General Belief Measures

Emil Weydert
Max-Planck-Institute for Computer Science
Im Stadtwald, D-66123 Saarbrücken, Germany
emil@mpi-sb.mpg.de

Abstract
Probability measures by themselves are known to be inappropriate for modeling the dynamics of plain belief and their excessively strong measurability constraints make them unsuitable for some representational tasks, e.g. in the context of first-order knowledge. In this paper, we are therefore going to look for possible alternatives and extensions. We begin by delimiting the general area of interest, proposing a minimal list of assumptions to be satisfied by any reasonable quasi-probabilistic valuation concept. Within this framework, we investigate two particularly interesting kinds of quasi-measures which are not or much less affected by the traditional problems:

• Ranking measures, which generalize Spohn-type and possibility measures.
• Cumulative measures, which combine the probabilistic and the ranking philosophy, allowing thereby a fine-grained account of static and dynamic belief.

1 INTRODUCTION
The successful acting of cognitive agents in complex, opaque and dynamic worlds depends on their ability to manipulate huge amounts of incomplete and uncertain information in a reasonable way. The representation of soft and partial knowledge together with the formalization of the corresponding reasoning patterns are therefore major issues in artificial intelligence. Very roughly, we can distinguish the following approaches to the modeling of belief, i.e. of uncertain or revisable knowledge:

• Strictly qualitative.
Belief states are represented by specific prioritized sets of sentences (syntactic) [Weydert 92, Nebel 92], or by some preference relation on possible worlds structures inducing an epistemic-entrenchment-like ordering on propositions (semantic) [Gärdenfors, Makinson 88, 91].

• Strictly quantitative.
Belief states are characterized by classical probabilistic measures [Pearl 88] or by alternative numerical accounts, e.g. belief functions [Shafer 76].

• Semi-quantitative or semi-qualitative.
Belief states are described by rough qualitative measures assigning orders of magnitude [e.g. Spohn 88].

For most proposals, there are more or less suitable application contexts, but altogether they also exhibit some more or less serious shortcomings. Prioritized sentential belief models are interesting because of their simplicity and the absence of omniscience assumptions, but the lacking semantic foundations and the resulting ad hoc character relativize their usefulness. More generally, strictly qualitative syntax- or semantic-based comparative frameworks are not fine-grained enough to handle the complexities of real-world knowledge. For instance, we cannot easily model belief strength differences or dependency constraints and our decision-theoretic considerations have to stay rather rudimentary.

Among the numerical approaches, probability theory is certainly the best investigated and most successful formalism for modeling uncertain propositional knowledge. Even if its naive use might sometimes cause problems, it still appears to be the ideal reference formalism for more sophisticated or special-purpose representations, not to mention the practical importance of the powerful tools it provides. But there are several general problems which have to be addressed, in particular if we want to handle changing or first-order belief states.

• Agents cannot always associate exact numbers - e.g. probabilities reflecting betting intentions and/or statistical knowledge - with propositions. Very often, numerical values are undefined, unknown, irrelevant, cumbersome or untractable, and more transparent, simpler models of reality are required, lowering the computational costs.

• Standard conditionalization procedures in the traditional probabilistic context don't allow us to revise plain belief if we choose to implement it, quite naturally, by assigning (subjective) probability 1.

• The representation of complex forms of (e.g. first-order) knowledge dealing with infinite structures is frequently blocked by unsatisfiable measurability constraints.
• There are intuitive flaws, for instance the gap between impossibility and probability 0 or the nonexistence of uniform distributions on countably infinite sets.

To a certain extent, all these problems can be solved in semi-qualitative quasi-probabilistic formal accounts based on abstract ranking measures. Their characteristic feature is that the value associated with the union of two sets is assumed to be, w.r.t. a given ranking order, just the maximum of the values attributed to the individual sets. This means less commitments by lower precision. Homogeneous countably infinite sets and proper impossibility are no longer barred and ranking measure can be lifted to the full power set. Related concepts have been used fairly successfully for multiple revisions of plain belief [Spohn 88], default reasoning [Weydert 91, Dubois, Prade 91, Goldszmidt, Pearl 92] and belief modeling [Weydert 94]. Their main disadvantage is that they can only provide a very rough picture, suitable for interpreting defaults but not for a (really) fine-grained decision-theoretic analysis. So, in some sense, quantitative and qualitative formalisms seem to offer complementary features which we might try to combine in a new type of mixed framework, possibly powerful enough for more realistic dynamic models of graded first-order belief.

In this paper, we are going to develop such an integrated approach by considering natural generalizations of probability measures. We shall proceed in four steps. To begin with, we shall delimit the area of interest and single out the minimal conditions for reasonable quasi-probabilistic valuation formalisms. Next, we shall introduce and discuss two particularly interesting subclasses of nonclassical quasi-measures. On one hand, we have the ranking measures, a coarse-grained semi-qualitative notion. On the other hand, we have the cumulative measures, a fine-grained extended quantitative-qualitative measure concept, which tries to accumulate the best of both worlds. To conclude, we shall sketch how cumulative measures may be used to model static and dynamic belief.

2 QUASI-MEASURES

Traditional probabilistic measure spaces can be seen as triples of the form \((\mathcal{B}, \mathcal{P}, \mathcal{V})\), where \(\mathcal{B} = (\mathcal{B}, \cup, \cap, -, 0, 1)\) is a \(\sigma\)-algebra of events, i.e. a boolean algebra closed under countable joins (unions) and meets (intersections) with top 1 and bottom 0, \(\mathcal{V} = (\mathbb{R}^+, +, \cdot, 0, 1, \leq)\) is the positive half of the ordered real number field, and the probability measure \(\mathcal{P}\) is a function from \(\mathcal{B}\) to \(\mathbb{R}^+\) s.t. \(\mathcal{P}(1) = 1\) and \(\mathcal{P}(\cup A_i) = \sum_{i} \mathcal{P}(A_i)\) for every countable set of \(\cap\)-disjoint \(A_i\) (\(\sigma\)-additivity). For each such space, there is a corresponding conditional probability measure \(\mathcal{P}(1) : \mathcal{B} \times \mathbb{R}^+ \to \mathcal{B}\), which verifies \(\mathcal{P}(A \cap B) = \mathcal{P}(A | B) \times \mathcal{P}(B)\) for \(B \in \mathcal{B}^0 = \{ \emptyset \in \mathcal{B} | \mathcal{P}(B) = 0 \}\). In the context of \(\mathcal{P}\), \(A_1 ... A_n\) are called (conditionally) independent given \(B \in \mathcal{B}^0\) iff \(\mathcal{P}(\cap A_i | B) = \mathcal{P}(A_i | B) \times ... \times \mathcal{P}(A_i | B)\) for each subsequence \(A_{i1} ... A_{in}\) of \(A_1 ... A_n\).

Our first task now will be to see how far we can relax all these requirements without giving up practical relevance for belief valuation or basic features of the probability calculus. In particular, we want to keep reasonable notions of conditioning and independency. On the other hand, our framework should be general enough to cover semi-qualitative Spohn-type formalisms like those introduced for modeling defaults and revision. In the following, we are going to propose a minimal list of assumptions to be satisfied by any structure claiming to support a reasonable form of generalized probabilistic reasoning. Before, let's recall some useful algebraic notions.

\((G, \ast)\) is called a semi-group iff \(G \neq \emptyset\) and \(\ast\) is an associative operation on \(G\). In \(G, (\ast, e)\), we call \(e\) neutral or the identity iff for all \(x \in G\), \(e \ast x = x \ast e = x\) and \(n\) absorptive iff for all \(x \in G\), \(x \ast n = n \ast x = n\). A semi-group \((G, \ast)\) is a group iff it has an identity \(e\) and for all \(x \in G\), there is \(y \in G\) s.t. \(x \ast y = y \ast x = e\). \((G, (\ast, <))\) is called an ordered group iff \((G, \ast)\) is a group, \(<\) is a linear ordering on \(G\) and for all \(x, y \in G\), \(x < y\) implies \(x \ast y < x \ast y\).

Definition 2.1 \((\mathcal{B}, \mathcal{R}, \mathcal{V})\) is a quasi-measure space with event algebra \(\mathcal{B} = (\mathcal{B}, \cup, \cap, -, 0, 1)\), valuation algebra \(\mathcal{V} = (\mathcal{V}, \#, \ast, \cdot, 0, 1, \leq)\) (generalizing \((\mathbb{R}^+, +, \cdot, 0, 1, \leq)\)) and quasi-measure \(\mathcal{R} : \mathcal{B} \to \mathcal{V}\) iff the postulates I.0 - I.9 hold.

I.0 Boolean structure:
\(\mathcal{B} = (\mathcal{B}, \cup, \cap, -, 0, 1)\) is a boolean algebra.

The usual propositional connectives are fundamental reasoning tools. Infinitary closures are quite exotic.

I.1 Value function:
\(\mathcal{R} : \mathcal{B} \to \mathcal{V}\) is a function with \(\mathcal{R}(0) = n\), \(\mathcal{R}(1) = e\) and \(\mathcal{R}(A \cup B) = \mathcal{R}(A) \# \mathcal{R}(B)\) for \(\cap\)-disjoint \(A, B\).

By measure intuitions. If \(A \cap B = 0\), \(\mathcal{R}(A \cup B)\) should only depend on \(\mathcal{R}(A)\) and \(\mathcal{R}(B)\). To grasp infinitesimal probabilities, we have to give up (full) \(\sigma\)-additivity because of missing upper bounds in nonstandard analysis.

I.2 Additive structure:
\((\mathcal{V}, \#)\) is a commutative semi-group with identity \(n\).

Induced by the boolean properties of \(\cup\) and \(\emptyset\) together with the I.1 characterization of \(\mathcal{R}\). Strictly speaking, this requirement only constrains the possible values of \(\mathcal{R}\), but we are going to extrapolate these algebraic properties to the whole structure.

I.3 Multiplicative structure:
\((\mathcal{V}, \ast)\) is a commutative semi-group with identity \(e\) and absorptive \(n\).

Based on the boolean properties of \(\cap, 0, 1\) and the desire to set \(\mathcal{R}(A \cap B) = \mathcal{R}(A) \ast \mathcal{R}(B)\) for intuitively inde-
General Belief Measures 577

pendent events A, B (\(\ast\) is meant to play the role of x), in particular for A, B \(\in\{0, 1\}\).

I.4 Distributivity :
\[ \forall v, v', w \ (w \ast (v \# v') = (w \ast v) \# (w \ast v')). \]
Backed by the distributivity of \(\cap\) over \(\cup\) and our considerations above.

I.5 Linearity :
\( (V, \ast) \) is a linear ordering.

For decision-theoretic reasons and formal convenience. If necessary, we can model partiality by sets of valuations.

I.6 Additive monotony :
\[ \forall v, v', w \ (v \leq v' \rightarrow w \leq v' \# w). \]
Suggested by the monotonic character of measures. To account for ranking measures, we need the weak version \(\leq\).

I.7 Multiplicative monotony :
\[ \forall v, v', w \ (v \leq v' \rightarrow w \leq w \ast v' \# w). \]
Conditioning on an independent, non-empty event should not affect \(\leq\)-relationships.

I.8 Additive accessibility :
\[ \forall v, v' \ (v \leq v' \rightarrow \exists w \ v \# w = v' \) \]
It should be possible to attribute an arbitrary lower value to a proper subvent (valuation freedom). Then, the principle follows from I.1.

I.9 Multiplicative accessibility :
\[ \forall v, v' \ (v \leq v' \rightarrow \exists w \ v' \# w = v) \]
This condition is necessary if we accept valuation freedom and want a suitable conditional version of \(I\).

To get a better understanding of these structures, we now consider some interesting easy consequences, e.g. the fact that \(\leq\) or \(\triangleright\) can be defined from \(\ast\).

Theorem 2.1 Assuming I.0 - I.9, we have,
1. \( \forall v \ n \leq v, \)
2. \( \forall v, v' \ (v \leq v' \& v' \neq n \rightarrow \exists w \ v' \# w = v) \)
3. \( \forall v, v' \ (v = v' \rightarrow n \leq n \rightarrow v = n) \)
4. \( \forall v, v' \ (v \leq v' \rightarrow \exists w \ leq w = v) \)
5. \( \forall v, v' \ (n \leq v, \varepsilon \leq v' \rightarrow n \ast v \ast v' \leq \varepsilon) \)
6. \( \forall A \in B \ n \leq \mathcal{R}(A) \leq \varepsilon \)

Proof :
1. By multiplicative accessibility, \( v \ast n \) would imply that there is a w s.t. \( v = n \ast w = n \), which is impossible. Using linearity, this gives us \( n \leq v \).
2. Because of multiplicative accessibility, we only have to show uniqueness. Suppose, we had \( v \leq v', v' \neq n \) and \( w \neq w' \) s.t. \( v' \ast w = v' \# w' = v \). W.l.o.g., we may assume \( w \leq w' \). But this would contradict multiplicative monotony.
3. Given \( n \leq v' \) and \( v' \ast n = n \). 2.1.2 allows us to infer from \( v' \ast v = n \) and \( v' \neq n \) that \( v = n \).
4. If there is a w s.t. \( v' \# v = w \) and \( w \leq n \), either \( v = n \leq v' \), and we are done, or \( n \leq v, v', w \) (by 2.1.2).
5. Multiplicative monotony and \( n \leq v, v' \leq n \) imply that \( n \leq v' \# v \ast v' \# n = n \).
6. By 2.1.2 and additive monotony, \( n \leq \mathcal{R}(A) = n \) by the restriction of \( I(V) \) to the interval \([n, e]\), \( n, e \leq v \) imply that \( n \leq \mathcal{R}(A) \leq n \). 2.1.2 allows us either \( n \leq v \), multiplicative monotony gives us either \( v' \# v \leq v \# v' \# n \leq n \), which is impossible, or \( n \leq v \). But then, \( v' \# w = n \) for \( w \leq n \).

We can justify this notation by the following result.

Theorem 2.2 If \( (B, \mathcal{R}, V) \) is a quasi-measure space, \( B \in B, \mathcal{R}(B) \neq n \) and \( \mathcal{R}(X) = \mathcal{R}(X \backslash B) \) for all \( X \in B \), then \( (B, \mathcal{R}(B), V) \) is also a quasi-measure space.

Proof : We only have to verify I.1. Obviously \( \mathcal{R}(0) = n \), \( \mathcal{R}(0 \uparrow B) = n \leq \mathcal{R}(0 \uparrow B) \). We still must prove is that \( \mathcal{R}(A \cup A' \uparrow B) = \mathcal{R}(A \uparrow B) \# \mathcal{R}(A' \uparrow B) \) for \(A \cap A' = \emptyset\). Obviously, this holds for \( \mathcal{R}(B) = n \). For \( \mathcal{R}(B) \neq n \), by 2.1.2, it is enough to show that \( \mathcal{R}(A \cup A') \uparrow B) = \mathcal{R}(A \uparrow B) \# \mathcal{R}(A' \uparrow B) \). But \( \mathcal{R}(A \cup A') \# \mathcal{R}(A' \uparrow B) \) is a linear ordering of \( \mathcal{R}(A \uparrow B) \# \mathcal{R}(A' \uparrow B) \) by distributivity.

Assuming \( n \neq e \), the relevant ordered multiplicative substructure of a given valuation algebra \( V \) is characterized by the restriction of \( (V, \ast, \leq) \) to the interval \([n, e]\), \( e \leq \leq \ast \), \( n \) being the converse of \( \ast \).

Theorem 2.3 \( I(V) \) is the positive half of an ordered commutative group.

Proof : Based on our definitions and 2.1.5, it is possible to show that \( I(V) \) is an ordered commutative semi-group with identity and minimum \( e \). Now, exploi-
One way to distinguish valuation algebras is to consider the interaction patterns they induce between \# and \&. So, for every valuation context, we would like to know how much we have to add to an abstract quantity to get a bigger one. This amounts to investigating for every valuation algebra \( \mathcal{V} = (\mathbb{V}, \# , \& , n, e, \prec) \) the corresponding additive magnitude ordering \( \ll \), defined by \( v \ll w \) iff \( w \# v = w \).

It is easy to see that \( \ll \leq \leq \) (by additive monotony) is both transitive (associativity of \#) and anti-symmetric (commutativity of \#), but not necessarily reflexive, e.g., if \( \mathcal{V} = (\mathbb{R}^+ , + , 0, 0, <) \). Notice however that \( n \ll n \).

Furthermore, it follows from additive monotony and accessibility that \( \ll \) is \( \leq \)-extendible, i.e., if \( v \leq x \ll x' \ll v' \), then we get \( v \ll v' \) (by additive accessibility, \( v' = x' \# w \) and therefore \( v \leq v' \ll v = x' \# w \ll x \# x' \# w' = x \# w' = v' \), i.e., \( v \ll v' \)).

In general, \( \ll \) is much coarser than \( \leq \). Setting \( S(w) = \{ v \mid v \ll w \} \), by extendibility an initial \( \leq \)-segment, there are basically three possibilities for each \( w \) : (1) \( S(w) = \{ n \} \), (2) \( \{ n \} \subseteq S(w) \subseteq \{ n, w \} \), (3) \( \{ n, w \} \subseteq S(w) \). In fact, we are going to show that it is enough to consider \( w = e \). In the following, let \( \text{SP}, \text{SH} \) and \( \text{SR} \) be the principles obtained by stating (1), (2) resp. (3) for \( w = e \).

**Theorem 2.4** If \( n \ll e \), for \( i = 1, 2, 3 \), (i) holds for \( e \) iff (i) holds for all \( x \neq n \).

**Proof:** The right-to-left direction is obvious.

1. Suppose \( S(e) = \{ n \} \) and \( n \neq x \ll x \# y \) for some \( y \neq n \). W.l.o.g. we may assume \( y \leq x \), otherwise, additive monotony would give us \( x \ll x \# x \ll x \# y = x \# y \) and we could replace \( y \) by \( x \). By multiplicative accessibility, there is \( v \neq n \) s.t. \( y = x \ll v \). Because \( e \ll e \# v \) and \( x \neq n \), we have \( x = x \ll e \ll e \# (e \# v) \ll x \ll (e \# v) = x \ll y \), which is impossible. Hence, (1) must hold for all \( x \neq n \).

2. Suppose \( x \neq n \) and \( e \ll v \ll e \# e \# v \) for some \( v \) with \( n \ll x \# e \). Then \( x = x = e = e \# (e \# v) \ll x \ll (e \# e) = x \ll x \) and \( n = x \ll n \ll x \ll e \ll x = x \). In other words, there is \( y \ll v \) s.t. \( n \ll y \ll x \) and \( x = x \ll (e \# v) \ll x \ll e \ll x \ll v \ll v \ll x \ll x \ll n \# x \ll n \# x \# n \). Consequently, (2) is verified for all \( x \neq n \).

3. From \( e = e \# e \), we get \( x = x = (e \# e) = x \# x \). That is, (3) is valid for all \( x \).

It seems natural to assume that \( \ll \) should define a linear hierarchy of additive magnitudes. This amounts to requiring that \( \ll \) should be a (transitive) modular ordering, i.e., for each \( w, v \ll v' \) should imply \( w \ll v' \) or \( v \ll w \).

**Definition 2.3** A valuation algebra \( \mathcal{V} = (\mathbb{V}, \# , \& , n, e, \ll) \) is called *hierarchical* iff \( \ll \) is modular.

Standard (probabilistic) and non-standard real valuation algebras - based on the ordered real number field \( (\mathbb{R}^+, +, 0, 1, <) \) resp. one of its elementary extensions including infinitesimals - will verify \( \text{SP} \). Here, \( \# \) is a very sensitive connective and adding non-zero quantities always makes a difference. In this context, \( \ll \) is just \( ([n]xV\{n\}) \) and therefore modular. That is, these valuation algebras are trivially hierarchical. The remaining two configuration types \( \text{SH} \) and \( \text{SR} \) reflect different qualitative-quantitative philosophies, which we are going to investigate now.

### 3 Ranking Measures

The principle \( \text{SR} \), which is equivalent to idempotence for \# characterizes what might be called pure semi-quantitative quasi-probabilistic accounts. Adopting this condition considerably simplifies our valuation structure. First, it allows us to identify \# with \( \max_n \), because for \( n \ll x \ll y \), we have \( y \leq x \# y \ll y \# y = \max_n(x \# y) \).

With 2.15 and the absorptive character of \( n \), this guarantees that \( \{ n, e \} \) is closed for \# and \&. Furthermore, 2.1.6 gives us \( R^*B \subseteq \{ n, e \} \). Therefore, w.l.o.g., we may concentrate on \( \text{SR} \)-type valuation algebras of the form \( \{ n, e \}, \max_n, \& , e, \ll \). From 2.3 we know that this structure can be completely described by the positive half of an ordered commutative group which has been topped by an absorptive \( n \). The converse of the resulting extended order is just our valuation ordering \( \ll \). Another interesting consequence is that we can now drop the disjointness condition in 1.1. Because we want to ensure that \( R(A) = n \) encodes real impossibility, unions of impossible events should not be allowed to embrace a possible one. All this suggests the following definition extending previous proposals in [Weydert 91, 92].

**Definition 3.1** \( (\mathfrak{B}, R, \mathcal{V}) \) is called a *ranking-measure space* iff

1. \( \mathfrak{B} = (\mathfrak{B}, \sqcup, \cap, \neg, 0, 1) \) is a boolean algebra.
2. \( \mathcal{V} = (\mathbb{V}, \#, \ll) \) (ranking algebra) is s.t.
   - \( (\mathbb{V}\{n\}, \#, \ll) \) is the positive half of an ordered commutative group with identity \( e \).
   - \( n \ll \) is minimal and absorptive for \( \ll \).
3. \( R : \mathfrak{B} \to R(\text{V ranking measure}) \) is a function satisfying
   - \( R(0) = n \), \( R(1) = e \) and
   - \( R(A \cup B) = \max_n(R(A), R(B)) \).
4. If \( A = \cup \{ P_i \mid i \in I \} \) and for all \( i \in I \), \( R(P_i) = n \), then
   - \( R(A) = n \) (coherence).

Observe that coherence is automatically satisfied for compact boolean algebras (i.e., where every covering of an event has a finite subcovering). What we still have to prove is that ranking measures really fit into our general quasi-measure framework.

**Theorem 3.1** Ranking-measure spaces are quasi-measure spaces and ranking algebras are hierarchical valuation algebras.
**Theorem 3.2** Let \((\mathcal{B}, \mathcal{R}, \mathcal{V})\) be a ranking measure space where \(\mathcal{B}\) is a set algebra on \(S\), the valuation ordering \(\prec\) is complete (i.e., \(\llcorner\) exists and \(n \prec \inf_{\mathcal{R}^{(P)}}(\mathcal{R}(P) \mid P \in B, n \prec \mathcal{R}(P))\). Then there is a canonical extension \(\mathcal{R}^*\) of \(\mathcal{R}\) to the powerset structure \(\mathcal{B}^* = (2^S, \cup, \cap, \cdot, 0, S)\) on \(S\), s.t. \((\mathcal{B}^*, \mathcal{R}^*, \mathcal{V})\) is a ranking measure space and \(\mathcal{R}^*\) is the maximal extension w.r.t. the \(\prec\)-induced pointwise ordering on ranking-measures from \(\mathcal{B}^*\) to \(\mathcal{V}\).

**Proof sketch:** \(\mathcal{R}^*\) defined by \(\mathcal{R}^*(A) = \inf_{\mathcal{R}^{(P)}}(\mathcal{R}(P) \mid A \subseteq P, P \in B)\) for \(A \subseteq S\) as desired.

In the literature, ranking-measure-like notions have been considered for different purposes. In [Weydert 91], we have provided a simple ranking measure semantics for default conditionals and comparative modalities (formalizing negligibility) together with a corresponding axiomatic characterization. In [Weydert 93], this approach has been extended to default quantifiers, i.e., genuine first-order contexts. A semantics for belief based on sets of ranking measures, different from what we discuss in this paper, can be found in [Weydert 94].

A very influential and sophisticated framework has been proposed by Spohn [88, 90]. His ordinal conditional functions (OCF) are intended to measure relative disbelief and model iterated belief changes. Their domains have the form \(2^S \setminus \{\emptyset\}\) and the valuation structure is \((\text{Ord}, +, \prec)\), with \text{Ord} being the class of ordinal numbers \(\{0, 1, 2, \ldots, \omega, \omega + 1, \ldots, \omega + \omega, \omega + \omega + 1, \ldots\}\), \(\omega\) the ordinal number addition and \(\prec\) the converse of the ordinal number ordering \(\langle \ldots \prec \omega \prec \ldots \prec 1 \prec 0 \rangle\). Because there is no absorptive value for \(\emptyset\) and, in particular, because \(+\) is not commutative \((\omega + 1 \neq \omega = 1 + \omega)\), his approach is not directly subsumed by ours. We could use another connective \(+^*\) based on polynomial representations \((\alpha = a_0 \omega^0 b_0 + \ldots + a_1 \omega^1 b_1 + \ldots a_0)\), which is commutative. But we still wouldn’t get accessibility because there is no \(x\) with \(1 +^* x = \omega\). Of course, for Spohn’s natural conditional functions (NCF), which take their values from \text{Nat}, this problem doesn’t arise. Here, the valuation structure basically becomes \(\mathcal{V}(Z)\). For many practical purposes, this might be enough. In [Goldszmidt, Pearl 92], for instance, this formalism has been used for modeling nonmonotonic inference with variable-strength defaults.

Possibility measures [Dubois, Prade 91] are another related and quite popular quasi-probabilistic tool for reasoning about uncertainty. Here, the domains are power sets and the valuation structure is of the form \([\{0, 1\}, \prec\). Because conditional possibility measures are defined through min, stronger \(\prec\)-like connectives are not considered. Spohn’s account and possibility theory both assume that their measures are fully additive, i.e., \(\mathcal{R}(A) = \sup_{\mathcal{R}(\{a\}) \mid a \in A}\). However, this shouldn’t be a general postulate. If we want to avoid an infinitary form of the lottery paradox, we will need semi-quantitative measures \(\mathcal{R}\) compatible with homogeneously small singletons, i.e., with \(v \in \mathcal{V}\) s.t. \(\mathcal{R}(\{a\}) = v \prec e\) for all \(a \subseteq S\).

### 4. CUMULATIVE MEASURES

Probabilistic and ranking measures are situated at the extremes of the quasi-measure range. Both fairly popular, they are characterized by different strengths and weaknesses. Probabilities are well suited for a fine-grained, e.g. decision-theoretic analysis in finite or continuous contexts where precise numbers are available. For their semi-qualitative counterparts, quite the opposite is true. They are most appropriate when we are interested in a simple, clear, rough, mainly number-free evaluation of plausibility, e.g. in default reasoning or (full) belief revision. When modeling plain belief, for instance, we want to avoid any confidence loss when conjoining two beliefs. Within the classical probabilistic framework, this can only be achieved by attributing measure zero to \(\neg \varphi\) if \(\varphi\) is believed. However, when we have to revise our beliefs based on new evidence supporting \(\neg \varphi\) we get into trouble because subsets of a null set cannot be differentiated probabilistically. For ranking measures, this problem doesn’t arise. Here, believing \(\varphi\) means associating a non-maximal, but not necessarily minimal ranking measure value to \(\neg \varphi\). This guarantees that beliefs are closed under conjunction. At the same time, Spohn-type conditionalization allows us to formulate reasonable revision procedures. Furthermore, we do no longer have to care about measurability constraints, the prohibition of uniform countable sample spaces and the representation of impossibility. On the other hand, everybody will agree that the representational and inferential power of ranking measures is rather restricted. In fact, it seems as if both formalisms were largely complementary. Consequently,
we should try to combine them in a powerful framework sharing the basic advantages of both.

The major issue is to extend the usual probabilistic framework far enough to allow differentiation among and especially conditioning on sets of vanishing probability. The idea is to introduce a hierarchy of mainly classical valuation contexts such that from the perspective of any given level, lower ranks are considered negligible or irrelevant, i.e. would get probability zero within a standard interpretation. Our strategy is to use a ranking algebra \( V = (V, \cdot, \prec) \) for describing the rough, global hierarchical structure and the real probabilistic valuation algebra \( R_+ = (IR_+, +, x, 0, 1, \prec) \), i.e. the standard positive real number algebra, for the more conventional, fine-grained local structure. The main task is now to merge these structures in a suitable way. This is done by the following definition.

**Definition 4.1** Let \( V^\circ = (V^\circ, \max, _\circ, 0^\circ, e^\circ, \varnothing^\circ) \) be a ranking and \( R_+ = (IR_+, +, x, 0, 1, \prec) \) be the real valuation algebra. Then we call \( \mathcal{H}(V^\circ, R_+) = (H(V^\circ, R_+), \#, \cdot, \n, e, \prec) \) a cumulative algebra with global structure \( V^\circ \) and local structure \( R_+ \) iff

- \( H(V^\circ, R_+) = \{(n, 0) \cup V^\circ\setminus\{n\} \times IR_+ \setminus\{0\} \), and for all \((a, r), (a', r') \in H(V^\circ, R_+)\).\n- \((a, r) \# (a', r') \) if \(a = a'\),
  - \(= (a', r') \) if \(a < 0 a', r \times r')\),
  - \((a, r) = (a', r') \) if \(a = a', r \times r')\),
- \(n = (n^\circ, 0) \) and \(e = (e^\circ, 1)\),
- \( (a, r) \prec (a', r') \) iff \(a < 0 a'\) or \(a = a'\) and \(a < a'\).

The definition of \( H(V^\circ, R_+) \) as a proper subset of \( V^\circ \times IR_+ \) excluding \((n^\circ, r) \) for \(r \neq 0\) and \((a, 0) \) for \(a \neq 0^\circ\) ensures that there are no zero-divisors \((x, y) \neq 0^\circ \) with \(x \times y = 0^\circ\), e.g. for \(x = (a, 0)\) and \(y = (n^\circ, r) \) or trivial violations of multiplicative monotony by \((a, 0) \neq 0^\circ \) or \((e, r) = (e^\circ, 1)\) and \((a, 0) \neq (e, r) = (a, 0) \neq (e, 1)\). For the multiplicative connective, the componentwise strategy is the obvious one to guarantee those properties which are required for encoding independency (cf part 2). The lexicographic ordering reflects our basic stratification philosophy. Finding an appropriate additive connective, however, is less straightforward. Within single ranks (\(a = a'\), we must adopt the pointwise approach \((a, r) \# (a', r') = (\max a, a', r + r')\), because distributivity and the intended isomorphism between \(\{e\} \times R_+ \setminus\{0\} \) and \(R_+ \setminus\{0\} \) (i.e. real valuation algebra at the top) allow the derivation of \((a, r) \# (a', r') = (a, 0) \# \{(a, 1) \# (e, r) \# \{(a, 1) \# (e, r) \# (e, r)\}) = (a, 1) \# \{(e, r) \# (e, r)\}) = (a, 0) \# (e, r + r') = (a, r + r')\). But additions involving different ranks cannot be handled that way. To satisfy additive accessibility, e.g. to pass from \((a, 1) \# (e, r)\) with a \(\prec e\) to the bigger value \((e, r)\) for \(r \leq 1\), we would have to introduce negative numbers on the right-hand-side, which simply doesn't make sense. In fact, it would again bring in oddities like \((a, 0) \# a \neq 0^\circ\). Notice that multiplicative accessibility already has brought us to consider \((a, r) \) with \(1 < r \) to get from \((e, 1/r) \) to \((a, 1)\). Taking all these precautions, our structures are well-behaved.

**Theorem 4.1** Cumulative algebras are hierarchical valuation algebras. They verify SH if the ranking algebra is non-trivial \((V^\circ \neq \{0^\circ, e^\circ\})\).

**Proof:** Let \( H(V^\circ, R_+) = (H(V^\circ, R_+), \#, \cdot, \n, e, \prec) \) be a cumulative algebra.

1. **\( H(V^\circ, R_+) \) is a valuation algebra.**

1.2 **Additive structure.** Commutativity and the neutrality of \(n = (n^\circ, 0)\) are obvious. To see that associativity holds, consider any \(((a, r) \# (a', r')) \# (a'', r'')\) and \((a, r) \# ((a', r') \# (a'', r''))\).

1.3 **Multiplicative structure.** Obvious because of the corresponding properties for \((V^\circ, \cdot, 0^\circ)\) and \((R_+, \times, 1)\).

1.4 **Distribution.** If \(a = a''\), then \((a, r) \# ((a', r') \# (a'', r'')) = (a \circ a', r \times (r' + r'')) = (a \circ a', r \times r' + r \times r'') = (a \circ a'', r \times (r' + r'')) = (a \circ a'', r \times r' + r \times r'')\).

1.5 **Linearity.** Follows directly from that of \(\prec \circ \) and <.

1.6 **Additive monotony.** If \(a \circ b \prec 0 a'\), then \((a, r) \# (a', r') \) implies \((a, r) \# (b, s) = (b, s) \# (a', r') = (a', r') \# (b, s)\). The remaining possibilities are immediate.

1.7 **Multiplicative monotony.** Immediate from the multiplicative monotony of \(V^\circ\) and \(R_+\), and \(n = (n^\circ, 0)\).

1.8 **Additive accessibility.** If \((a, r) \# (a', r')\), then we have either \(a < 0 a'\) and \((a, r) \# (a', r') = (a', r')\), or \(a = a'\) and \(r < r'\) and the additive accessibility of \(R_+\) gives us \(r'\) with \((a, r) \# (a', r') = (a', r')\).

1.9 **Multiplicative accessibility.** By the corresponding feature of \(V^\circ\) and \(R_+\) and the definition of \(H(V^\circ, R_+)\).

2. **\( H(V^\circ, R_+) \) is hierarchical.** \((a, r) \# (a', r') \) iff \((a', r') \# (a, r) = (a', r') \# a < 0 a'\) or \((a, r) = (n^\circ, 0)\), which describes a modular relation.

3. **SH holds.** \((e^\circ, 1) \# (a, r) = (e^\circ, 1)\) for some \((a, r) \) with \((n^\circ, 0) \# (a, r) \prec (e^\circ, 1)\), which exists by \(V^\circ\) nontriviality.

The next definition now introduces our main concept.

**Definition 4.2** A quasi-measure whose valuation algebra is a cumulative algebra is called a **cumulative measure.**
Notice that real measures can be seen as special cumulative measures using the trivial ranking algebra, i.e. where \( \nu^o = (\{n^o, s^o\}, \ast^o, s^o) \). The main purpose of our cumulative constructions is to provide a general, quasi-probabilistic framework supporting something like shifting granularity. Locally, we are dealing with a classical numerical context, but globally, we adopt the ranking perspective. Conditioning allows us to pass between levels and, as we shall see, to model plain belief revision without being forced to give up all the useful classical probabilistic or measure-theoretic tools. The idea is to use the (top-ranked) local probabilities for decision-theoretic purposes and the global ranking structure for revision tasks. Of course, we have to expect that some features get lost. Completeness, for instance, can at best be guaranteed within bounded segments of ranks. That is, there might well be disjoint \( A_i \), e.g. with \( \mathcal{R}(A_i) = (a, i) \) for \( i \in \text{Nat} \), where \( \text{sup}_x(\mathcal{R}(A_i)) \) doesn't exist. On the other hand, we are no longer troubled by measurability constraints. Every cumulative (and therefore every probabilistic) measure on a set-algebra can be extended to a cumulative measure on the whole powerset.

**Theorem 4.2** Let \( (B, \mathcal{R}, H(\nu, \mathcal{R}^+)) \) be a cumulative measure space where \( B \) is a set algebra on \( S \) and \( \mathcal{R}^+ = (IR^+, +, x, 0, 1, <) \). Then there is an extension \( \mathcal{R}^* \) of \( \mathcal{R} \) to the powerset structure \( B^* = (2^S, \cup, \cap, \sim, \emptyset, S) \) on \( S \) and a ranking algebra \( \nu^* \) s.t. \( \mathcal{R}^* \) is a cumulative measure and \( \nu \) a substructure of \( \nu^* \).

This result is different from our corresponding theorem 3.2 for ranking measures insofar as here, in general, we don't have neither a canonical extension nor constant valuation algebras (i.e. \( \nu = \nu^* \)).

Another proposal to reconcile the advantages of probabilistic and ranking measures has been made by Boutilier [Boutilier 93]. Roughly speaking, the structures he considers could be interpreted as special, restricted instances of cumulative measure spaces. Basically, his approach has been designed for finite boolean algebras, but - to a certain extent - it can be generalized to infinite ones. The idea is to have a possibility ranking on atoms and a probability measure \( \mathcal{P}_r \) attached to each level \( r \) which vanishes for measurable propositions built up from atoms of other ranks. Given an event \( A \), we look for the highest rank \( r \) where its probability \( \mathcal{P}_r(A) \) achieves strictly positive. Boutilier doesn't present a fully integrated account with combined valuation scales, but in our terminology, \( A \) would get the value \( (r, \mathcal{P}_r(A)) \).

The approach, as it stands, has several shortcomings. First of all, it relies on possibility measures, which are not general enough for some purposes (e.g. for modeling infinite boolean algebras based on countably many small atoms). Secondly, his definition forces us to use well-founded rankings, which seems to be a quite artificial restriction. In fact, the only well-founded ranking algebras are \( \nu(Z) \) and the trivial one. But the main problem with Boutilier's theory is that it doesn't really fit into the quasi-measure framework. Because the local values are supplied by probabilistic valuations, they cannot exceed 1. Consequently, multiplicative accessibility will fail, i.e. there will be no smooth account of conditional measures and our valuation freedom will be severely restricted. For instance, we may encounter valued events, less "probable" than their complements, for which there can't be another event which is exactly two times "probable". So, it seems justified to say that Boutilier's formalism achieves only a partial integration of the quantitative and the qualitative viewpoint.

**5 BELIEF STRUCTURES**

Traditionally, epistemic states of ideal agents have often been represented or approximated by probabilistic or discrete ranking measures on world sets. Unfortunately, as mentioned before, these pure approaches exhibit serious weaknesses. The first one, for instance, fails either to represent (by admitting full beliefs with probabilities below 1) or to revise plain belief (trivialization by conditioning on null sets). It also runs into problems because of measurability constraints affecting the realization of countably infinite uniform belief spaces. The second one, on the other hand, is unsuitable for fine-grained representation and decision-taking tasks where it makes a difference whether one option is as or twice as plausible as another one (which cannot be expressed by ranking measures). In particular, there is no possibility to model statistical knowledge directly by Spohn-type degrees of belief. Basically, these problems call for refined valuation structures backing the necessary differentiations. Of course, this can be done in many different ways. Either indirectly and hidden by a sophisticated machinery [Bacchus et al. 92], or directly, by combining complementary uncertainty formalisms, as we do. In the following, we are going to sketch some ingredients of an epistemic framework based on cumulative measures, which allows a correct handling of our benchmarks.

First of all, we have to address the question about the nature of epistemic states. Here, we want to adopt a liberal attitude. Epistemic states should not be identified with any kind of measures. This would be an oversimplification contradicting our subjective experience of what constitutes real-world belief (inconsistency, partiality, vagueness, mixture of qualitative and quantitative contents, deductive incompleteness, ...). What we will assume, however, is that any given epistemic state \( s \) can be interpreted or evaluated so as to provide a cumulative measure \( \mathcal{R}s \) describing the corresponding official, surface belief structure, to be exploited for practical purposes like decision-tasks. This approach might be called the projection model of belief. The domain of \( \mathcal{R}s \) is formed by the propositions of our actual language (state) \( Ls \), which we assume to be boolean. The valuation algebra is assumed to be non-trivial. Backed by the pleasant properties of cumulative measures, we can now represent plain belief in the traditional way without having to care about the shortcomings of the probabilistic account. Let's express "within the epistemic state \( s \), \( A \) is plainly believed" by \( s \models B(A) \).
Plain belief:

\[ s \models B(A) \iff \mathcal{Rs}(A) = (\varepsilon^0, 1) \]

What's important is that this doesn't force us to stipulate \(\mathcal{Rs}(-A) = (\varepsilon^0, 0)\), i.e. the impossibility of \(-A\), because \((\varepsilon^0, 1) = \mathcal{Rs}(T) = \mathcal{Rs}(A \lor \neg A) = \mathcal{Rs}(A) \# \mathcal{Rs}(\neg A) = (\varepsilon^0, 1) \# (a, r) \) is possible and true for \((\varepsilon^0, 1) \lt (a, r) \lt (\varepsilon^0, 1)\). Plainly believing \(A\) is not the same as rejecting the possibility of \(-A\). Also, we can differentiate between beliefs with regard to their revisability or epistemic entrenchment by attributing either bigger or smaller measure values to their complements (cf. below). A basic characteristic of plain beliefs is their closure under conjunctions.

Conjunctive closure:

If \(\mathcal{Rs}(A), \mathcal{Rs}(A') = (\varepsilon^0, 1)\), then \(\mathcal{Rs}(A \& A') = (\varepsilon^0, 1)\)

This holds because \(\mathcal{Rs}(-A \lor \neg A') = \max_a \mathcal{Rs}(-A), \mathcal{Rs}(-A')) \lt (\varepsilon^0, 1)\). Of course, we may also adopt less certain attitudes and assign intermediate values. Indifference w.r.t. \(A\), for instance, amounts to state \(\mathcal{Rs}(\neg A) = \mathcal{Rs}(A) = (\varepsilon^0, 0.5)\). In first-order contexts, it is easy to imagine situations where the additive structure of classical valuation algebras becomes inappropriate. For instance, let \(\Psi(x)\) be an opaque number-theoretic property for which we know that there is exactly one solution, but where we cannot give any bounds. Then, it seems reasonable to adopt a symmetric attitude, which can easily be realized within our cumulative framework. Let's assume that \(\mathcal{Ls}\) is powerful enough to express \(\Psi(n)\), the \(n\) being natural number constants.

Infinitary uniformity:

\[ \mathcal{Rs}(\Psi(j)) = \mathcal{Rs}(\Psi(l)) = \ldots = (a, 1) \lt (\varepsilon^0, 1)\]

In addition, we need of course a possibly nondeterministic revision mechanism \(\mathcal{N}\) which, given an epistemic state \(s\) and a new item of information \(i\), defines a set of possible updates \(\mathcal{N}(i|s)\). In general, \(i\) will be a constraint for the cumulative measures \(\mathcal{Rs}'\) of admissible revision states \(s'\), e.g. \(i = [\mathcal{R}(A) = (\varepsilon^0, 1)]\). A minimal condition for this \(\mathcal{N}(i)\) can be borrowed from the classical paradigm.

Top-conditionality:

\(\mathcal{Rs}'(B) = \mathcal{Rs}(B \mid A)\) if \(\mathcal{Rs}(B \mid A)\) has the form \((\varepsilon^0, r)\)

We cannot require full conditionality because this would force us to accept \(\mathcal{Rs}'(-A) = \mathcal{Rs}(\neg A \mid A) = (\varepsilon^0, 0)\), precluding nontrivial updating with \(\neg A\). That is, we would be confronted against the problems of the traditional framework, we want to escape.

Acknowledgements

This work was supported by the German Ministry of Research and Technology under grant ITS 9102, project LOGO. Thanks to the anonymous referees, whose comments have helped to make this paper at least slightly more intelligible.

References

[Bacchus et al 92] F. Bacchus, A.J. Grove, J.Y. Halpern, D. Koller. From statistics to belief. In Proceedings of AAAI 92. Morgan Kaufmann, 1992.

[Boutilier 93] C. Boutilier. The probability of a possibility. Adding uncertainty to default rules. In Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann, 1993.

[Darwiche, Ginsberg 92] A.Y. Darwiche, M.L. Ginsberg. A symbolic generalization of probability theory. In Proceedings of AAAI 92, Morgan Kaufmann, 1992.

[Dubois, Prade 91] D. Dubois, H. Prade. Possibilistic logic, preferential models and related issues. In Proceedings of the 12th International Conference on Artificial Intelligence, Morgan Kaufmann, 1991.

[Gärdenfors, Makinson 88] P. Gärdenfors, D. Makinson. Revisions of Knowledge Systems using Epistemic Entrenchment. In Theoretical Aspects of Reasoning about Knowledge, M. Vardi (ed.), Morgan Kaufmann, 1988.

[Goldszmidt, Pearl 92] M. Goldszmidt, J. Pearl. Reasoning with qualitative probabilities can be tractable. In Proceedings of the Eight Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann, 1992.

[Makinson, Gärdenfors 91] D. Makinson, P. Gärdenfors. Relations between the logic of theory change and nonmonotonic logic. In The Logic of Theory Change, A. Fuhrmann, M. Morreau (eds.), Springer-Verlag, 1991.

[Nebel 92] B. Nebel. Syntax-based approaches to belief revision. In Belief Revision, P. Gärdenfors (ed.), Cambridge University Press, 1992.

[Pearl 88] J. Pearl. Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann 1988.

[Shafer 76] G. Shafer. A Mathematical Theory of Evidence, Princeton University Press, 1976.

[Spoohn 88] W. Spoohn. Ordinal conditional functions: A dynamic theory of epistemic states. In Causation in Decision, Belief Change and Statistics, W.L. Harper, B. Skyrms (eds.), Kluwer Academic Publishers, Dordrecht.

[Spoohn 90] W. Spoohn. A general non-probabilistic theory of inductive reasoning. In Proceedings of the Fourth Conference on Uncertainty in Artificial Intelligence, R.D. Shachter et al. (eds.), North Holland, Amsterdam, 1990.

[Weydert 91] E. Weydert. Qualitative magnitude reasoning. Towards a new syntax and semantics for default reasoning. In Nonmonotonic and Inductive Logics, Dix et al. (eds.), Springer-Verlag, 1991.

[Weydert 92] E. Weydert. Relevance and Revision. About generalizing syntax-based belief revision. In Proceedings of the Third European Workshop on Logics in AI, JELIA 92, Springer-Verlag, 1992.

[Weydert 93] E. Weydert. Default Quantifiers. In Working Notes of the First Dutch-German workshop on Nonmonotonic Reasoning Techniques and their Applications, W. Nejdl (ed.), Aachen 1993.

[Weydert 94] E. Weydert. Doxastic normality logic: A qualitative probabilistic modal framework for defaults and belief. In Logic; Action and Information, A. Fuhrmann, H. Rott (eds.), de Gruyter, Berlin - New York.