0$, I rather believe or fear to fail). However, I am quite confident that I would get that promotion (i.e., at least I do not believe that I would fail). So, $\kappa(P \mid \bar{A}) = 0$ and $\kappa(J \mid \bar{A}) = y$. In this way, my options A and $\bar{A}$ are ranking prospects, and according to Theorem 3.1, their conditional expected utilities are $\nu(A) = \min(x + 0,\ 0 + 8) = \min(x, 8)$ and $\nu(\bar{A}) = \min(0 + 3,\ y + 5) = 3$. This shows that the potential doubts y of not getting the promotion have no weight. So, it all depends on my doubts to receive tenure in the research position. If the doubts are small enough

[11] Note that a prospect is the better, the lower its ν-value. This is why ν is better called a 'disutility' function.

[12] Dubois, Prade, and Sabbadin [3], Theorem 5 (p. 475) and Theorem 6 (p. 476), prove analogous representation theorems for their pessimistic and optimistic criterion mentioned in footnote 10. However, they make no claims about the uniqueness of their representation. Given their characterization of 'expected utilities' in purely ordinal terms, I conjecture that their utility function is measured only weakly on an ordinal scale.

[13] I have used this example already in Spohn [19]. However, there I still conceived utilities in the format of Giang and Shenoy [5].

($x < 3$), I should accept the offer; if they are large ($x > 3$), I should not; and if $x = 3$, my doubts just outweigh the advantage of T over P. This is a coarse-grained, but fairly plausible representation of my decision situation.

The example demonstrates the analogy between standard and ranking-theoretic decision theory. There may be a difference, though. In Luce and Raiffa's lotteries the probabilities were supposed to be objective and extraneously given. There was no measurement problem for them. Savage [15] was dissatisfied with this conception. For him, probabilities were just as subjective as utilities. Hence, he proved a much more demanding representation theorem, inferring both probabilities and utilities from preferences (most conspicuously presented in Fishburn [4, ch. 14]).

Now it seems that we should have the same concern as Savage. In prospects, the ranks were also extraneously given. However, there are no objective ranks; they just represent the subjective beliefs of the agent.[14] So, we should also provide a simultaneous measurement of ranks and disutilities, i.e., of epistemic and normative deviancies. Presently, I don't know how to do this; this is for future work.

Let me point out, though, that the concern is not so pressing. Above, I had mentioned that there exists a rigorous measurement procedure for ranks. And in principle, we can combine this with the present measurement of disutilities. The only drawback is that the two measurement procedures follow completely different methodologies.[15] So, it would still be nice to have a copy of Savage.

A final remark: One may wonder why the 'disutility' function ν is uniquely determined in Theorem 3.1. Aren't we used to the fact that utility functions are measured only on an interval scale? Well, the reason is that ν is directly measured by the ranks in the prospects, as Lemmas 3.1 and 3.2 display. This would be different in a joint measurement of epistemic and normative deviancies. As mentioned, epistemic ranks are measured on a ratio scale; the zero point is fixed, but the unit may vary. And then the 'disutilities' would covary with the epistemic ranks.

5. The bridge to deontic logic

Is rank-based decision theory more than a nice formal possibility? It can be fully developed, but is it worth doing so? I don't know whether it can help interpreting and explaining recalcitrant empirical data which had led to the development of alternatives to standard decision theory in the first place. As yet, nobody has looked at the data from this perspective. This might be rewarding, given the fact that the notion of belief is so deeply entrenched in our everyday practice and that ranking theory is a good, or perhaps the best, theory to capture rational belief and its dynamics, while this notion is not well grasped, in fact not even considered, by those alternatives. So, it is an open issue whether there is support from the empirical side for our version of qualitative decision theory.

Still, there is strong support from a very different, perhaps unexpected, side. Our theory may help advancing deontic logic. Of course, one may ask in turn how useful deontic logic is. However, it is taken seriously not only by philosophical logicians. We may derive at least relative support from this connection. Let me explain this point.

Standard deontic logic (see, e.g., McNamara and Van De Putte [11, sect. 2]) operates on the set L of sentences of propositional logic having truth conditions in a set W of possible worlds, and it adds a deontic operator O. Oφ is to mean that φ ought to be the case. Syntax easily allows to iterate the deontic operator. Semantics, though, has difficulties with this iteration. What exactly should it mean that it ought to be the case that φ ought to be the case? Let's better abstain from iterating O and take Oφ to be well-formed only if $\varphi \in L$, i.e., if φ does not contain O.

We may interpret the deontic operator in a perfectly subjective way, i.e., as expressing the normative conception of some agent. The agent has aims, desires, interests, also moral ideas, etc.; in sum, she wants the world to be a certain way. There is no need to think about obligations in a more objective sense. However, it is suggestive to take the legislator or a legal system to be such an 'agent'. In this interpretation, deontic logic becomes potentially interesting for legal theory.

Standardly, then, O simply follows the modal logic KD, i.e., it requires that O applies to a consistent and deductively closed set of sentences from L. That's all. Note that the doxastic logic of the belief operator Bel is exactly the same; logic requires belief just to be consistent and deductively closed.[16] (I am a bit more explicit in Section 6.)

So far, this is not very substantial. Standard extensions concern, e.g., the inclusion of actions and the relation to time (see, e.g., McNamara and Van De Putte [11, sect. 5.2–3]). Here, I am interested in the extension to conditional obligations. I am obliged (or I want or desire) not to hurt anyone, but *if* I do, I am obliged (or I want or desire) to help her. Normative conflicts usually arise from conditional obligations or wants. Under some conditions I should do a, under other conditions I should do b instead. If there is a thunderstorm outside, I should stay in the house. If there is a fire in the house, I should leave it immediately. Now, what if both conditions obtain at the same time?[17]

The standard account of conditional obligations as-

[14] For possible ways of objectivization see, however, Spohn [18, ch. 15].

[15] A further technical incoherence lies in the fact that the account of Hild and Spohn [8] works with real-valued ranking functions, while in this paper ranks are natural numbers.

[16] Beware, though. Consistency and deductive closure, or other axioms, of Bel and of O are not a matter of logic, properly understood. Rather, they are rationality postulates.

[17] This is structurally the same problem as the famous Nixon diamond in default logic.

sumes a kind of preference order over the set of worlds. Ideally, one of the best worlds should obtain. Given that such an ideal is not (to be) realized, one of the second-best worlds should obtain. And so on. For the standard logic for conditional obligations resulting from this assumption see the final Section 6.

Now note that the semantics of conditional obligations can just as well be stated in ranking-theoretic terms. The paradigm is the ranking-theoretic treatment of conditional belief introduced in Section 2. This treatment works exactly the same in normative terms. We can start with a point function ν on the set W of worlds (or outcomes) and extend it to a ranking function defined for all propositions in $\mathcal{P}(W)$ via (3). Thus, a preference order as standardly assumed is established as well. Then, if A and B are, respectively, the truth conditions of φ and ψ, the truth condition of $\mathrm{O}(\psi \mid \varphi)$ is $\nu(\bar{B} \mid A) > 0$ (literally: given A, non-B is forbidden; i.e., B is mandatory). The logic resulting from this ranking-theoretic semantics is precisely the standard logic for conditional obligations (see Section 6).

However, this treatment alone does not yet settle the various ambiguities besetting the notion of conditional obligation (want, desire). Here is no place for unfolding this very confusing field; it is as tentative as the field of qualitative decision theory. In Spohn [20, sect. 6] I have attempted a broader comparative discussion. Here, I can only present my own account. The point is to explain how this account is entailed by the rank-based decision theory presented in the previous sections.

In this confusing field, the most illuminating puzzle seems to me to be the paradox of Chisholm [1]:

(i) Jones ought to go to the aid of his neighbors.

(ii) If Jones goes to the aid of his neighbors, then he ought to tell them that he is coming.

(iii) If Jones does not go to the aid of his neighbors, then he ought not to tell them that he is coming.

(iv) Jones does not go to the aid of his neighbors.

(i) and (ii) seem to imply that Jones ought to tell his neighbors that he is coming, while (iii) and (iv) imply that he ought not. Contradiction. The puzzle is that intuitively (i)–(iv) are consistent; in particular, (i) and (iv) are not contradictory. The usual response to this puzzle is that there are two different detachments at work here. (iii) and (iv) deploy factual detachment, while (i) and (ii) appeal to some kind of deontic detachment. And then the issue is: when is which kind of detachment appropriate? This is how Nute [12] frames the topic.

I follow here Spohn [20, sect. 4–5], where I argue that the ambiguity runs deeper, that there are indeed two different deontic operators at work. (ii) should more explicitly read as

(v) If Jones ought to go to the aid of his neighbors, then he ought to tell them that he is coming.

But we are often sloppy and use the simpler (ii) to express (v). If we read (ii) as (v), then it becomes clear that (i) and (ii) express what is desirable as such or by itself. They describe a purely normative point of view, a *purely normative* ought O^-, which can be conceived as being governed by the above ranking function ν, which might be called purely normative as well. This purely normative point of view is also full of normative fallback positions that are provided by the preference order contained in ν. Indeed, if we want to account for rational change of the pure norms, we better do it in terms of the ranking function ν and not in terms of the contained preference order. This is a lesson from the ranking-theoretic account of belief change.

By contrast, (iii) and (iv) do not address the purely normative point of view. They address what the norms are when facts diverging from what ought to be the case are assumed, not what they are when other norms are assumed. They are about fact-regarding (or rather belief-regarding) norms, about a *fact-regarding* ought O^+. So, if two different oughts are involved in Chisholm's paradox, then no contradiction can follow from (i)–(iv).

How, though, are we to account for the fact-regarding ought O^+? Well, we have already done so in Section 3. I just proposed to extend the normative point function ν to all propositions via (3). This embodies the purely normative point of view. However, in Section 3 the extension of the point function ν for outcomes to prospects worked in a different way. It took the epistemic values assigned to the outcomes or worlds in prospects into account as well and derived expected 'disutilities' in the way specified in Section 3.

Now we can conceive of propositions also as prospects. Instead of writing the epistemic ranks directly into the prospects, we can assume an epistemic ranking function κ for worlds (and propositions). Then a proposition A is a prospect in the sense that it presents the worlds in A normatively evaluated by ν with an epistemic deviancy provided by κ. Thus conceived, our representation theorem 3.1 results in a *fact-regarding normative* ranking function ν_κ pointwise defined by

(7) $\nu_\kappa(w) = \kappa(w) + \nu(w) - \min_{w' \in W}\big(\kappa(w') + \nu(w')\big)$

– the last term is just a normalizing factor – and then extended to propositions in the usual way by

(8) $\nu_\kappa(A) = \min_{w \in A} \nu_\kappa(w)$.

We had already explained how best to understand this fact-regarding normative ranking function ν_κ: each world w is more or less epistemically deviant, as measured by κ, and (purely, we may add now) normatively deviant, as measured by ν. And these two deviancies add up to a (normalized) overall deviancy, as measured by ν_κ. So, all three functions, the epistemic κ, the purely normative ν, and the fact-regarding normative ν_κ, are ranking functions.[18]

Let me illustrate the interaction of these three ranking functions with the venerable practical syllogism, which is often thought to be the most basic rule of practical reasoning. Here is an example:

(9) Premise 1: I ought not to get wet.

Premise 2: I would get wet without an umbrella.

Conclusion: I ought to take an umbrella.

Clearly, this inference mixes normative and descriptive premises: Premise 1 is normative, Premise 2 is descriptive, and the conclusion is normative in turn. With the above distinctions this inference can be reconstructed in the following way:

Premise 1 expresses a pure norm (in that context), represented by ν (where W = "I get wet" and U = "I take an umbrella"):

ν	W	$\bar{W}$
U	2	0
$\bar{U}$	2	0

Thus, $\nu(W) = 2$, i.e., $\bar{W}$ ought to be the case. However, from the purely normative point of view I am indifferent regarding the umbrella, i.e., $\nu(U) = \nu(\bar{U}) = 0$.

Premise 2 expresses an empirical belief, represented by the epistemic ranking function κ:

κ	W	$\bar{W}$
U	3	0
$\bar{U}$	0	3

In fact, this represents the biconditional "I would get wet only without an umbrella". (This makes the presentation of the example easier.).

Given the mixed premises, we should construe the conclusion as stating a fact-regarding norm. According to (7), ν and κ add up to ν_κ:

ν_κ	W	$\bar{W}$
U	5	0
$\bar{U}$	2	3

This represents that, from a fact-regarding point of view, both U and $\bar{W}$ are mandatory (since $\nu_\kappa(\bar{U} \text{ or } W) > 0$) and thus confirms the conclusion, as desired.

Note, however, that the practical syllogism is *not* logically valid. It is only a defeasible inference; i.e., it may be defeated by additional premises (such as "I get an umbrella only by stealing it"). This defeasibility can well be accounted for by the ranking-theoretic representation.

[18] Note that in Section 3, I used ν indiscriminately for a 'disutility' function. Now, after distinguishing a purely normative and a fact-regarding normative ranking function, it turns out that the ν from Section 3 is rather the present fact-regarding ν_κ.

6. The resulting logic

Why define the fact-regarding normative ranking function ν_κ by (7) and (8) and not in any other way? This is justified precisely by our representation theorem 3.1. Spohn [20, sect. 7] offers a different justification. According to (4), we can introduce conditional ranks for each of the three ranking functions. For the fact-regarding normative ν_κ, however, a condition A can be given in three different ways: as an epistemic condition (what if A were the case?), as a purely normative condition (what if A as such ought to be the case?), or as a normative condition in the fact-regarding sense (what if, in view of the facts, A ought to be the case?).

This results in three different definitions:

(10) $\nu_\kappa(B \,|_{\mathrm{e}}\, A) = \min_{w \in A \cap B}\big[\nu(w) + \kappa(w \,|\, A)\big] - \min_{w \in A}\big[\nu(w) + \kappa(w \,|\, A)\big]$,

(11) $\nu_\kappa(B \,|_{\mathrm{p}}\, A) = \min_{w \in A \cap B}\big[\nu(w \,|\, A) + \kappa(w)\big] - \min_{w \in A}\big[\nu(w \,|\, A) + \kappa(w)\big]$, and

(12) $\nu_\kappa(B \,|_{\mathrm{f}}\, A) = \min_{w \in A \cap B}\big[\nu(w \,|\, A) + \kappa(w \,|\, A)\big] - \min_{w \in A}\big[\nu(w \,|\, A) + \kappa(w \,|\, A)\big]$,

where $|_{\mathrm{e}}$ denotes epistemic conditionalization, $|_{\mathrm{p}}$ denotes purely normative conditionalization, and $|_{\mathrm{f}}$ denotes fact-regarding normative conditionalization. The second terms, respectively, are again normalization terms ensuring that $\nu_\kappa(A \,|_{\mathrm{e}}\, A) = \nu_\kappa(A \,|_{\mathrm{p}}\, A) = \nu_\kappa(A \,|_{\mathrm{f}}\, A) = 0$, as it should be. This looks like an overkill of conditionalizations. However, it is easy to show that the distinctions make no difference and that all three agree with the definition (4) of conditional ranks:

(13) $\nu_\kappa(B \,|_{\mathrm{e}}\, A) = \nu_\kappa(B \,|_{\mathrm{p}}\, A) = \nu_\kappa(B \,|_{\mathrm{f}}\, A) = \nu_\kappa(B \,|\, A) = \nu_\kappa(A \cap B) - \nu_\kappa(A)$.

This may explain why we carelessly switch between (ii) and (v).

In Spohn [20, sect. 7] I suggest reversing the picture. There it is proven that, if ν_κ is any function of ν and κ and if (13) holds, then ν_κ must take the additive form (7). So, this is another reason for accepting the above definition of fact-regarding normative ranks.

Let me summarize the logic that results from all these considerations, as far as it is presently known. As said, we start with the set L of sentences of propositional logic having truth conditions in a set W of possible worlds. We extend our language by the operators Bel, O^-, O^+, and a conditional $\rhd$ to a language L_{C} in the following way: If φ and ψ are sentences of L, then $\mathrm{Bel}(\varphi)$, $\mathrm{O}^-(\varphi)$, $\mathrm{O}^+(\varphi)$, $\mathrm{Bel}(\varphi) \rhd \mathrm{Bel}(\psi)$, $\mathrm{O}^-(\varphi) \rhd \mathrm{O}^-(\psi)$, $\mathrm{Bel}(\varphi) \rhd \mathrm{O}^+(\psi)$, $\mathrm{O}^-(\varphi) \rhd \mathrm{O}^+(\psi)$, and $\mathrm{O}^+(\varphi) \rhd \mathrm{O}^+(\psi)$ are sentences of L_{C}. Finally, if φ and ψ are sentences of L_{C}, then propositional combinations of φ and ψ are also sentences of L_{C}. We keep things simple and have no nestings of operators or conditionals. Instead of introducing binary operators like $\mathrm{Bel}(\cdot \,|\, \cdot)$ I have introduced a single conditional $\rhd$, which, however, is syntactically restricted in the way specified.[19] Let L_{Bel} be the restriction of L_{C} to sentences not containing O^- or O^+, and let L_{O^-} and L_{O^+} be defined similarly. So much for the syntax.

The semantics for propositional combinations is familiar and need not be specified. And we have already indicated the semantics for the operators. It runs as follows: Let W be the set of all propositional evaluations of the sentences in L. Let A and B, respectively, be the truth conditions of φ and ψ in L (i.e., the set of propositional evaluations in which φ and ψ are true). And let κ be an (epistemic) ranking function for W, ν be a (purely normative) ranking function for W, and ν_κ as defined by (7) and (8). Then

(14) $\text{Bel}(\varphi)$ is true wrt. κ iff $\kappa(\bar{A}) > 0$,

$\text{Bel}(\varphi) \rhd \text{Bel}(\psi)$ is true wrt. κ iff $\kappa(\bar{B} \mid A) > 0$ or $\kappa(\bar{B} \mid A)$ is not defined,

$\text{O}^-(\varphi)$ is true wrt. ν iff $\nu(\bar{A}) > 0$,

$\text{O}^-(\varphi) \rhd \text{O}^-(\psi)$ is true wrt. ν iff $\nu(\bar{B} \mid A) > 0$ or $\nu(\bar{B} \mid A)$ is not defined,

$\text{O}^+(\varphi)$ is true wrt. ν and κ iff $\nu_\kappa(\bar{A}) > 0$,

$\text{Bel}(\varphi) \rhd \text{O}^+(\psi)$ is true wrt. ν and κ iff $\nu_\kappa(\bar{B} \mid_{\text{e}} A) > 0$ or $\nu_\kappa(\bar{B} \mid_{\text{e}} A)$ is not defined,

$\text{O}^-(\varphi) \rhd \text{O}^+(\psi)$ is true wrt. ν and κ iff $\nu_\kappa(\bar{B} \mid_{\text{p}} A) > 0$ or $\nu_\kappa(\bar{B} \mid_{\text{p}} A)$ is not defined, and

$\text{O}^+(\varphi) \rhd \text{O}^+(\psi)$ is true wrt. ν and κ iff $\nu_\kappa(\bar{B} \mid_{\text{f}} A) > 0$ or $\nu_\kappa(\bar{B} \mid_{\text{f}} A)$ is not defined.

Finally, a sentence φ of L_{C} is logically true iff it is true in all propositional evaluations of W and wrt. all ranking functions κ and ν for W.

Partially, the logics for these semantics are well-known. For all languages considered, modus ponens is the only inference rule we need. For the restricted language L_{Bel} we have the following axioms (where $\rightarrow$, material implication, binds more strongly than $\rhd$):

(15) all tautologies (truths of propositional logic),

$\text{Bel}(\varphi) \leftrightarrow \text{Bel}(\top) \rhd \text{Bel}(\varphi)$, where $\top$ is any tautology,

$\text{Bel}(\varphi) \rhd \text{Bel}(\varphi)$,

$\text{Bel}(\neg\varphi) \rhd \text{Bel}(\varphi) \rightarrow \text{Bel}(\psi) \rhd \text{Bel}(\varphi)$,

$\text{Bel}(\varphi) \rhd \text{Bel}(\psi \rightarrow \chi) \rightarrow \big(\text{Bel}(\varphi) \rhd \text{Bel}(\psi) \rightarrow \text{Bel}(\varphi) \rhd \text{Bel}(\chi)\big)$ (conditional deductive closure),

$\neg\big(\text{Bel}(\varphi) \rhd \text{Bel}(\neg\psi)\big) \rightarrow \big(\text{Bel}(\varphi \wedge \psi) \rhd \text{Bel}(\chi) \leftrightarrow \text{Bel}(\varphi) \rhd \text{Bel}(\psi \rightarrow \chi)\big)$ (rational monotony).

This logic is the basic system V of Lewis [9, sect. 6.1] (with our syntactical restrictions), which is proved there to be sound and complete. Mutatis mutandis, these assertions exactly apply to the other two restricted languages L_{O^-} and L_{O^+}.

Of course, the interesting question is: what is a sound and complete logic for the full language L_{C}? Alas, I don't have an answer. Of course, all the axioms for the restricted languages are also axioms for L_{C}. Moreover, we can't have axioms connecting Bel and O^- by themselves, since they are governed by independent ranking functions. However, both interact with O^+ via (7) and (8). Trivial axioms are

(16) $\big(\text{O}^+(\varphi) \rhd \text{O}^+(\psi)\big) \leftrightarrow \big(\text{Bel}(\varphi) \rhd \text{O}^+(\psi)\big)$ and $\big(\text{O}^+(\varphi) \rhd \text{O}^+(\psi)\big) \leftrightarrow \big(\text{O}^-(\varphi) \rhd \text{O}^+(\psi)\big)$.

This is just the consequence of (13). But beyond this, it is an open issue how this interaction is reflected in suitable axioms. The issue looks difficult, since, as far as I know, it is not just a matter of transferring and adapting familiar logics and semantics. It must be left for future research.

However, the point is that we now have a precise model for the interaction of beliefs and norms. This was wanting so far. My aim was to explain how this flows from an investigation of an exotically looking corner of decision theory and how we thereby reach out to quite a different field potentially leading to a better understanding of legal logic. This is certainly a good reason to continue with this investigation.

[19] It is unconventional to avoid factual sentences φ and ψ as arguments of the conditional $\rhd$. I do this because I focus here on the use of the conditional in epistemic and normative contexts and do not engage in discussing the intricacies of subjunctive and counterfactual conditionals.

REFERENCES

[1] Roderick M. Chisholm. "Contrary-to-Duty Imperatives and Deontic Logic". In: *Analysis* 24 (1963), pp. 33–36. DOI: 10.2307/3327064.

[2] Didier Dubois and Henri Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. New York: Plenum Press, 1988.

[3] Didier Dubois, Henri Prade, and Régis Sabbadin. "Decision-Theoretic Foundations of Qualitative Possibility Theory". In: *European Journal of Operational Research* 128 (2001), pp. 459–478. DOI: 10.1016/S0377-2217(99)00473-7.

[4] Peter C. Fishburn. "Utility Theory". In: *Management Science* 14 (1968), pp. 335–378. DOI: 10.1287/mnsc.14.5.335.

[5] Phan Hong Giang and Prakash P. Shenoy. "A Qualitative Linear Utility Theory for Spohn's Theory of Epistemic Beliefs". In: *Uncertainty in Artificial Intelligence*. Ed. by C. Boutilier and M. Goldszmidt. Vol. 16. San Francisco: Morgan Kaufmann, 2000, pp. 220–229. DOI: 10.48550/arXiv.1301.3858.

[6] Moisés Goldszmidt and Judea Pearl. "Qualitative Probabilities for Default Reasoning, Belief Revision, and Causal Modeling". In: *Artificial Intelligence* 84.1–2 (July 1996), pp. 57–112. DOI: 10.1016/0004-3702(95)00090-9.

[7] Joseph Y. Halpern. *Reasoning about Uncertainty*. Second edition. Cambridge, Mass.: MIT Press, 2017.

[8] Matthias Hild and Wolfgang Spohn. "The Measurement of Ranks and the Laws of Iterated Contraction". In: *Artificial Intelligence* 172 (2008), pp. 1195–1218. DOI: 10.1016/j.artint.2008.03.002.

[9] David Lewis. *Counterfactuals*. Oxford: Blackwell, 1973.

[10] R. Duncan Luce and Howard Raiffa. *Games and Decisions: Introduction and Critical Survey*. New York: Wiley, 1957.

[11] Paul McNamara and Frederik Van De Putte. "Deontic Logic". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Fall 2022 edition. 2022. URL: https://plato.stanford.edu/archives/fall2022/entries/logic-deontic/.

[12] Donald Nute, ed. *Defeasible Deontic Logic*. Dordrecht: Kluwer, 1997.

[13] Judea Pearl. "System Z: A Natural Ordering of Defaults with Tractable Applications to Default Reasoning". In: *Proceedings TARK-90*. Ed. by Rohit Parikh. San Mateo, Calif.: Morgan Kaufmann, 1990, pp. 121–135.

[14] Hans Rott. "Shifting Priorities: Simple Representations for Twenty-seven Iterated Theory Change Operators". In: *Towards Mathematical Philosophy*. Ed. by David Makinson, Jacek Malinowski, and Heinrich Wansing. Dordrecht: Springer, 2009, pp. 269–296.

[15] Leonard J. Savage. *The Foundations of Statistics*. 2nd ed. Dover 1972. New York: Wiley, 1954.

[16] Gerhard Schurz. *The Is–Ought Problem: A Study in Philosophical Logic*. Dordrecht: Kluwer, 1997.

[17] Wolfgang Spohn. "Ordinal Conditional Functions: A Dynamic Theory of Epistemic States". In: *Causation in Decision, Belief Change, and Statistics*. Ed. by W. L. Harper and B. Skyrms. Vol. II. Dordrecht: Kluwer, 1988, pp. 105–134.

[18] Wolfgang Spohn. *The Laws of Belief: Ranking Theory and Its Philosophical Applications*. Oxford: Oxford University Press, 2012.

[19] Wolfgang Spohn. "Knightian Uncertainty Meets Ranking Theory". In: *Homo Oeconomicus* 34 (2017), pp. 293–311. DOI: 10.1007/s41412-017-0060-5.

[20] Wolfgang Spohn. "Defeasible Normative Reasoning". In: *Synthese* 197 (2020), pp. 1391–1428. DOI: 10.1007/s11229-019-02083-2.

[21] Sek-Wah Tan and Judea Pearl. "Qualitative Decision Theory". In: *Proceedings of the Twelfth National Conference on Artificial Intelligence*. Ed. by B. Hayes-Roth and R. Korf. Vol. 2. Menlo Park, Calif.: AAAI Press, 1994, pp. 928–933.

[22] Georg Henrik von Wright. "Deontic Logic". In: *Mind* 60 (1951), pp. 1–15. JSTOR: 2251395.

[23] Peter P. Wakker. *Prospect Theory for Risk and Ambiguity*. Cambridge: Cambridge University Press, 2010.